The NLC Design Group

C. Adolphsen, R. Aiello, R. Alley, R. Assmann, K.L. Bane, T. Barklow, V. Bharadwaj, J. Bogart, G.B. Bowden, M. Breidenbach, K.L. Brown, D.L. Burke, Y. Cai, G. Caryotakis, R.L. Cassel, P. Chen, S.L. Clark, J.E. Clendenin, C. Corvin, F.-J. Decker, A. Donaldson, R.A. Early, K.R. Eppley, S. Ecklund, J. Eichner, P. Emma, L. Eriksson, Z.D. Farkas, A.S. Fisher, C. Foundoulis, W.R. Fowkes, J. Frisch, R.W. Fuller, L. Genova, S. Gold, G. Gross, S. Hanna, S. Hartman, S.A. Heifets, L. Hendrickson, R.H. Helm, H.A. Hoag, J. Hodgson, J. Humphrey, R. Humphrey, J. Irwin, R.K. Jobe, R.M. Jones, L.P. Keller, K. Ko, R.F. Koontz, E. Kraft, P. Krejcik, A. Kulikov, T.L. Lavine, Z. Li, W. Linebarger, G.A. Loew, R.J. Loewen, T.W. Markiewicz, T. Maruyama, T.S. Mattison, B. McKee, R. Messner, R.H. Miller, M.G. Minty, W. Moshammer, M. Munro, C.D. Nantista, E.M. Nelson, W.R. Nelson, C.K. Ng, Y. Nosochkov, D. Palmer, R.B. Palmer, J.M. Paterson, C. Pearson, R.M. Phillips, N. Phinney, R. Pope, T.O. Raubenheimer,[1] J. Rifkin, S.H. Rokni, M.C. Ross, R.E. Ruland, R.D. Ruth, A. Saab, H. Schwarz, B. Scott, J.C. Sheppard, H. Shoaee, S. Smith, W.L. Spence, C.M. Spencer, J.E. Spencer, D. Sprehn, G. Stupakov, H. Tang, S.G. Tantawi, P. Tenenbaum, F. Tian, K.A. Thompson, J. Turner, T. Usher, A.E. Vlieks, D.R. Walz, J.W. Wang, A.W. Weidemann, D.H. Whittum, P.B. Wilson, Z. Wilson, M. Woodley, M. Woods, Y.T. Yan, A.D. Yeremian, F. Zimmermann; *Stanford Linear Accelerator Center, Stanford, California, USA*

A. Jackson, W.A. Barletta, J.M. Byrd, S. Chattopadhyay, J.N. Corlett, W.M. Fawley, M. Furman, E. Henestroza, R.A. Jacobsen, K.-J. Kim, H. Li, H. Murayama, L. Reginato, R.A. Rimmer, D. Robin, M. Ronan, A.M. Sessler, D. Vanecek, J.S. Wurtele, M. Xie, S.S. Yu, A.A. Zholents; *Lawrence Berkeley National Laboratory, Berkeley, California, USA*

L. Bertolini, K. Van Bibber, D. Clem, F. Deadrick, T. Houck, M. Perry, G.A. Westenskow; *Lawrence Livermore National Laboratory, Livermore, California, USA*

M. Akemoto, T. Higo, K. Higashi, K. Kubo, K. Oide, K. Yokoya; *KEK National Laboratory, Tsukuba, Japan*

L. Rinolfi; *CERN, Geneva, Switzerland*

J.A. Holt; *Fermilab National Laboratory, Batavia, Illinois, USA*

V. Telnov; *Budker Institute for Nuclear Physics, Novosibirsk, Russia*

T. Takahashi, T. Ohgaki; *Hiroshima University, Hiroshima, Japan*

J. Rosenzweig; *University of California, Los Angeles, Los Angeles, California, USA*

S. Lidia; *University of California, Davis, Davis, California, USA*

N.M. Kroll; *University of California, San Diego, San Diego, California, USA*

A.J. Dragt, R.L. Gluckstern; *University of Maryland, College Park, Maryland, USA*

S.R. Hertzbach; *University of Massachusetts, Amherst, Massachusetts, USA*

G. Giordano; *University of Milano, Milan, Italy*

R.E. Frey; *University of Oregon, Eugene, Oregon, USA*

D.D. Meyerhofer; *University of Rochester, Rochester, New York, USA*

[1] Editor-in-chief.

Acknowledgments

This design for the Next Linear Collider (NLC) relies heavily on the first linear collider—the Stanford Linear Collider (SLC)—and on all the people who proposed, designed, and commissioned this pioneering accelerator. Without their work, none of this would have been possible (or necessary). In addition, the design owes much to the extensive international research effort that is investigating the different technological paths to a future linear collider. Finally, the NLC Design Group is a list of people, authors and non-authors, who have contributed significantly to this design; our apology in advance to anyone whose name may have inadvertantly been omitted.

In the process of completing this "Zeroth-Order Design" for the NLC, we have held two internal reviews and, more recently, an international external review. All of these have been very important to the design and we thank all of those who assisted. In particular, we thank the members of the external review panel which consisted of: Gerry Dugan, Helen Edwards, Hans Frischholz, David Gurd, Tom Himel, Steve Holmes, Norbert Holtkamp, John Ives, Robert Jameson, Katsunobu Oide, Satoshi Ozaki, John Rees, Nobu Toge. A number of useful comments were made, some of which already have been incorporated into the design. Most members of the internal review committees are listed in the contributors list; additional useful suggestions were given by H. DeStaebler, J.T. Seeman, R.H. Siemann, U. Wienands, and M. Zisman, along with many other members of the LBL, LLNL, and SLAC communities.

Finally, we would like to thank Angie Seymour, Marge Bangali, and Mary Litynski for administrative support, Terry Anderson, Sylvia MacBride, and Jim Wahl for creating the illustrations, and Jamie Walker and Laurie Gennari for producing this document.

Contents

Preface

This "Zeroth-Order Design Report" (ZDR) for the Next Linear Collider (NLC) is being created at a time of both great opportunity and uncertainty in the future directions that will be taken by the world-wide community of high-energy physics. There is exciting news that the Large Hadron Collider project has been approved for construction at CERN, and the planned involvement by physicists and engineers from countries around the globe will make this the first accelerator to be designed and built by a truly world-wide collaboration. By contrast, the cancellation of the SSC has demonstrated the necessity of international collaboration on such large scientific projects. The community of scientists and engineers at work on the accelerator physics and technologies of high-energy electron-positron colliders has recognized this need, and has made concerted effort to coordinate research activities to optimize our combined understanding and knowledge. This ZDR is one further step in this process.

The first electron-positron linear collider, the Stanford Linear Collider (SLC), began operation in 1989 with the dual purpose to explore the particle physics of the Z^0 boson and to develop the accelerator physics needed for a future TeV-scale linear collider. The SLC program has proven to be quite successful on both counts. Experiences gained and lessons learned from this prototype collider are a firm foundation for the design and implementation of a next generation machine. Developments at laboratories around the world have led to several choices of technologies to efficiently accelerate beams of electrons and positrons to high energy, and major test facilities presently nearing completion will soon allow evaluation of complete systems of these acceleration techniques. Additional test facilities already, or soon will, provide demonstrations and experience with techniques to create and control the delicate beams required to achieve the high luminosities needed for particle physics at the TeV-scale.

This NLC ZDR has been completed in the above context as a feasibility study for a TeV-scale linear collider that incorporates a room-temperature accelerator powered by rf microwaves at 11.424 GHz—similar to that presently used in the SLC, but at four times the rf frequency. The purpose of this study is to examine the complete systems of such a collider, to understand how the parts fit together, and to make certain that every required piece has been included. The "design" presented here is not fully engineered in any sense, but to be assured that the NLC can be built, attention has been given to a number of critical components and issues that present special challenges. More engineering and development of a number of mechanical and electrical systems remain to be done, but the conclusion of this study is that indeed the NLC is technically feasible and can be expected to reach the performance levels required to perform research at the TeV energy scale.

It is important to recognize that the contents of this ZDR include the work of many people not acknowledged as authors in the subsections of the report. This ZDR is the result of many years of discussion and investigation with scientists and engineers from around the world. References have been given in the text, but it is not always possible to accurately identify the true source of many of the notions and ideas included in a work of this type. The authors of this report apologize in advance for omissions. Effort has been made to use technical definitions in this ZDR that conform as widely as possible to those used in the recently completed International Linear Collider Technical Review Committee Report (The TRC Report, edited by G. Loew, SLAC Report–471, 1996). The ideas and parameters that appear in this ZDR have evolved from those given in the TRC report. Even so, the TRC report is a valuable companion to this document.

1

Introduction

1.1 Goals for the Next Linear Collider

For the past 25 years accelerator facilities with colliding beams have been the forefront instruments used to study elementary particle physics at high energies (Figure 1-1). Both hadron-hadron and electron-positron colliders have been used to make important observations and discoveries. Direct observations of the $W^{\pm}$ and Z^{0} bosons at CERN and investigations of the top quark at Fermilab are examples of physics done at hadron colliders. Electron-positron colliders provide well-controlled and well-understood experimental environments in which new phenomena stand out and precise measurements can be made. The discoveries of the charm quark and τ lepton at SPEAR, discovery of the gluon and establishment of QCD at PETRA and PEP, and precision exploration of electroweak phenomena at the SLC and LEP are highlights of the results produced by experiments at electron-positron colliders.

The ability to study nature with these two different kinds of instruments has proven essential to the advancement of our understanding of particle physics. This will remain true as we seek answers to questions posed at the TeV energy scale:

- What is the top quark, and what are its interactions?
- Why is the symmetry of the electroweak interaction broken, and what is the origin of mass?
- Do Higgs particles exist? If so, how many, and what are their structures and interactions?
- Is the world supersymmetric, and if so, what is its structure, and is it part of a larger unification of nature?
- Are quarks, leptons, and gauge bosons fundamental particles, or are they more complex?
- Are there other new particles or interactions, and what might nature contain that we have not yet imagined?

The Large Hadron Collider (LHC) in Europe offers an entry into the TeV energy regime with significant opportunity for discovery of new phenomena. The planned participation in the design, construction, and utilization of this collider by nations around the world will make the LHC the first truly global facility for the study of particle physics. This will be an exciting and important step in the continuing evolution of our science.

The companion electron-positron collider at this next step in energy, a Next Linear Collider (NLC), will provide a complementary program of experiments with unique opportunities for both discovery and precision measurement. To understand the nature of physics at the TeV scale, to see how the new phenomena we expect to find there fit together with the known particles and interactions into a grander picture, both the LHC and an NLC will be required.

Studies of physics goals and requirements for the next-generation electron-positron collider began in 1987-88 in the United States [Ahn 1988, Snowmass 1988, Snowmass 1990], Europe [LaThuile 1987, DESY 1990], and Japan [JLCI 1989, JLCII 1990]. These regional studies have evolved into a series of internationally sponsored and organized workshops [Finland 1991, Hawaii 1993, Japan 1995] that continue to build an important consensus on the goals

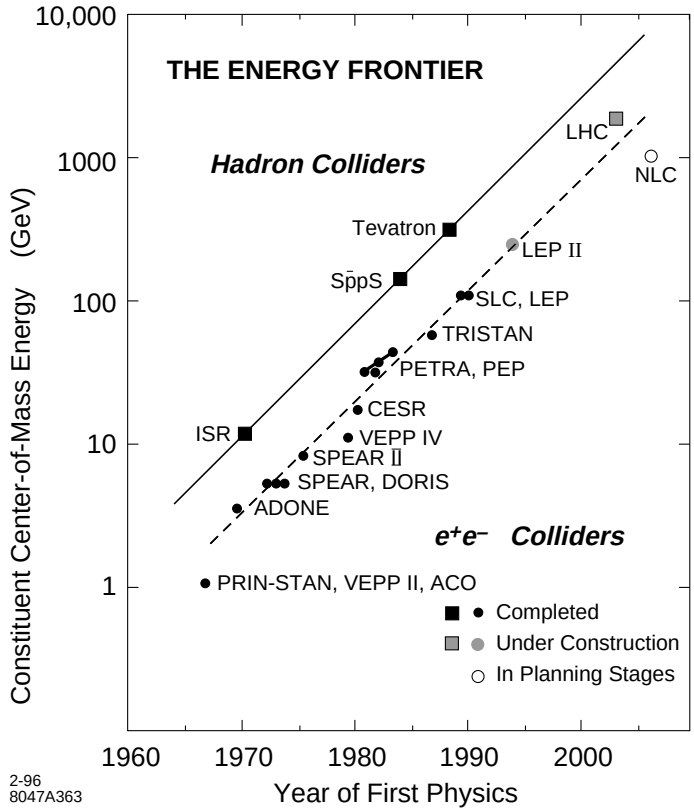


Figure 1-1. *The energy frontier of particle physics. The effective constituent energy of existing and planned colliders and the time of first physics results from each.*

and specifications of a Next Linear Collider. More recently, a series of workshops were held over the past year at locations throughout the United States. Working groups were established at a first meeting in Estes Park, Colorado to provide a framework for people to participate in the discussions of various topics in physics and experimentation at linear colliders. These groups continued to meet at subsequent workshops held at Fermilab, SLAC, and Brookhaven Laboratories. The results of those meetings is presented in a companion document "Physics and Technology of the Next Linear Collider" [NLC Physics] which also provides a summary of the NLC design.

A picture has emerged of a high-performance collider able to explore a broad range of center of mass energies from a few hundred GeV to a TeV and beyond (Figure 1-2). The goals of particle physics at TeV energies require luminosities $\sim 10^{34}\text{cm}^{-2}\text{sec}^{-1}$ and reliable technologies that can provide large integrated data samples. It is important that the beam energy spread remain well controlled and that backgrounds created by lost particles and radiation from the beams be maintained at low levels. This will assure that the clean experimental environment, historically offered by electron-positron colliders, remains intact. Beam polarization, an additional tool available at a linear collider that provides new and revealing views of particle physics, is also a requirement for any future collider.

In this first chapter, we introduce the accelerator physics and technologies of the Next Linear Collider, discuss its design choices and philosophies, and provide a brief status report on the R&D program that is being carried out in support of the NLC design effort. The last section of this chapter will provide an overview of the NLC design and then the design is presented in the following seventeen chapters. This is supplemented with four appendices that describe an alternate power source that might be used for the 1.5 TeV upgrade, a possible second interaction region, ground

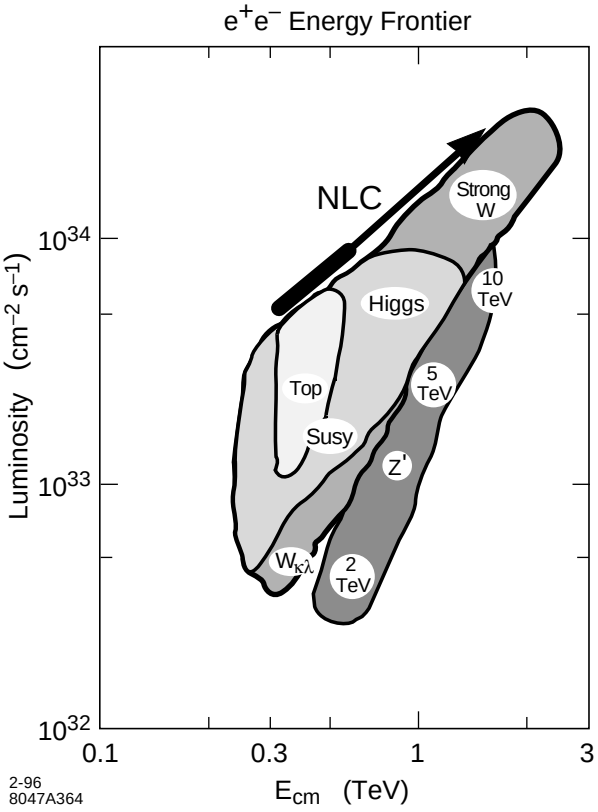


Figure 1-2. *Physics goals for a TeV-scale e^+e^- collider.*

motion studies, and a description of beam-based feedback which is utilized throughout the collider design. As stated, there is also a companion document "Physics and Technology of the Next Linear Collider", that contains results from a series of workshops studying the physics and experimental issues in the NLC as well as a shorter summary of the collider design.

1.2 Accelerator Design Choices

1.2.1 The Stanford Linear Collider

The Stanford Linear Collider (Figure 1-3) was conceived and built to accomplish two goals: to study particle physics at the 100-GeV energy scale and to develop the accelerator physics and technology necessary for the realization of future high-energy colliders. The SLC was completed in 1987 and provided a first look at the physics of the Z^0 in 1989. In time, the luminosity provided by this machine has grown steadily (Figure 1-4), and has allowed particle physicists to make unique and important studies of the Z^0 and its decays.

The design of the Next Linear Collider (NLC) presented in this document is intimately connected with experiences gained at the SLC. Our choices of technologies and philosophies of design have direct links to these experiences and

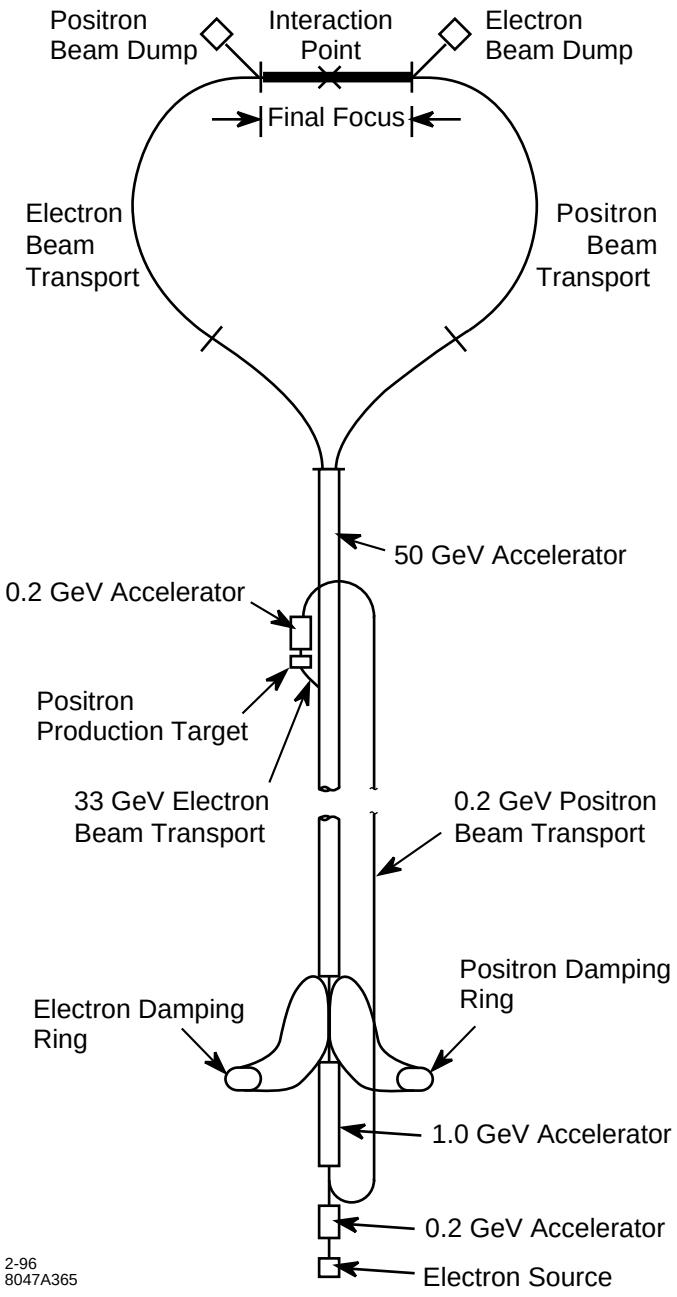


Figure 1-3. *The Stanford Linear Collider (SLC).*

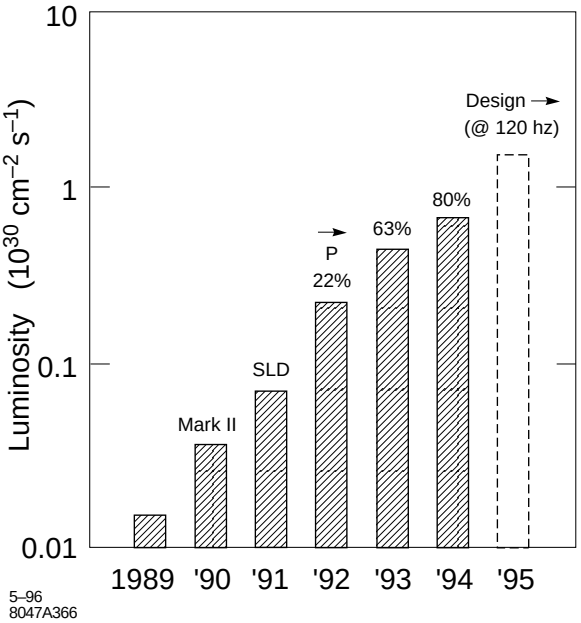


Figure 1-4. *Performance of the SLC from early commissioning. Polarization of the electron beam is also shown.*

considerable overlap with them. Lessons have been learned and techniques developed at the SLC that are relevant to the design and implementation of every part and system of the NLC:

- Injectors
  - Stabilized high-power electron sources
  - Polarized electrons
  - High-power targets and positron capture
- Damping Rings
  - Stabilized fast (50 ns) injection and extraction systems
  - Sub-ps phase synchronization with linac rf systems
- Linear Acceleration
  - Management of large rf systems
  - rf phase control
  - “Time-slot” compensation
  - Short-range longitudinal wake compensation
  - Multibunch beam loading compensation
- Linac Emittance Preservation
  - Beam-based alignment
  - LEM—energy/lattice matching

  - BNS damping
  - Coherent wakefield cancelation
  - Dispersion-free steering
- Final Focus Systems
  - Second-order chromatic optics and tuning
  - Precision diagnostics
  - Beam-beam control and tuning
- Experimentation
  - Theory and modeling of backgrounds
  - Vulnerability of detector technologies
  - Collimation—theory and implementation
- Systems Performance and Operation
  - Precision instrumentation—BPMs and wirescanners
  - Feedback theory and implementation
  - Importance of on-line modeling and analysis
  - Automated diagnostics and tuning
  - Mechanical stabilization of supports and components
  - Thermal stabilization of supports and components
  - Reliability
  - History monitoring (from seconds to years)

1.2.2 Future Linear Colliders

The basic components of any linear collider are those already incorporated into the SLC. The energy of such a future collider must be five to ten times that of the SLC, and a TeV-scale collider must be able to deliver luminosities that are several orders of magnitude greater than those achieved at the SLC. A generic collider complex is diagrammed in Figure 1-5. Trains of bunches of electrons and positrons are created, condensed in damping rings, accelerated to high energy, focused to small spots, and collided to produce a brightness given by

$$L = \frac{nN^2Hf}{4\pi\sigma_x^*\sigma_y^*} \quad , \tag{1.1}$$

where

n = number of bunches per train,
N = number of particles per bunch,
H = L enhancement,
f = machine repetition rate,

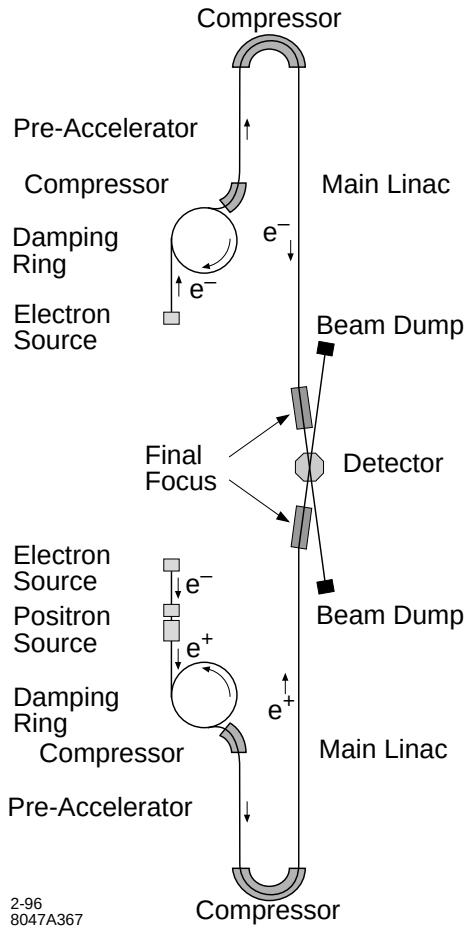


Figure 1-5. *Layout of a TeV-scale linear collider (not to scale).*

and σ_x^* and σ_y^* are the horizontal and vertical beam dimensions at the collision point. Equation 1.1 can be written as

$$L = \frac{1}{4\pi E}\frac{NH}{\sigma_x^*}\frac{P}{\sigma_y^*} \quad , \tag{1.2}$$

where P is the average power in each beam. The factor N/σ_x^* determines the number of beamstrahlung photons emitted during the beam-beam interaction, and since these photons can create backgrounds in experimental detectors, this factor is highly constrained. It is mainly the last ratio that can be addressed by accelerator technology; high luminosity corresponds to high beam power and/or small beam spots. These two parameters pose different, and in many cases contrary, challenges to the accelerator physicist, and several technologies that represent differing degrees of compromise between beam power and spot size are being developed. Table 1-1 summarizes the initial stage of the mainstream design choices.

Each of the technologies in Table 1-1 is being pursued by physicists and engineers at laboratories around the globe. This strong international effort is remarkably well coordinated through collaborations that combine to provide a set of test facilities to address each of the important aspects of the collider design and implementation. A summary of the facilities presently in operation or under construction is given in Table 1-2.

	Rf Freq (GHz)	Rf Grad (MV/m)	Total Length (km)	Beam Power (MW)	σ_y (nm)	Luminosity (10^{33} cm^{-2} s^{-1})
SuperC	1.3	25	30	8.2	19	6
S-Band	3.0	21	30	7.3	15	5
X-Band	11.4	50	16	4.8	5.5	6
2-Beam	30.0	80	9	2.7	7.5	5

Table 1-1. *Linear collider design parameters ($E_{cm} = 500$ GeV).*

Facility	Location	Goal	Operations
SLC	SLAC	Prototype Collider	1988–1998
ATF	KEK	Injector	1995
		Damping Ring	1996
TTF	DESY	SuperC Linac	1997
SBTF	DESY	S-band Linac	1996
NLCTA	SLAC	X-band Linac	1996
CTF	CERN	2-Beam Linac	1996
FFTB	SLAC	Final Focus Interaction Region	1994

Table 1-2. *Linear collider test facilities around the world.*

1.3 The Next Linear Collider

1.3.1 Technology Choice and Design Philosophy

The goal to reach 1 to 1.5-TeV center-of-mass energy with luminosities of 10^{34} cm^{-2} s^{-1} or better and our experiences with the SLC guide our choice of technologies for the NLC. We believe that the most natural match to these design goals is made with normal-conducting X-band (11.424 GHz) microwave components patterned after the S-band technology used in the SLC. A schematic of a section of the rf system for the NLC is shown in Figure 1-6. Our choice of technology has required the development of new rf klystrons and advanced pulse compression systems, but provides confidence that accelerating gradients of 50–100 MV/m can be achieved and used in the implementation of the collider. The technical risk of building a collider with new X-band technologies is perhaps greater than simply building a larger SLC at S-Band, but the goal to reach 1–1.5 TeV is substantially more assured, and capital costs to reach these energies will be lower.

The NLC is designed with a nominal center-of-mass energy of 1 TeV. It is envisaged to be built with an initial rf system able to drive the beams to 0.5-TeV center-of-mass energy, but with all the infrastructure and beam lines able to support 1 TeV. The rf system design incorporates the ability to replace and add modulators and klystrons without access to the accelerator beam line (dashed lines in Figure 1-6), so an unobtrusive, smooth, and adiabatic transition from 0.5 TeV to 1 TeV center-of-mass energy can be made with modest and expected improvements in X-band technology. This allows the collider to begin operation with the greatest of margins in cost and performance, and provides an excellent match

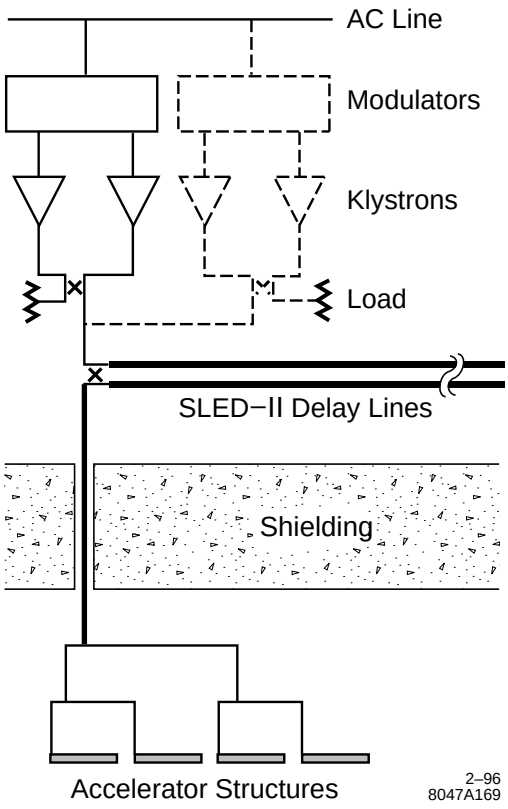


Figure 1-6. *Normal-conducting microwave rf system for the NLC.*

to the anticipated physics goals at the energy frontier (Figure 1-2). Our philosophy is akin to that taken previously in the construction of the SLAC linac which provided a 17-GeV electron beam at its inauguration, was improved to 35 GeV, and with continued advances in S-band technology, now provides 50-GeV electrons and positrons for the SLC.

The NLC design also incorporates multiple paths to further upgrade the center-of-mass energy to 1.5 TeV. The "trombone" shape of the collider layout would easily accommodate a straightforward albeit expensive increase in the length of the main accelerators without requiring extensive modification of the remainder of the complex. This final energy might also be accomplished by development of new, more efficient, X-band technologies; for example, gridded klystrons, cluster klystrons, or relativistic two-beam klystrons.

The highest level parameters of the NLC are listed in Table 1-3. At each of the nominal 0.5 and 1.0-TeV cms energies, three sets of parameters define the operating plane of the collider. The expected luminosity is constant over the operating plane, but is achieved with differing combinations of beam current and spot size. This provides a region in parameter space where the collider can be operated. Construction and operational tolerances for the various subsystems of the collider are set by the portion of the operating plane that is most difficult. For example, the more difficult parameters for the final focus are those of case (a), in which the beam divergences are large. In contrast, preserving the emittance of the beam in the linac is more difficult in case (c), in which the beam charge is highest and the bunch length longest. This design philosophy builds significant margin into the underlying parameters of the collider.

An important element in the design strategy of the NLC is the use of the beam to measure and correct errors in electrical and mechanical parameters of the accelerator. These techniques, many in extensive use at the SLC and FFTB, are able

	NLC-Ia	NLC-Ib	NLC-Ic	NLC-IIa	NLC-IIb	NLC-IIc
Nominal CMS Energy (TeV)		0.5			1.0	
Luminosity w/ IP dilutions (10^{33})	5.8	5.5	6.0	10.2	11.0	10.6
Repetition Rate (Hz)		180			120	
Bunch Charge (10^{10})	0.65	0.75	0.85	0.95	1.10	1.25
Bunches/RF Pulse		90			90	
Bunch Separation (ns)		1.4			1.4	
$\gamma\varepsilon_x$ at IP (10^{-8} m-rad)		400			400	
$\gamma\varepsilon_y$ at IP (10^{-8} m-rad)	7	9	11	9	11	13
β_x/β_y at IP (mm)	8/0.125	10/0.150	10/0.200	10/0.125	12/0.150	16/0.200
σ_x/σ_y at IP (nm)	264/5.1	294/6.3	294/7.8	231/4.4	250/5.1	284/6.5
σ_z at IP (μm)	100	125	150	125	150	150
Υ (Beamstrahlung Param.)	0.10	0.09	0.09	0.33	0.29	0.29
Pinch Enhancement	1.4	1.4	1.5	1.4	1.4	1.5
Beamstrahlung δ_B (%)	3.5	3.2	3.5	12.6	12.6	12.1
# Photons per e^-/e^+	0.97	1.02	1.16	1.65	1.77	1.74
Unloaded Gradient (MV/m)		50			85	
Effective Gradient (MV/m)	31.5	29.4	27.2	58.3	55.1	51.4
Active Linac Length (km)		8.15			8.90	
Min. Total Site Length (km)		23.8			30.5	
Max. Beam Energy (GeV)	267	250	232	529	500	468
Power/Beam (MW)	4.2	4.8	5.5	6.8	7.9	9.0
# of Klystrons		4528			9816	
Klystron Peak Pwr. (MV)		50			72	
Pulse Comp. Gain		3.6			3.6	
RF System Efficiency (%)		28			37	
Total AC Power (MW)		121			193	

Table 1-3. *Present IP and linac parameters of NLC designs.*

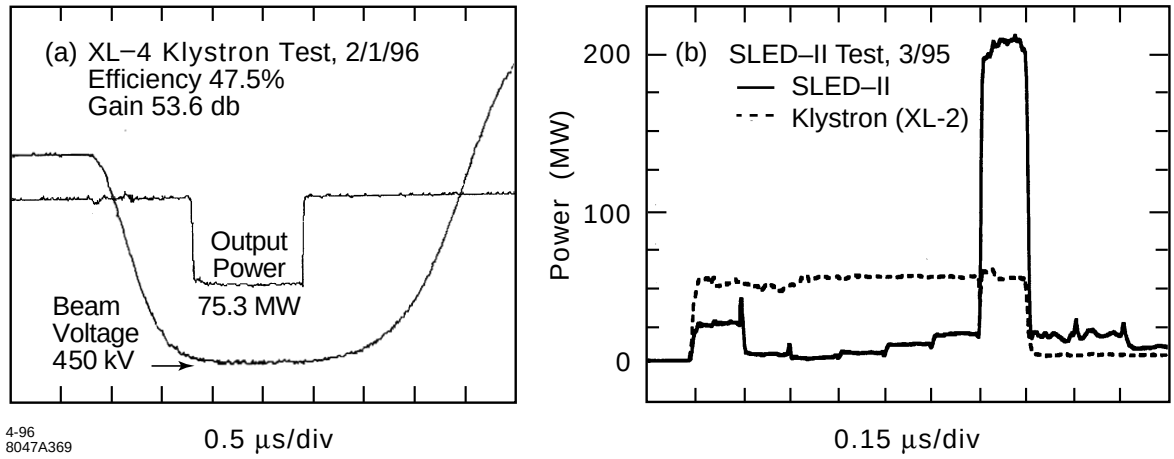


Figure 1-7. *Results of tests of X-band components: (a) high-power klystrons, and (b) rf pulse compression systems.*

to achieve far greater accuracy than is possible during fabrication and installation of components. For example, the use of optical matching and beam-based alignment algorithms considerably loosen tolerances required on magnet strengths and positioning. These procedures require accurate measurement of the properties of the beam and extensive online modeling and software. The existence of instrumentation suitable for these purposes is an important aspect of the readiness of technologies for the collider.

Additional performance overhead has been included in the designs of most subsystems of the NLC and errors that we anticipate occurring during machine tuning and operations have been taken into account. For example, the injector systems, *i.e.*, the particle sources, damping rings, and bunch compressors, are designed to provide 20% more charge than is indicated in Table 1-3 providing a substantial margin over the required performance. Similarly, fabrication and alignment tolerances for the main linac structures are specified without assuming benefit from certain global tuning methods such as coherent wakefield cancelation. These are powerful techniques in routine practice at the SLC, but our philosophy is to use them only to provide operational margin. We also recognize that the beam-based tuning cannot be done with perfect accuracy. For example, we have analyzed the tuning procedure for the final focus and estimated a 30% increase in the spot size at the IP due to errors that we anticipate will occur in measuring and correcting aberrations inherent in the optics. (This is included in Table 1-3.) This layered approach to specification of collider performance is an important part of our design philosophy.

1.3.2 Status Report on Technologies for the NLC

Progress in development of X-band rf components has been impressive in recent years. Prototype klystrons now produce 50 MW, 1.5 μs-long pulses with performance characteristics that are correctly modeled by computer codes. This exceeds the requirements of the initial 0.5-TeV stage of the NLC. In addition, the most recent prototype produces 75 MW, 1.1 μs-long pulses, as required for upgrading the NLC to 1-TeV center-of-mass energy. Tests of pulse-compression transformers have exceeded most goals of the NLC, and high-power rf windows and mode converters that allow high-efficiency transfer of power between components have been successfully tested. Examples of these results are shown in Figure 1-7.

The voltage gradient that can be used in a particle accelerator can be limited by the dark current created when electrons are drawn from the surfaces of the accelerator structures and captured on the accelerating rf wave. For a given rf

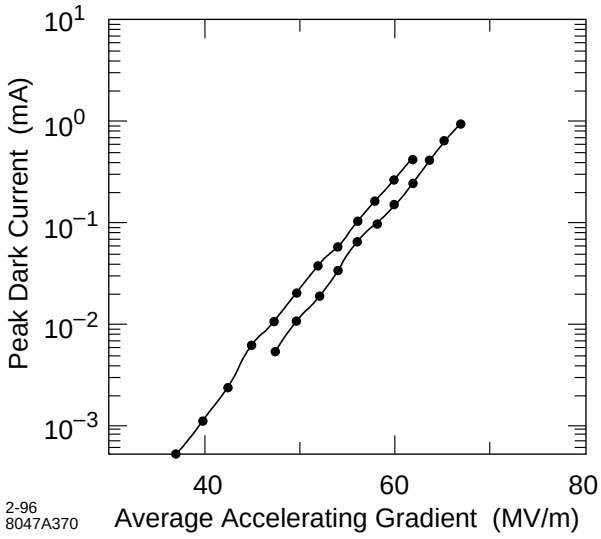


Figure 1-8. *Processing of X-band accelerator structures to high gradient.*

frequency, there is a well-defined gradient beyond which some electrons emitted at rest will be captured and accelerated to relativistic velocities. The capture gradient at S-band is about 16 MV/m, and scales to 64 MV/m at X-band. These are not actual limits to gradients that can be utilized in an accelerator since much of the charge is swept aside by the focusing quadrupoles of the machine lattice, but the dark current will grow rapidly above these values, and may adversely affect the primary beam or interfere with instrumentation needed for tuning. Gradients somewhat above the capture field are likely to be useful in practice, but the operational limits are not well known since no large-scale high-performance facility has been operated significantly above capture gradient. Expected thresholds of dark currents in S-band and X-band structures have been confirmed and it has been shown that (unloaded) gradients as high as 70 MV/m can be attained at X-band (Figure 1-8).

The electro-mechanical design of the main linac accelerator structures must not only produce the desired gradient, but must also minimize wakefields excited by the passage of the beam. The retarded electromagnetic fields left by each particle can disrupt the trajectories of particles that follow it through the accelerator. Many techniques to control the effects of the short distance intrabunch wakefields have been developed, tested, and put into use at the SLC. It will be necessary to also control long-range wakefields at the NLC in order to allow trains of closely spaced bunches to be accelerated on each rf pulse.

Structures that suppress wakefields by careful tuning of their response to the passage of the beam have been developed, and tests have been performed at a facility (ASSET) installed in the SLAC linac (Figure 1-9). Agreement with theoretical expectations is excellent and lends confidence to the design and manufacture of these structures. A more advanced design that further mitigates the long-range wakefields by coupling non-accelerating rf modes to external energy-absorbing materials has been completed, and a prototype of this new structure is being readied for testing in ASSET as well.

Work remains to be done on X-band rf technologies, but with prototype components now in hand, tests of completely integrated systems have begun. A fully engineered accelerator is under construction at SLAC that will allow optimization of rf systems and provide experience with beam operations at X-band frequencies. This test accelerator will be a 40-m-long section of six, 1.8 m X-band structures powered by 50–75 MW klystrons to an accelerating gradient of 50-85 MV/m. Commissioning of this facility has begun, and operations are expected to be under way by the end of this year (Table 1-2).

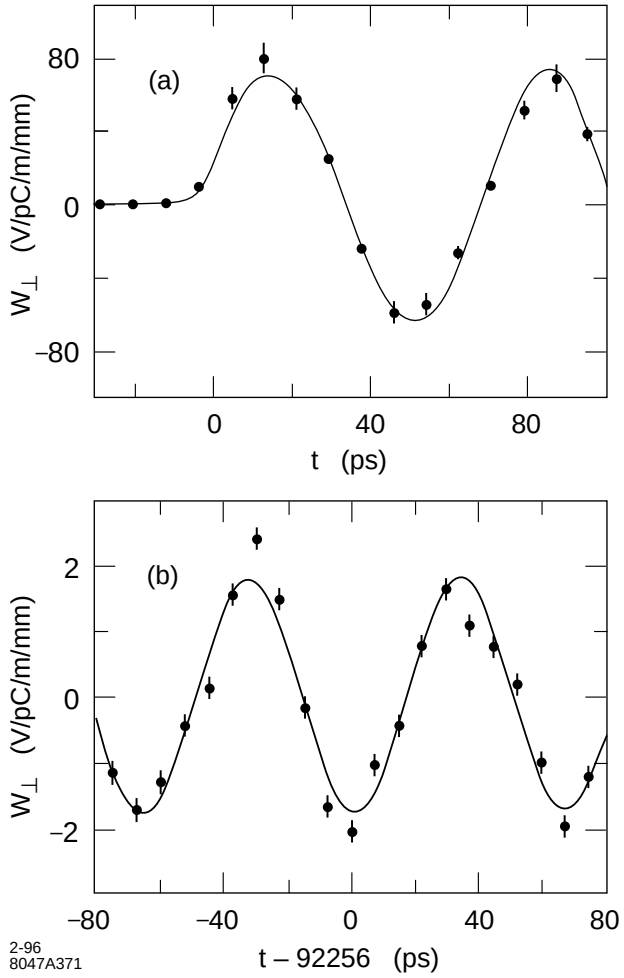


Figure 1-9. *Measured (circles) and predicted (line) transverse dipole wakefields in a 1.8-m-long X-band accelerator structure where (a) shows the wakefield immediately following a bunch and (b) shows the wakefield 92 ns later; note the large decrease in amplitude.*

The spot sizes that must be produced at the interaction point of the NLC represent significant extrapolations from those achieved at the SLC. It is important to demonstrate that it is possible to demagnify a beam by the large factor needed in the NLC. An experiment has been performed by the Final Focus Test Beam Collaboration to show that such large demagnifications can be achieved. The Final Focus Test Beam (FFTB) is a prototype beam line installed in a channel located at the end of the SLAC linac at zero degrees extraction angle. The FFTB lattice is designed to produce a focal point at which the beam height can be demagnified by a factor of 380 to reduce the SLC beam ($\gamma\varepsilon_y = 2\times10^{-6}$ m-rad) to a size smaller than 100 nm. The demagnification factor of the FFTB beam line is well in excess of that needed for the NLC.

The FFTB optics are chromatically corrected to third-order in the beam energy spread. (The SLC is corrected to second order.) All magnetic elements are mounted on precision stages that can be remotely positioned with step size of ≈ 0.3 micron, and beam-based alignment procedures were developed that successfully place these elements to within 5–15 microns of an ideal smooth trajectory. New state-of-the-art instruments were developed and used to measure the FFTB beam positions and spot sizes. Following a brief shake-down run in August of 1993, data were taken with the FFTB

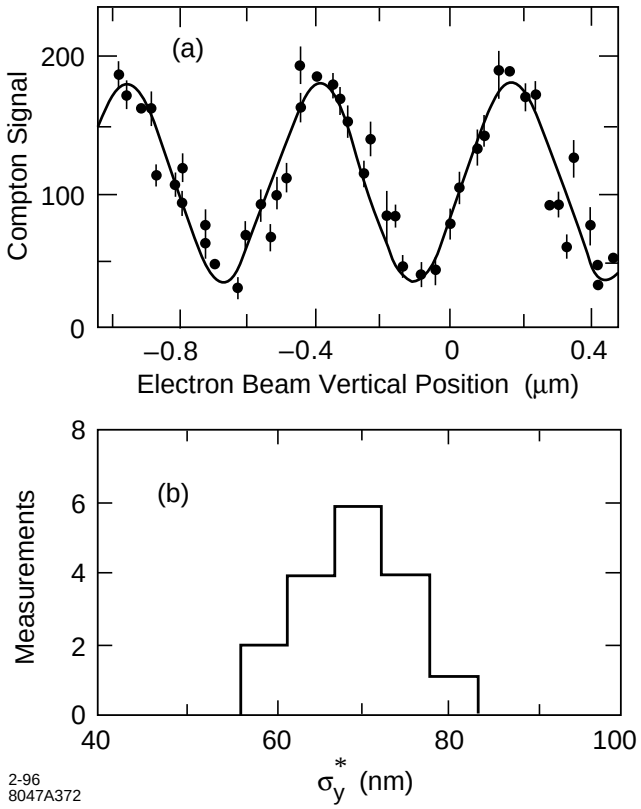


Figure 1-10. *Measurement of 70-nm beam spots with a laser-Compton beam size monitor in the FFTB. (a) The rate of Compton scatters from a laser interference pattern used to determine the beam size, in this case 73 nm. (b) Repeatability of spot measurement over periods of several hours.*

during a three-week period in April and May of 1994. Beam demagnifications of 320 and spot sizes of 70 nm were controllably produced during this period. Measurement of these beams is shown in Figure 1-10. The design of the NLC final focus follows that of the FFTB, and the experiences gained from the FFTB are incorporated into the tuning strategies of the NLC.

Important advances have also been made in instrumentation required to measure and control properties of the beam. The SLC control system has evolved dramatically over the past years to include extensive online modeling and automation of data analysis and tuning procedures. Scheduled procedures use sets of wire scanners to make complete measurements of the beam phase space, and provide recorded histories of machine performance. Online data analysis packages are able to reconstruct fully coupled non-linear optical systems. Beam-based feedback and feedforward are in routine operation in the SLC with over 100 loops providing control of beam trajectories and energies. Beam position monitors have been developed for the FFTB that achieve pulse-to-pulse resolutions of 1 micron, and new position monitors have recently been installed that are able to measure beam motions of 100 nm. The FFTB focal point spot monitors have demonstrated techniques to measure beam sizes of 30–40 nm, and extrapolation of these techniques to sizes as small as 10 nm can be expected to be successful.

1.4 Outlook for the Next Linear Collider

As the SLC has systematically increased its luminosity, the accelerator physics and technologies of linear colliders has matured. Experiences and lessons learned from the task of making this first collider perform as an instrument for particle physics studies make a firm foundation on which to base the design and technology choices for the next linear collider. At the same time, the essential demonstrations of new collider technologies have either taken place or soon will be underway. The experimental program with the FFTB is providing the experience needed to evaluate limitations to designs of final focus and interaction regions. The ability to demagnify beams by the amount required for the NLC has already been achieved. Microwave rf power sources have exceeded requirements for the initial stage of the NLC and critical tests assure us that this technology can be expected to drive beams to center-of-mass energies of a TeV or more. Fully integrated test accelerators are presently under construction at CERN, DESY, KEK, and SLAC that will soon provide answers to questions of technical optimization and costs of the major components of a TeV-scale collider.

Given the great international interest and commitment to the goals of a TeV-scale high-performance e^+e^- collider, it is certain that its final design, construction, and utilization will be a global effort. It is important that the scientific community put into place foundations for such a collaboration. The international character of the linear collider project is already reflected in the collaborations at work on the accelerator physics and technology of linear colliders and in the process of international discussion and review of progress in the field [Loew 1995]. It is essential that we continue to build on this base of understanding and cooperation and make certain that all involved in this enterprise are full party to its final realization.

1.5 NLC Systems Overview

The Next Linear Collider consists of a set of subsystems—injectors, linacs, beam delivery, and interaction regions. These are responsible for creating intense and highly condensed beams of positrons and polarized electrons, accelerating them to high energy, focusing them to small spots, and colliding them in an environment that allows sensitive particle detectors to operate for physics. In this section, we introduce these various parts of the collider while pointing to the more detailed discussions in the following text. We will also describe the energy upgrade and other more global issues.

A schematic of the NLC is shown in Figure 1-11. The physical footprint of the collider complex is approximately 30 km in length and less than 1 km wide. This includes all beam transport lines in the injectors and linacs necessary to obtain 1 TeV center-of-mass energy, and all space need in the beam delivery sections to accommodate 1.5 TeV center-of-mass energy. To reach 1.5 TeV, however, it may be necessary to extend the “trombone” layout of the collider to provide additional length for the linac. This can be done without moving the injectors or damping rings.

Injectors

The injector systems prepare the beams for injection into the main linacs at a beam energy of 10 GeV. The injectors consist of the polarized electron and positron sources, the damping rings, which reduce the transverse emittances of the beams, and the bunch compressors, which reduce the bunch lengths. Because they must provide a stable and reliable platform for the rest of the linear collider, we have been careful to design the systems with substantial operating margin.

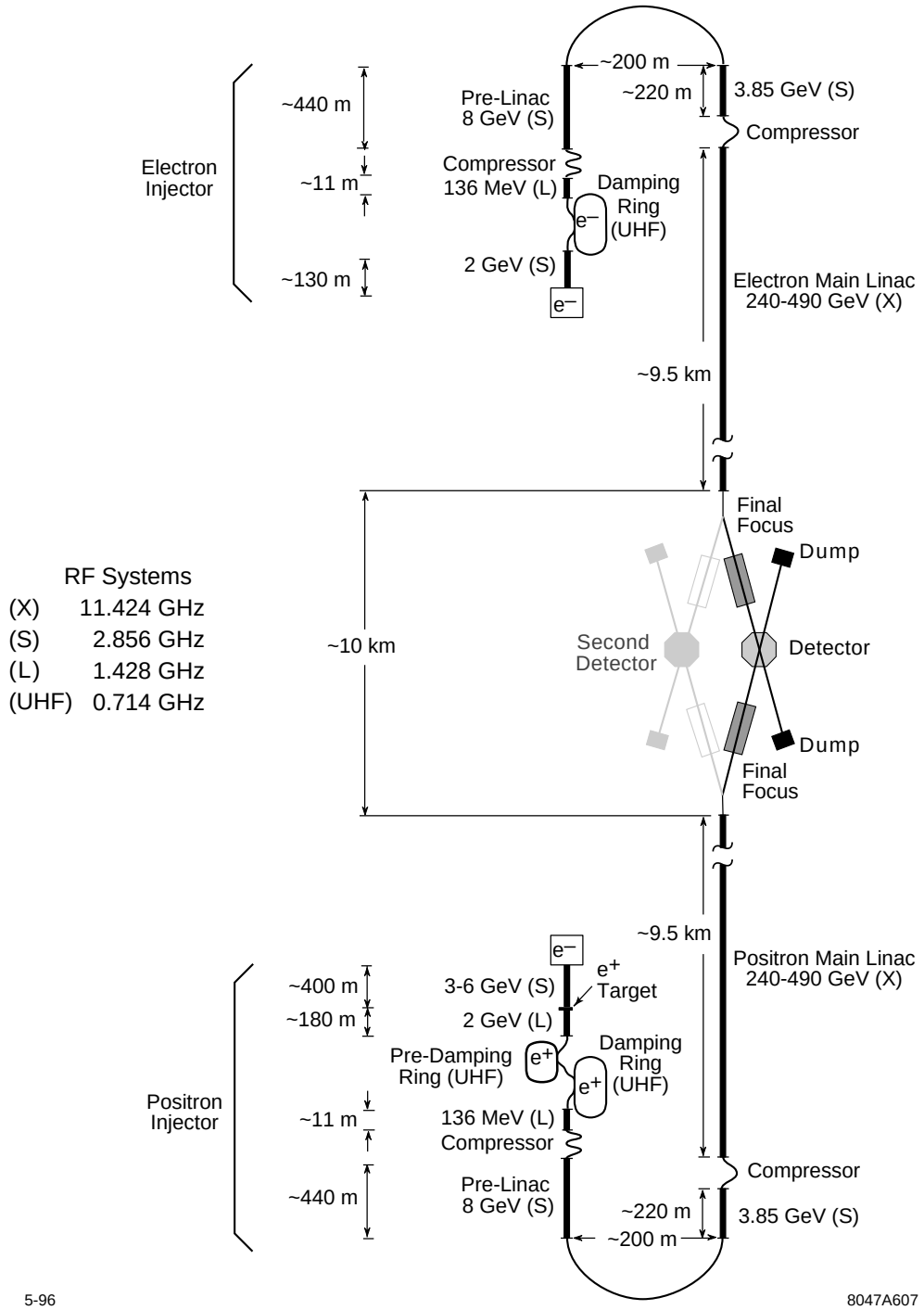


Figure 1-11. *Schematic layout of NLC systems (not to scale).*

The polarized electron source (Chapter 2) for the NLC is copied from the present SLC system which generates beams with $> 75\%$ polarization and operates very reliably. The NLC source includes a polarized laser, a photocathode electron gun, and a non-relativistic subharmonic bunching system. The new challenge for the NLC source is that it must produce trains of 90 bunches spaced by 1.4 ns at the machine repetition rate of 120–180 Hz. This is done by modulating a 126-ns laser pulse with a series of resonant rf Pockels cells. To get a relatively rectangular intensity profile, the modulation is performed using the first and third harmonic of the bunch spacing.

The bunched beam from the electron source is accelerated in an S-band (2.856 GHz) booster linac to 2 GeV and then injected into a damping ring (Chapter 4). The damping ring stores four trains of 90 bunches at once, extracting the oldest train as a new one is injected. It is designed to damp the transverse phase space of the beam from the incoming normalized emittances of $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad to $\gamma\varepsilon_x = 3 \times 10^{-6}$m-rad and $\gamma\varepsilon_y = 3 \times 10^{-8}$m-rad. These damped horizontal and vertical emittances are much smaller than those achieved in the SLC damping rings, and are comparable to those in the present generation of synchrotron light sources such as the Advanced Light Source at Lawrence Berkeley National Laboratory. To attain the small emittances, the ring is designed with a strong focusing lattice and uses beam-based techniques to achieve the required alignment of the ring magnets. The damping ring design incorporates lessons and techniques learned from operation of the SLC as well as experiences from the more modern synchrotron light sources. Finally, many of the components required in the rings have already been or soon will be designed and tested in other storage rings. This includes the wiggler magnets, which are included in the NLC design to reduce damping times, multibunch feedback and rf systems, and the vacuum system.

The technique for production of positrons that is planned for the NLC (Chapter 3) is also largely copied from that used at the SLC. Here, positrons are created in electromagnetic showers produced by targeting a 3–6 GeV electron bunch train onto a rotating $W_{75}Re_{25}$ target; the 3–6 GeV electron beam is produced with a conventional thermionic electron source and accelerated in a separate S-band linac. The large emittance positron beam is captured and accelerated to 2 GeV in a large-aperture L-band (1.428 GHz) linac. This lower frequency rf system has a smaller beam loading and leads to a good capture efficiency because of the increased longitudinal and transverse apertures.

At the 2-GeV point, the positron beam, with transverse emittances of 0.06 m-rad, is cycled through a large-aperture pre-damping ring (Chapter 4). The pre-damping ring damps the incoming positron emittances to $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad prior to injection into a main ring which is identical to that used for the polarized electron beam. Finally, the overall layout also includes transport lines that will allow the drive electron beam to bypass the positron production target and pre-damping ring to allow the study of e^-e^- or $\gamma\gamma$ collisions (with both electron or photon beams polarized).

Next, the lengths of the electron and positron bunches become too great in the damping rings for the beams to be successfully accelerated in the main X-band linacs, so they must be reduced (Chapter 5). This is done in two stages. A first bunch compressor, located immediately after the damping rings, reduces the bunch length from 4 mm to 500 μm to optimize injection into an S-band linac that accelerates the beams to 10 GeV. The bunch lengths are then further compressed to 100–150 μm in a compressor that rotates the longitudinal phase space by 360° and reverses the direction of travel; this arc allows for future upgrades of the main linac length and permits feedforward of the bunch charge and trajectory made in the damping rings and compressors.

Finally, the rf acceleration throughout the injector systems, *i.e.*, the bunch compressor rf, the 8-GeV prelinacs, and the 2-GeV booster linacs, is performed with relatively low frequency rf. In particular, the positron booster linac and the first bunch compressors operate at L-band (1.4 GHz) and the electron booster linac, the prelinacs, and the second bunch compressor operate at S-band (2.8 GHz). Although these systems are, in general, longer and more expensive than higher frequency accelerator systems, we feel that they are needed at the lower beam energy. In particular, we have designed the systems so that they have small beam loading and relatively loose alignment tolerances. This is important to provide the reliability and stability that is desired in the injector systems. Of course, such lightly-loaded systems are inefficient and could not be used to accelerate the beams to very high energy.

To minimize the long-range transverse wakefields, all of the accelerator structures are damped and detuned structures, similar to those developed for the X-band main linacs. In addition, the beam loading is primarily compensated using the ΔT (early injection) method which is discussed in Chapter 6. This has the advantage of canceling the bunch-to-bunch energy deviations locally within an accelerator structure thereby minimizing the transverse emittance growth due to dispersive and chromatic effects. Details of the multibunch beam loading compensation in the injector linacs as well as the low frequency accelerator structures and power sources is presented in Chapter 6.

Main Linacs

The main linacs of the NLC use normal-conducting traveling-wave copper structures operating at X-band (11.424 GHz). The choice of such a high frequency, relative to existing high-energy linacs, allows higher accelerating gradient, shorter linac length, and lower AC power consumption for a given beam energy. Considering the size, weight, cost, and availability of standard microwave components, we have chosen a frequency in the X-band for a design that is upgradeable from an initial 500-GeV center-of-mass energy to 1 TeV or more.

In the 500-GeV center-of-mass design, the required rf power is generated by 50 MW klystrons in 1.2 μs pulses which is then compressed to 0.24 μs, with higher peak power, by a passive rf transformer (SLED-II); this provides an unloaded acceleration gradient of 50 MeV/m. The upgrade to 1-TeV center-of-mass energy is accomplished by doubling the number of modulators and by replacing each 50-MW klystron with a pair of 75-MW klystrons. This attains an unloaded gradient of 85 MV/m. To achieve the full 1-TeV center-of-mass design, the total active length of linac must be increased slightly from 16,300 m to 17,700 m. This would be done by replacing spool pieces at the ends of the linac beam lines with accelerator structures. In addition, the planned upgrade includes improvements in the modulator and pulse compression systems to increase the rf system efficiency. The rf system for both the 500-GeV design and the 1-TeV upgrade is described in Chapter 8.

The rf system operates with a repetition frequency of 120–180 Hz. On each rf pulse, a train of 90 bunches is accelerated in each linac. While the number of particles in each of these bunches is small compared to that in the SLC, the total charge accelerated by each rf pulse is more than an order of magnitude greater. This multibunch design and the correspondingly larger fraction of energy transferred from the power sources to the beam is one of the important differences between the SLC and the NLC.

Extensive effort has been made on the design of an accelerator structure and rf system to ensure that both the beam loading can be compensated to the level required and the long-range transverse wakefield does not significantly increase the transverse emittances. The beam loading is compensated by optimizing the temporal profile of the rf pulse by shifting the phase of the rf from the klystrons. The long-range transverse wakefields have been reduced by tailoring the cell-to-cell dipole mode frequency to yield a Gaussian-like decay of the wakefield and by coupling each of the cells to four damping manifolds; this accelerator structure is referred to as a Damped-Detuned Structure (DDS). The technical designs of these rf components are described in Chapter 8 while many of the issues associated with the multibunch operation are summarized in Chapter 13.

The design and layouts of the linacs and the beam dynamics studies are presented in Chapter 7. The focusing lattice of the linac is designed to allow the center-of-mass energy to vary from 300 GeV to 1 TeV. Thus it does not have to be modified for the 1-TeV upgrade; the upgrade to 1.5-TeV will require changes. To verify the state of the beam, five diagnostic stations, located along each of the two main linacs, will include laser wire scanners to measure the transverse phase space, beam-based feedbacks to correct for centroid shifts of the bunch train, multibunch BPMs and high-frequency kickers to measure and correct bunch-to-bunch position errors, and magnetic chicanes to provide non-invasive energy and energy spread measurements.

The components in the linacs will have to be aligned to a few microns of a tuned reference orbit to prevent excessive emittance growth. This will be done using beam-based alignment techniques and high resolution beam position

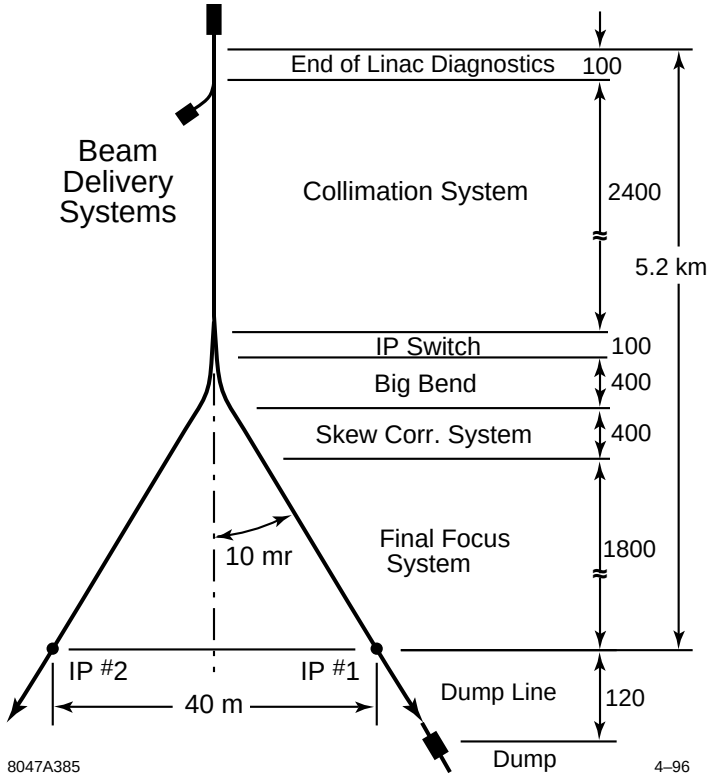


Figure 1-12. *Schematic layout of the beam delivery system.*

monitors. Studies of vibration and stability, based on measurements of the ground motion made at SLAC and described in Appendix C, show that the linac is not extremely sensitive to the motion. Fast feedback systems, similar to those used at the SLC and described in Appendix D, are required to stabilize the beam and the trajectories will need to be re-steered every 30 minutes; neither of these presents an operational limitation.

The Beam Delivery System

The beam delivery system consists of a collimation section, a switchyard, the final focus, the interaction region, and the beam extraction and dump. These are shown schematically in Figure 1-12. The overall lengths of the beam lines have been designed to allow for a final beam energy of 750 GeV.

After acceleration in the main linacs, the beams enter the collimation sections, described in Chapter 9, where particles at the extremities of the energy and transverse phase space are eliminated. The collimation is performed in both the horizontal and vertical planes at two betatron phases that differ by 90° to effectively cut the beam in both position and angle. This primary collimation is then followed by a secondary collimation in both planes and both phases to remove particles that are scattered by the edges of the primary collimators. The collimation regions have been designed to absorb 1% of the nominal beam power while surviving the impact of a full bunch train. The system consists of a series of spoilers and absorbers which have been optimized to reduce the transverse geometric and resistive-wall wakefields. Although the collimation sections are rather extensive, our experiences at the SLC and Final Focus Test Beam (FFTB) have proved the need to perform this function very carefully.

The collimation regions are followed by passive switching sections (Chapter 10) which are used to direct the beams to one of two possible interaction points (IPs). The beams are each deflected in the switchyard by 10 mr to produce a net 20-mr crossing angle at the IPs. This deflection introduces only little emittance growth, but greatly reduces detector backgrounds and provides the IP crossing angle required to collide the closely spaced bunches in each train.

The NLC final focus is described in Chapter 11. It follows the design of the FFTB at SLAC. The optics consist of a matching section with appropriate instrumentation to measure the incoming beam phase space, horizontal and vertical chromatic correction sections, a final transformer, a final doublet, and a diagnostic dump line for the exiting beam. The final focus system is chromatically corrected to third order. It demagnifies the transverse beam sizes by a factor of 80 horizontally and 300 vertically. A dedicated geometry adjustment section is included to facilitate the adiabatic upgrade from 350 GeV to 1.5 TeV center-of-mass without changing the IP position and requiring only minor transverse realignment of the beam line.

Although there is less demagnification than in the FFTB, the vertical spot size at the IP is roughly 15 times smaller than that attained in the FFTB. This imposes stringent tolerances which are achieved using beam-based tuning techniques and feedback systems. Measurement of the final spot size is a particular challenge, but can be performed with an advanced laser fringe monitor similar to that used in the FFTB and in-situ tuning of the spot sizes and luminosity can be done using measurements of the beam-beam deflections and techniques developed at the SLC. An extensive accounting of all the aberrations and the limitations of the tuning methods, including the different time scales required and the measurement accuracy, has been included in the luminosity calculations; these amount to a reduction in luminosity of 30–40%.

Interaction Region

Two interaction regions (IRs) are included in the layout of the NLC. The two important issues for the NLC IRs are the backgrounds and masking and the vibration stabilization of the final doublet; these are discussed in Chapter 12. The two IRs, of course, must share the available luminosity, but it will be possible to install two complementary experiments. As with all colliders, the interaction region will be a very crowded location. The final quadrupole magnets of the machine optics must be positioned as closely as possible to the collision point, and high-Z masking must be installed to protect elements of the experimental detector. The detector itself will require clear access to as much of the volume around the interaction point as possible. The design presented in this document includes quadrupoles two meters from the interaction point, and complete access for detector elements at polar angles greater than 150 mr. Calorimetric measurement of Bhabha scatters at smaller polar angles should also be possible, and are expected to provide precise determination of the luminosity-weighted center-of-mass energy spectrum.

The NLC is designed to collide electrons and positrons beams at a small (20 mr) crossing angle. This prevents the tightly spaced bunches of one beam from being disturbed as they approach the interaction point by bunches in the opposing beam that are leaving the interaction point. To avoid loss of luminosity due to this crossing angle, it is necessary to use a pair of small deflecting-mode rf cavities to “crab” the beams so that the bunches collide head-on. This is a new task not encountered at the SLC, but a system with reasonable specifications has been designed.

Global Considerations

Finally, the last five chapters discuss more global issues in the collider. As mentioned, the NLC will operate with long trains of bunches. Multibunch operation is one of the major differences between the NLC and the SLC. The issues associated with the multibunch operation are covered throughout the text but, because it is an important topic, we summarize these issues in Chapter 13.

Next, the control system and the instrumentation issues are outlined in Chapters 14 and 15. With the heavy utilization of beam-based alignment and tuning, the control system and the instrumentation design are very important for the NLC. The control system design uses many lessons learned from the SLC control system which has evolved substantially to meet the SLC requirements. Similarly, much of the required instrumentation is based on designs that have already been demonstrated. In the future, we will organize the description of the instrumentation into a unifying chapter but, at this time, the design and requirements are described throughout the text. Chapter 15 provides a short introduction to the philosophy of the instrumentation design and utilization.

The Machine Protection System (MPS) is described in the next chapter, Chapter 16. Because of the very high beam density and large beam power, the MPS is a difficult system and imposes significant constraints on the operation of the collider.

Finally, the NLC is a large and complex instrument and reliability is very important consideration. The SLC operates with an overall accelerator availability of roughly 80% and the NLC will be roughly ten times larger. Simple scaling of the SLC fault rates would suggest that the NLC would be effectively inoperable. Reliability considerations need to be addressed in the design from the onset and they are discussed in Chapter 17.

Lastly, the facilities and conventional systems for the NLC are described in Chapter 18. This includes the power and cooling distribution, the accelerator housing and the klystron gallery, as well as the campus and support facilities.

Appendices

In addition to the main body of the text, four appendices are included in this document. Appendices A and B present a possible rf system for a 1.5-TeV center-of-mass upgrade which is based on a Two-Beam Accelerator and a design for an interaction region to produce and study $\gamma\gamma$ collisions. These are followed by Appendices C and D which discusses ground motion theory and measurements and the implementation and limitations of beam-based feedback.

References

[Ahn 1988] C. Ahn, *et al.*, SLAC–Report–329, 1988.

[CERN 1987] Proceedings of the Workshop on Physics at Future Accelerators (La Thuile), CERN, Geneva, 1987.

[Snowmass 1988] Proceedings of the 1988 DPF Summer Study: Snowmass '88, High Energy Physics in the 1990s, E.L. Berger, ed., Snowmass, CO, 1988.

[JLCI 1989] Proceedings of the First Workshop on Japan Linear Collider (JLC I), S.Kawabata, ed., KEK, 1989.

[Snowmass 1990] Proceedings of the 1990 DPF Summer Study on High Energy Physics: Research Directions for the Decade, F. Gilman, ed., (Snowmass, CO, 1990).

[JLCII 1990] Proceedings of the Second Workshop on Japan Linear Collider (JLC II), (KEK, 1990).

[DESY 1990] Workshop on Electron-Positron Collisions at 500 GeV: The Physics Potential, (DESY, 1990).

[Finland 1991] Proceedings of the First International Workshop on Physics and Experiments with Linear Colliders, R. Orava, ed., (Saariselka, Finland, 1991).

[Hawaii 1993] Proceedings of the Second International Workshop on Physics and Experiments with Linear Colliders, F. Harris, et al., eds., (Waikoloa, Hawaii, 1993).

[Japan 1995] Proceedings of the Third International Workshop on Physics and Experiments with Linear Colliders, (Iwate, Japan, 1995).

[LaThuile 1987] Proceedings of the 1987 LaThuile Meeting: Results and Perspectives in Particle Physics, M. Greco, ed., (Gif-sur-Yvette, France, 1987).

[NLC Physics] "Physics and Technology of the Next Linear Collider: A Report Submitted to Snowmass 1996", SLAC Report–485; BNL–52–502; Fermilab–Pub–96/112; LBNL–Pub–5425; UCRL–ID–124160.

[Loew 1995] G. Loew, et al., International Linear Collider Technical Review Committee Report, SLAC Report–471 (1996).

2

Electron Source

Contents

2.1 Introduction

The electron injector for the Next Linear Collider (NLC) is based on the design of the injector for the SLC because many of the performance requirements for the NLC are similar to the SLC. The main difference is that the NLC injector has to deliver a train of bunches, therefore multibunch issues must be addressed.

The SLC polarized source which has operated so successfully since 1992, can be duplicated for the NLC with almost no changes to the polarized gun. We expect that the polarized source system will operate at better than 99% uptime efficiency based on the SLC experience. With ultra-high vacuum achieved in the SLC gun, the cathode lifetime has improved such that now it is greater than thousands of hours. Similar to the SLC, greater than 80% electron polarization will be achieved in the NLC. Some integrated experiments need to be performed to demonstrate that the polarized cathodes and cathode handling techniques available today can be applied to the cathode for the NLC polarized gun to achieve the required current for the NLC at greater than 80% polarization (Appendix 2.B).

There will be two electron injectors on the NLC: one for the collision and another for the positron-production electrons. The collision electron injector will have a polarized electron gun while the electron injector for positron production will have a thermionic gun. The injector for the collision electrons goes up to 2 GeV and has a lattice to match the beam into the linac-to-ring transition. The drive linac for the positron source goes up to 3.11 GeV in the first phase and 6.22 GeV in the second phase, and has the necessary optics to transport the beam to one of the two parallel beam lines for the positron production systems.

The low-energy transport portion of both electron injectors will be identical to provide operational flexibility and the flexibility of exchanging the thermionic electron gun for a polarized gun for future upgrades where two polarized electron sources are needed for $\gamma - \gamma$ collisions.

Figure 2-1 shows the injector for the collision electrons from the gun up to the linac to ring transition point. The bulk of this chapter addresses the collision electron injector beam line but Section 2.6 addresses the positron drive injector in as far as it is different from the collision electron injector.

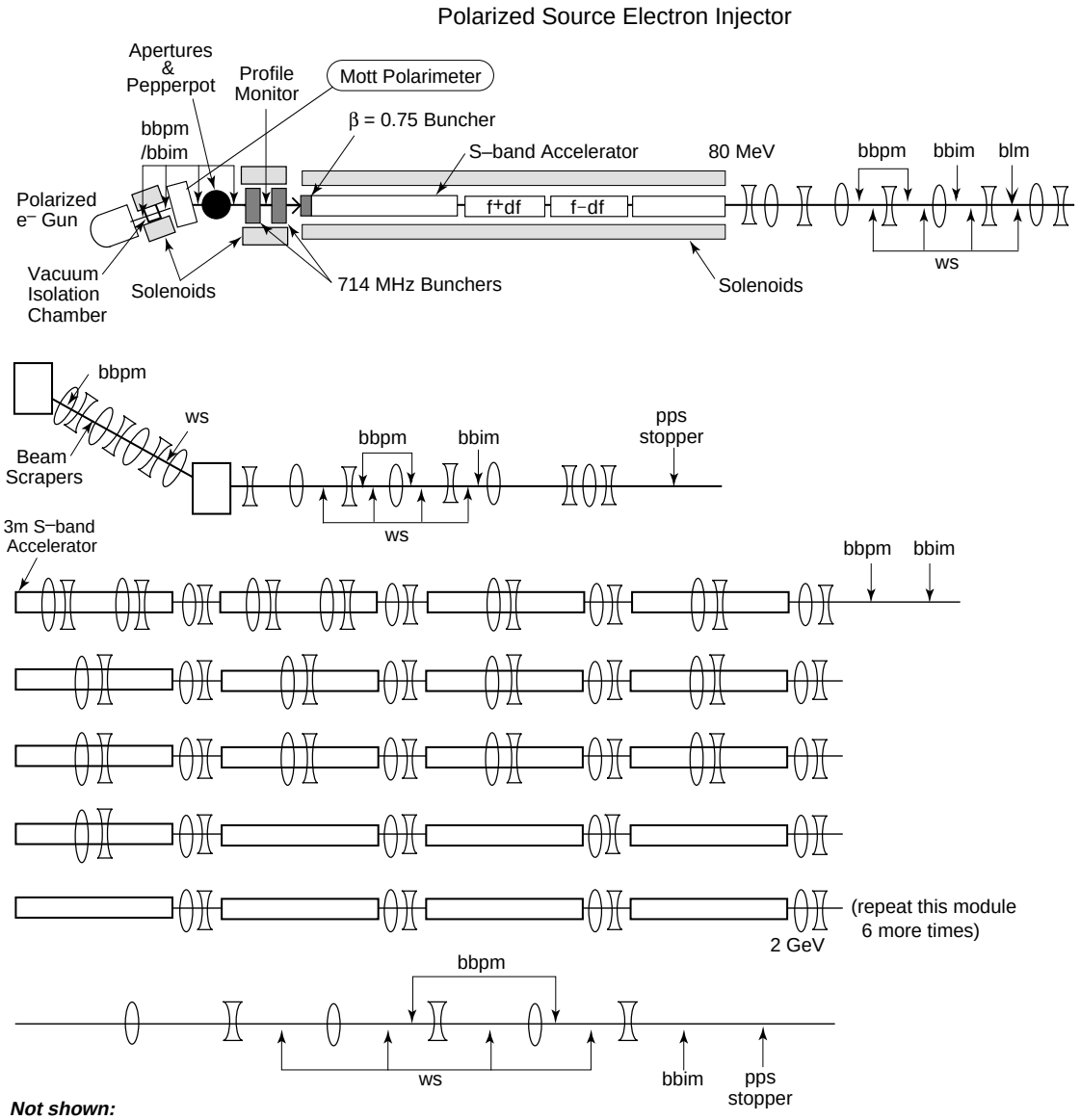
The baseline approach for the NLC injector will be a conventional system with a DC polarized gun and bunching system. The conventional subharmonically bunching injector approach [James 1981] is a proven, mature technology, used on injectors around the world, including on the SLC. The polarized electron gun and laser system are similar to the one which has been very successfully and reliably used on the SLC.

Some differences in this system are necessary to provide for the long bunch train. The laser system will produce a 126-ns-long train of bunches which are nearly rectangular in shape with a full-width-half-maximum value of 700 ps, and 1.4-ns apart. The beam intensity at 80 MeV needs to be about $1.9 \times 10^{10} e^-$ in 22 ps for each bunch, an average pulse train current of 1 A. The rms intensity jitter needs to be within 0.5% integrated over the entire train.

The polarized source rf gun injector approach should be studied in parallel (Appendix 2.A) to take advantage of the extremely low emittances achievable with rf guns, but at this point the technology of producing polarized electrons with rf guns has yet to be developed and is not the baseline approach for NLC.

The injector described in this section is shown in Figure 2-1. Our goal is to design an injector beam line such that it can produce the required beam parameters for both NLC-I and NLC-II as shown in Table 2-1. It has a 120-kV DC polarized electron gun which needs to produce $2.8 \times 10^{10} e^-$ in each bunch for NLC-II. The gun will house a 3-cm^2 cathode. All the current SLC polarized guns have a 3-cm^2 cathode, though the original gun with a 2-cm^2 cathode has been used for SLC operation.

A vacuum isolation chamber and a 20° bend protect the polarized electron gun from the downstream environment. The 20° bend also allows us to switch the beam into a Mott polarimeter station for occasional source polarization



Not shown:

Long ion chambers to detect beam loss will be located along the entire beam line from 40 MeV on up in energy.

12 beam loss monitors will be used to detect losses from the gun to the pepperpot following the first bend.

Train beam position monitors will be at each quadrupole, except for those wrapping around the accelerator.

A bbpm and bbim every 4 girders in the linac.

A bbpm, bbim, a wire scanner & a synchrotron streak camera station will be needed at high dispersion points following the linac.

Many dipole steering coils from the gun to the end of the injector.

Legend: ws = wire scanner
blm = bunch length monitor
bbim = bunch-to-bunch intensity monitor
bbpm = bunch-to-bunch beam position monitor

5–96
8047A125

Figure 2-1. *2-GeV polarized electron source injector.*

measurements. The train-to-train intensity jitter of the beam at the gun may be close to the allowed threshold of 0.5% due to the incoming laser intensity jitter on the photocathode. This will limit energy jitter in the linac to 0.25% as described in Chapter 1. An aperture 24-cm downstream of the 20° bend will be used to further reduce the electron beam intensity jitter of the bunched beam. Two 714-MHz standing-wave subharmonic bunchers and an S-band buncher-accelerator section compress the beam from the gun such that 83% of the charge is captured in 18° of S-band or 17.5 ps. PARMELA simulations were conducted from the gun through the second accelerator section, where the beam reaches 80 MeV $\pm$0.6% in energy. This energy spread is for a single bunch not including the beam-loading effect due to the long bunch train.

Beam-loading effects in the subharmonic bunchers are mitigated by using cavities with low R/Q to minimize the beam loading and inject the beam during the fill time of the cavity. Axial magnetic field focusing will be used on the first two accelerator sections (up to 80 MeV) where the beam is expected to have a larger energy spread than in the rest of the injector linac. At 80 MeV, there will be an isochronous and achromatic bend to connect the low-energy beam line to the main injector beam line. This bend allows us to construct another low-energy injector line in the future which may be needed either for more reliability in the polarized source or possibly for the first polarized source rf gun injector for the NLC. There is a set of wire scanners on both sides of the achromatic bend to measure emittance of the beam from the low-energy beam line and the emittance going into the injector linac after the bend. In addition, the chicane will have a set of scrapers at the high-dispersion point to clip away tails which are not in the main bunch and a wire scanner following the scrapers to measure the energy spread. From 80 MeV up to 2 GeV, the linac will use 3-m accelerator sections and Δt method for beam-loading compensation. Each module will consist of four 3-m accelerator sections powered by 2 S-band Klystrons similar to the SLAC 5045 klystrons except for the slightly longer 4-μs pulse width. Quadrupole doublets between each section and some wrapped around the first few sections will be used to control the transverse dimension of the beam.

There are only a couple of technological-breakthrough issues associated with the baseline injector chosen for the NLC. We need to demonstrate in an R&D program that we are able to produce cathodes with the 80% polarization and the up to 3.2-A average current needed for NLC-II. Currently available photocathodes with 80% polarization and 3-cm^2 emitting area are capable of producing only one fourth as much charge as is required for NLC-I with the (Appendix 2.B). It would be important to develop an 80% polarization cathode that will produce at least the required charge for NLC-1. If such a cathode is not developed in time then it is possible to increase the charge by doubling the area of the cathode with some compromise to the beam emittance at the gun to produce at least half the required charge for NLC-1. Other options for increasing the charge from the cathode could include development of the gun technology such that higher electric field would be tolerable at the cathode.

The beam-loading compensation both in the subharmonic bunchers and in the injector accelerator need to be simulated in detail, and physically realizable cavities need to be designed for the bunchers and the accelerators. These technologies will be demonstrated in the Next Linear Collider Test Accelerator (NLCTA) injector upgrade.

The development of a reliable and stable polarized-source-rf-gun injector deserves a concentrated parallel effort in the laboratory. If it could be demonstrated that the polarized-source rf gun system with the much lower beam emittance can be reliably and stably operated, then it will simplify the damping ring design or at least make it easier to operate. If the demonstration programs prove successful, the polarized-source rf gun could become the baseline injector for the NLC in which case we might also use an rf gun for the positron drive injector to make it similar in operation to the collision electron beam line.

Parameters	NLC-I			NLC-II			Overhead	SLC
	A	B	C	A	B	C	20%	achieved
N/bunch at IP (10^{10})	0.65	0.75	0.85	0.95	1.1	1.25	1.5	
N/bunch at Gun (10^{10})	1.2	1.4	1.6	1.75	2.0	2.3	2.8	8.8
Iave at gun (A)	1.4	1.7	1.9	2.0	2.4	2.7	3.2	
N/bunch at intensity jitter aperture (10^{10})	0.95	1.15	1.3	1.4	1.63	1.85	2.2	
N/bunch at 80 MeV (10^{10}] in 18 ps	0.83	0.97	1.1	1.2	1.4	1.6	1.96	
Iave at 80 MeV (A)	0.95	1.11	1.26	1.37	1.58	1.8	2.2	
N/bunch at damping ring (10^{10})	0.75	0.88	1.0	1.1	1.28	1.45	1.7	5.5

Parameters	NLC Required	Simulation	SLC no train
Bunch train durations (ns)	126	126	
Bunch separation (ns)	1.4	1.4	
PW at the gun FWHM (ps)	700	700	2000
PWedge after bunching (ps)	18	18	18
ε n,rms at 40 MeV (10^{-4} m-rad)	1.0	0.5	1 to 1.3
$\Delta E/E_{edge}$ at entr. to DR. (%)	$\pm$0.6	$\pm$0.6	$\pm$0.6 to 0.8
Train-to-train intensity jitter (%)	$<$0.5%	negligible w/ intensity jitter limiting aperture	0.8 at gun 0.5 after bunching

Table 2-1. *NLC polarized electron source parameters up to the damping ring.*

2.2 Polarized Electron Gun

2.2.1 Specifications

The principal requirements of the polarized electron gun are given in Table 2-2. To achieve $1.5 \times 10^{10}\ e^-$/bunch at the interaction point (IP), about $2.8 \times 10^{10}\ e^-$/bunch or a peak current of 4.5 A must be extracted from the cathode. The maximum current that can be extracted is established by the space-charge limit of the gun. The gun will be designed for a space charge limit of 13 A at the operating voltage. Assuming adequate laser energy, the ability to extract high peak currents at a given voltage depends on the internal charge limit of the cathode, which in turn depends primarily on the success in achieving a negative electron affinity surface. Achieving a successful negative electron affinity surface requires an ultrahigh vacuum system with total system pressure $\leq 10^{-12}$ Torr (excluding H_2).

Field emission associated with a DC-biased cathode can be destructive to a negative electron affinity surface. Thus the average dark current should be 25 nA or less. To achieve this goal, the electric field on the cathode electrode surface should be kept below 7 MeV/m at the operating voltage.

The intensity variations in the extracted electron pulse, including the jitter, are dominated by the properties of the laser pulse. On the other hand, the energy variations depend on the space-charge forces and the properties of the high-voltage power supply system for biasing the cathode. The energy variation within the electron pulse should be $\leq$1%, and the intensity variation $\leq$0.5% at the gun.

Gun vacuum	$\leq 10^{-12}$ Torr (excluding H_2)
Operating voltage	120 kV
Cathode area	3 cm^2
Extraction field	2.0 MV/m
Maximum field on cathode electrode	<7 MV/m
Space charge limit	13 A
Peak current at gun NLC I	2.5 A or $1.6 \times 10^{10} e^-$/bunch
NLC II	4.5 A or $2.8 \times 10^{10} e^-$/bunch
Number of bunches NLC I	90
NLC II	90
Bunch width at gun	700 ps FWHM
Bunch spacing	1.4 ns
Macropulse repetition rate NLC I	180 pps
NLC II	120 pps
Dark current	<25 nA DC
Intrapulse energy variation	≤1%
Gun normalized rms emittance	$< 5 \times 10^{-6}$ m-rad
Beam loss before bend	<0.1%
Cathode lifetime	≤200 h
Operating efficiency in any 1-week period	∼98%
Polarization	≤80%

Table 2-2. *Gun specifications.*

As with the SLC source, it is assumed that if the normalized emittance of the electron beam at the gun is a factor of 10 less than that of the bunched beam, the bunched beam will not be affected by small changes in the gun emittance.

The 1/e decay rate of the quantum efficiency of the cathode should be ≥200 h to ensure a high-operating efficiency. The lifetimes may be a moot question if it is determined that the cathodes can be cesiated at the operating high voltage without damaging the surface of the semiconductor crystal. The cathode in the SLC gun has a lifetime >1000 hrs.

The electron beam should have the highest possible polarization, preferably >80%, while still meeting all other NLC specifications, in particular the intensity requirements. While the present-generation strained GaAs photocathodes used for the SLC have satisfactory polarization performance, they are incapable of generating the long, high-intensity bunch train required by the NLC owing to an inherent cathode charge limit phenomenon [Alley 1995]. A detailed description of the charge limit phenomenon, its implications on long pulse charge production, and the steps to be taken towards meeting the polarization and charge requirements of the NLC source are given in Appendix 2.B. The insufficient charge performance may be remedied by substantially increasing the doping density in the strained GaAs. Unfortunately, the increased doping density also degrades the beam polarization. Research is underway to develop a strained GaAs cathode capable of both >80% polarization and NLC charge production with reasonably sized cathodes.

2.2.2 Design Philosophy

Except for the issues associated with long pulse (on the order of 100 ns) operation, the SLC polarized electron source has been shown to meet the gun specifications given above. Thus it is planned to use much of the SLC-source design features for the NLC source. The features include an extremely good vacuum system to ensure long cathode lifetimes, use of a loadlock for installing and removing cathodes in the gun, and provision for testing backup cathodes and guns.

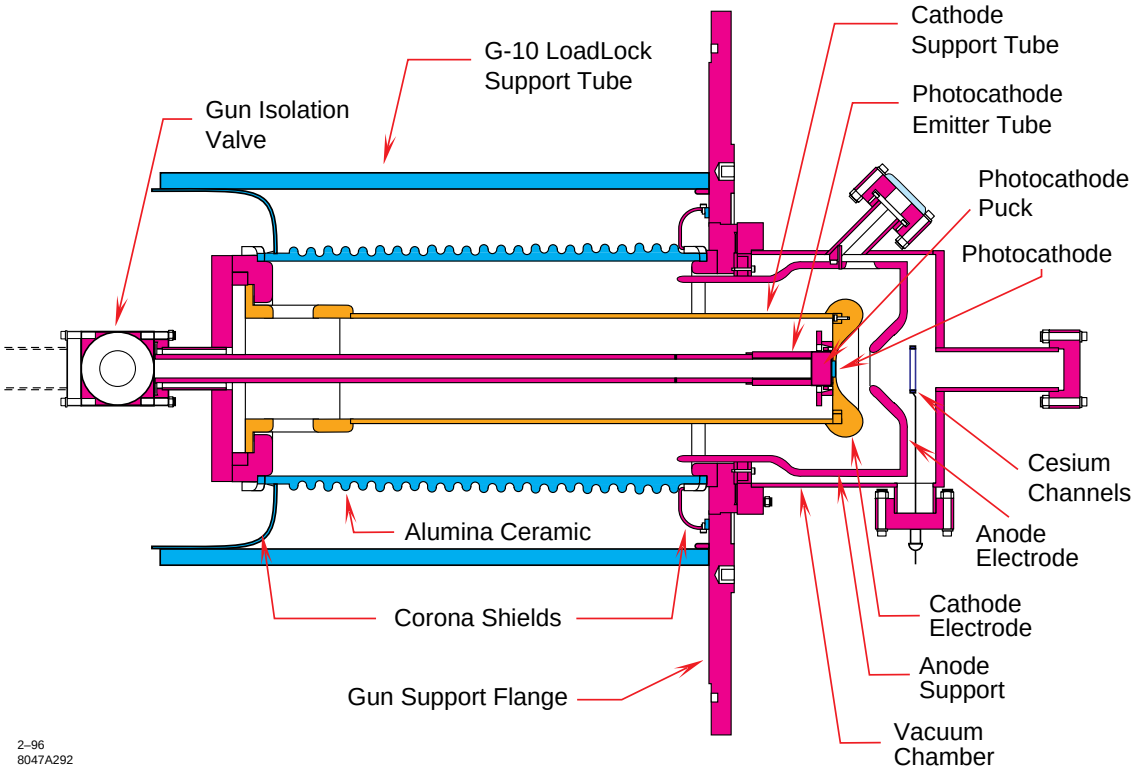


Figure 2-2. *Cross section of the SLC polarized electron gun (loadlock not shown).*

The SLC polarized electron gun with a loadlock will be the baseline design for the NLC source. Its performance and reliability have been well established with three years of smooth and nearly maintenance-free operation on the SLC. The "inverted-geometry" gun, which has a grounded body, will be considered as a logical upgrade if it is proven to be as reliable as the SLC gun. This gun has the advantage of being more compact with a simplified high-voltage operation, but it requires more testing in an accelerator environment. The cross sections of the SLC gun and the inverted-geometry gun are shown in Figures 2-2 and 2-3. The beam characteristics for the two guns are essentially identical because of similar focusing electrode design and the use of identical beam optics components. Both types of guns are configured to use the same gun support bench and will mate the downstream beam line in an identical manner. Therefore, switching between the two types of guns is as easy as switching between two guns of the same type, which should take about one day.

The preparation of photocathodes for polarized guns is a highly sophisticated operation that is not compatible with the accelerator environment. Thus it is necessary to have a cathode preparation laboratory that is separate from the accelerator. The preparation laboratory will be equipped with an ultrahigh vacuum (UHV) cathode test system capable of measuring the polarization as well as the quantum efficiency of a cathode sample at low voltage. Cathodes prepared in this laboratory for either the NLC operating source or the spare will be transported under vacuum.

Additional facilities, including certified clean rooms, will be necessary for the assembly and repair of guns.

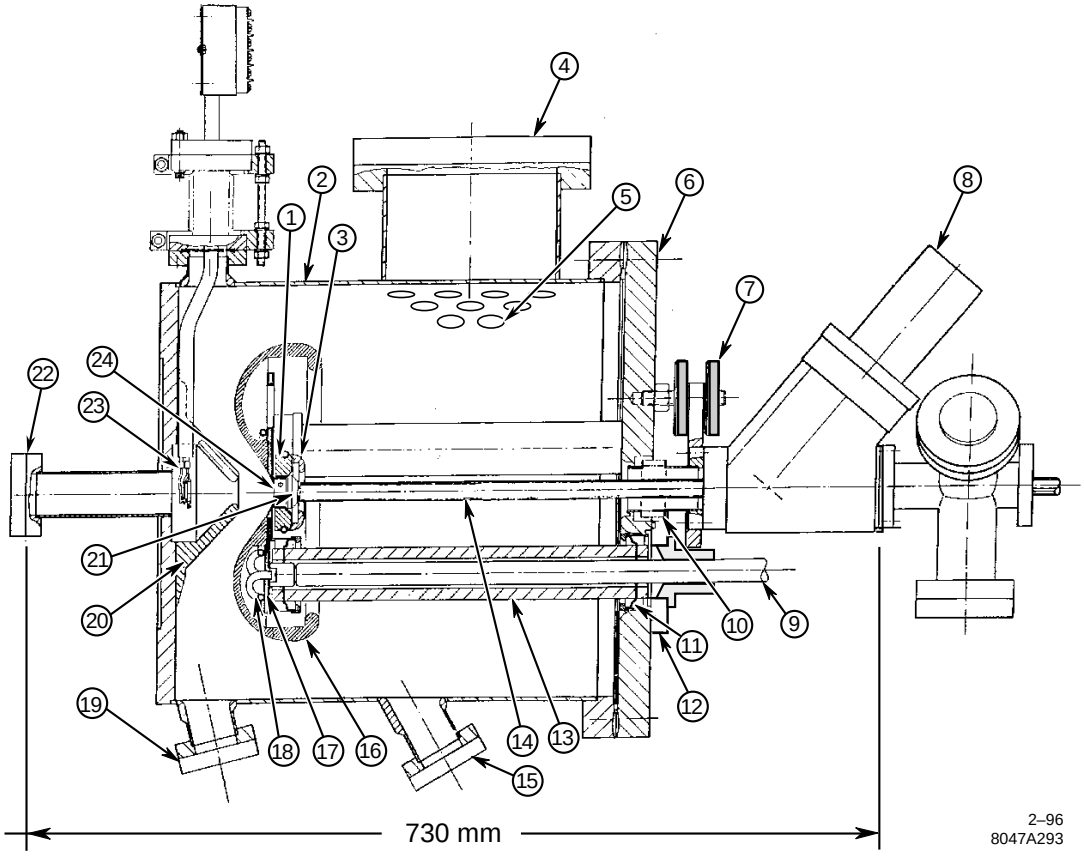


Figure 2-3. *Cross section of inverted gun: (1) docking mechanism and puck holder, (2) gun body, (3) shielding rings, (4) ion pump port, (5) pumping holes, (6) end flange, (7) jacks, (8) access valve, (9) high-voltage cable, (10) bellows, (11) Kovar weld ring, (12) gas seal flange, (13) ceramic insulator, (14) transporter rod (only in place while changing cathode), (15) Viewing port, (16) cathode electrode, (17) support plate, (18) gas tubes, (19) illumination window, (20) anode electrode, (21) puck, (22) exit port to beam line, (23) cesiator, (24) photocathode.*

2.2.3 Photocathode

Due to their superior polarization performance, photocathodes of strained GaAs epitaxially grown on GaAsP will be the primary candidates for use on the NLC. As is discussed in Appendix 2.B, the intrinsic charge limit effect inherent in p-type negative electron affinity semiconductor photocathodes may lead to inadequate charge performance for the NLC bunch train, from SLC-type uniformly doped strained GaAs cathodes with $>$80% polarization. However, there is room for significant improvement on the photoemission characteristics for such cathodes. More cathode research and development is expected to yield high-polarization, high-intensity cathodes that will meet the NLC requirement.

Conventional cathode preparation requires heat cleaning at 600°C for 1 hour. This high temperature results in some unwanted reconstruction at the surface. In addition, providing the high temperature in an UHV system places a definite limit on the reliability of the system due to the possibility of cathode contamination during the accompanying high pressures, and also because of the high potential to open a vacuum leak. Several techniques for low-temperature cathode preparation are being explored, including capping (following fabrication of the cathode) with a low-temperature

protective layer, and cleaning at low temperature in a cathode preparation chamber using an atomic hydrogen (H^*) beam.

Reducing the dark current in the SLC source to <50 nA has usually been difficult to achieve. Ultrahigh-purity water cleaning techniques now under investigation may make it easier to achieve the low levels of dark current desired for the NLC. It is possible to use a pulsed high-voltage cathode bias if necessary.

As described in Appendix 2.B, a successful low-temperature heat-cleaning technique will allow operation with a cathode having high-dopant density at the surface. Such a cathode may greatly increase not only the cathode charge limit, but also the quantum efficiency (QE) achievable when cesiating. With high initial QE achieved with the high dopant density, and long lifetimes achieved by having a clean UHV system and low dark current, it should be possible to operate the NLC source for many days without recesiating. The SLC has already demonstrated that such sources can operate continuously for many months without cathode maintenance other than recesiation. A recesiation, which takes 15 to 30 minutes, can be done by machine operators through the computer control system. The interval between cesiations on the SLC is about one week and it gets longer as the gun remains under vacuum.

The quantum efficiency decay rate, that is, the cathode lifetime is an important operating parameter. Lifetimes of over 1000 hours have been achieved with the SLC. A lifetime of 200 hours corresponds roughly to a drop of 10%/day in the value of the QE.

Other types of photocathodes are being investigated in various laboratories around the world including SLAC. The gun can accommodate any of these if they should prove superior to the SLC cathodes. A fundamental requirement for choosing a new type of photocathode material is that the polarization of the electron beam should be at least as high as the highest achieved in SLC with the GaAs-GaAsP strained lattice cathodes, and the charge limit associated with the cathode must not prohibit achievement of the desired electron intensities. In addition, there must be a source laser capable of producing the required energy at the wavelength that provides the highest polarization.

2.2.4 Polarization

The polarization of the electrons extracted into vacuum from the 100-nm SLC strained-lattice photocathodes is about 80% for the high peak intensity SLC pulses, and about 85% for very low intensity pulses for which the cathode QE can be very low. The effort to improve the GaAs-GaAsP cathodes is aimed at producing cathodes with greater than 80% polarization at NLC beam intensities with moderate-sized (3 to 9-cm^2) cathodes.

The polarization of the NLC cathodes during full current extraction must be tested for each cathode crystal before it is installed at the injector. To perform such tests, a Mott polarimeter and spin rotator, optimized for the operating energy of the NLC source, will be an integral part of the source. The Mott polarimeter will be calibrated against a polarimeter known to have an absolute accuracy of 1% or better.

There are conditions at the source that can change the polarization of the electron beam slightly. Thus, as a diagnostic the operating source as well as the development system will be equipped with a Mott polarimeter as shown in Figure 2-1.

2.2.5 Vacuum and Mechanical Design

To achieve the desired high-voltage performance, it is critical to minimize field emission from the surface regions that see high electric fields. Extremely careful attention will be given to the fabrication of the cathode and anode electrodes

and the cleanliness of all surfaces exposed to the gun vacuum. The electrodes will be fabricated using stainless steel with low carbon content and low inclusion density. The machined electrodes will be polished with diamond paste to a 1-mm finish with zero tolerance on visible pits and scratches. All components that operate at high voltage will be flushed with high-purity water before final vacuum firing. The assembly and alignment of the gun vacuum components will be done in a Class 100 clean room to avoid air-borne contamination.

Both ion pumps and nonevaporable getter (NEG) pumps will provide the pumping power for all gas species with the exception of inert gases. A high-sensitivity residual gas analyzer (RGA) will be used to monitor the gun vacuum. Following assembly, the gun is to be baked until the room temperature vacuum is $\sim 10^{-12}$ Torr excluding H_2.

The conductance between the gun chamber and the injector where the vacuum is significantly poorer will be minimized by a vacuum isolation chamber characterized by limited aperture conductance and extremely high pumping speed. A modest bend of 20° after the vacuum isolation chamber will further isolate the gun vacuum system and prevent back-reflected electrons from hitting the cathode or induce gas desorption near the cathode.

To attain the desired high-voltage performance, the gun must be high-voltage processed following the vacuum bake to eliminate field emission spots on the cathode electrode surfaces. This process involves gradually increasing the voltage in the presence of a low pressure of N^2 if necessary—until the dark current at the operating voltage is well below 25 nA.

2.2.6 Loadlock and Cathode Preparation Chamber

While the gun provides a vehicle for extracting high-intensity electron beams from a photocathode and at the same time providing the necessary UHV conditions required for reliably operating a negative electron affinity photocathode, its functionality is fully realized only with the help of a loadlock and a cathode preparation chamber. The preparation chamber is a UHV system in which a cathode is prepared to have an negative electron affinity surface, while the loadlock allows the cathode to be installed in and removed from the preparation chamber and installed in the gun itself without breaking vacuum in either system.

Since the cathode preparation is time consuming, taking about 8 hours and even longer when using a loadlock, it is best done in a chamber separate from the gun. The necessary cleaning techniques and the recipe for applying Cs_2 and an oxide are well established.

2.2.7 Inverted Geometry Gun Tests

The prototype inverted structure gun at SLAC will be extensively tested to ensure it can match the SLC photocathode gun for good cathode performance. Based on the results of these tests, the details of the NLC gun upgrade design will be determined.

2.2.8 Reliability

Based on the SLC experience, it is believed that achieving near 100% operating reliability over a period of several years of continuous operation is possible. The provisions to be incorporated into the NLC source to ensure this result are:

- Use of a loadlock system to insert and remove cathodes into the gun without breaking the gun vacuum.
- Possibly using a double-gate valve for connecting the gun to the loadlock to reduce the chance of vacuum contamination of the gun should one gate valve fail.
- A vacuum isolation/differential pumping section downstream of the gun to protect it from the poorer vacuum system of the injector.
- Provision for a spare polarized electron source for completely testing cathodes for intensity, lifetime, and polarization.
- Provision of a transfer system for moving cathodes under vacuum between the NLC source, the spare source, and the cathode activation chamber.
- Provision of spare guns and spare cathode activation chambers.

2.2.9 Operation

Following the SLC experience, it is possible for the accelerator operators to remotely monitor and perform routine maintenance on the polarized electron source through a computer control system. For this purpose, the following displays will be available to the operators:

- The full residual gas spectrum of the gun and downstream vacuum systems.
- The transverse and longitudinal shape of the electron pulse at the gun.

A complete set of analog and digital signals will be available to the operators. These signals will also be history-buffered. The analog data will be collected locally for each pulse over the last few-minute period. A representative value of the data will be history-buffered unless the system indicates a problem, such as a vacuum burst, in which case data will be recorded for every pulse. The principle operational task for the operator is to periodically cesiate the cathode. As the SLC has thoroughly demonstrated, this process can be fully automated (but initiated by operator command).

2.3 NLC Laser System

The NLC laser requirements are driven by the requirements on the electron beam from the polarized source photocathode [Alley 1994]. The most important of these requirements are the polarization and the charge limit properties of the GaAs cathodes. The laser wavelength needed varies with different cathode types and is chosen to maximize these parameters. Thus a range of wavelengths from 760 to 890 nm is needed. A quantum efficiency of 0.1% is assumed for the cathode based on the SLC experience. Electron beam optics considerations determine the other properties of the laser system such as longitudinal and transverse shape, stability, and so on.

The proposed NLC laser system uses Titanium-doped sapphire (Ti:Sapphire) as the primary laser material. Pumping of the Ti:Sapphire is provided by a pair of commercial Nd:YAG lasers. Overall macropulse optical shaping is performed with an electronic feedback driving a Pockels cell attenuator. Micropulse shaping and selecting is done with a series of resonant Pockels cells. All of the technology used is fairly conventional, although the resulting system is somewhat

complex. A schematic of the laser system is shown in Figure 2-4. The laser system is expected to meet the following specifications:

Wavelength	760 to 890 nm
Polarization	$> 99.7\%$
Bunch train length	Up to 150 ns
Bunch spacing	714 MHz, 357 MHz or 178.5 MHz
Bunch width	< 1 ns as close to rectangular as possible.
Pulse contrast	$<$ 5:1000 for the 714 MHz spacing, more is tolerable for the 357 and 178.5-MHz spacing.
Spot size on the cathode	Variable from 10 to 20-mm in diameter.
Transverse uniformity	As close to rectangular as possible.
Bunch energy	20 mJ
Bunch train energy stability	0.5% rms
Repetition rate	180 Hz

2.3.1 Pump Lasers

The pump laser system is required to produce a total of 150 mJ, 180 Hz, at approximately 500-nm wavelength. This can be done using two commercial Q-switched frequency-doubled Nd:YAG lasers. We have chosen to use two Coherent Infinity YAG lasers. These lasers use a seed laser and a phase conjugate amplifier configuration to obtain good transverse mode quality at high repetition rates. The two pump lasers are each operated at 90 Hz, and interleaved to obtain the required 180-Hz output.

Beam interleaving is achieved by rotating the polarization of one of the pump lasers and then combining them with a polarizing splitter/combiner. A pulsed Pockels cell is used to control the polarization of the combined pump beam. A waveplate is used to adjust the splitting ration between the oscillator and amplifier. A second waveplate and polarizer is used to adjust the pump energy to each end of the amplifier.

2.3.2 Ti:Sapphire Oscillator

The oscillator produces a long ($\sim$150 ns) optical pulse with approximately 1 mJ of energy. Wavelength tuning is accomplished with a series of Brewster angle prisms (birefringent tuners would also work). A Pockels cell and the S-polarization loss of the prisms are used to Q-switch the cavity. The resulting output pulse is later chopped to produce the required pulse structure.

The primary source of output fluctuation in this system are changes in pump laser energy changing the output pulse energy and timing (due to changes in gain). The timing changes, coupled with the pulse chopping, produce additional intensity fluctuations. We plan to use a feedforward system which, on each pulse, adjusts the Q-switch time for the Ti:Sapphire laser to compensate for changes in the pump energy. If the system detects a high energy pump pulse, it delays the Q-switch time to compensate for the decreased build-up time (due to higher gain). The system is adjusted to slightly under-compensate for the gain change, the resulting slight timing shift (of the Q-switched pulse, not the output) compensates for the changes in Q-switched pulse energy.

A system of this type is used on the SLC source laser to provide an output with intensity fluctuations of $<$1% rms (0.6% has been demonstrated) with pump fluctuations of $>$3% rms. The limiting factor here is the stability of the

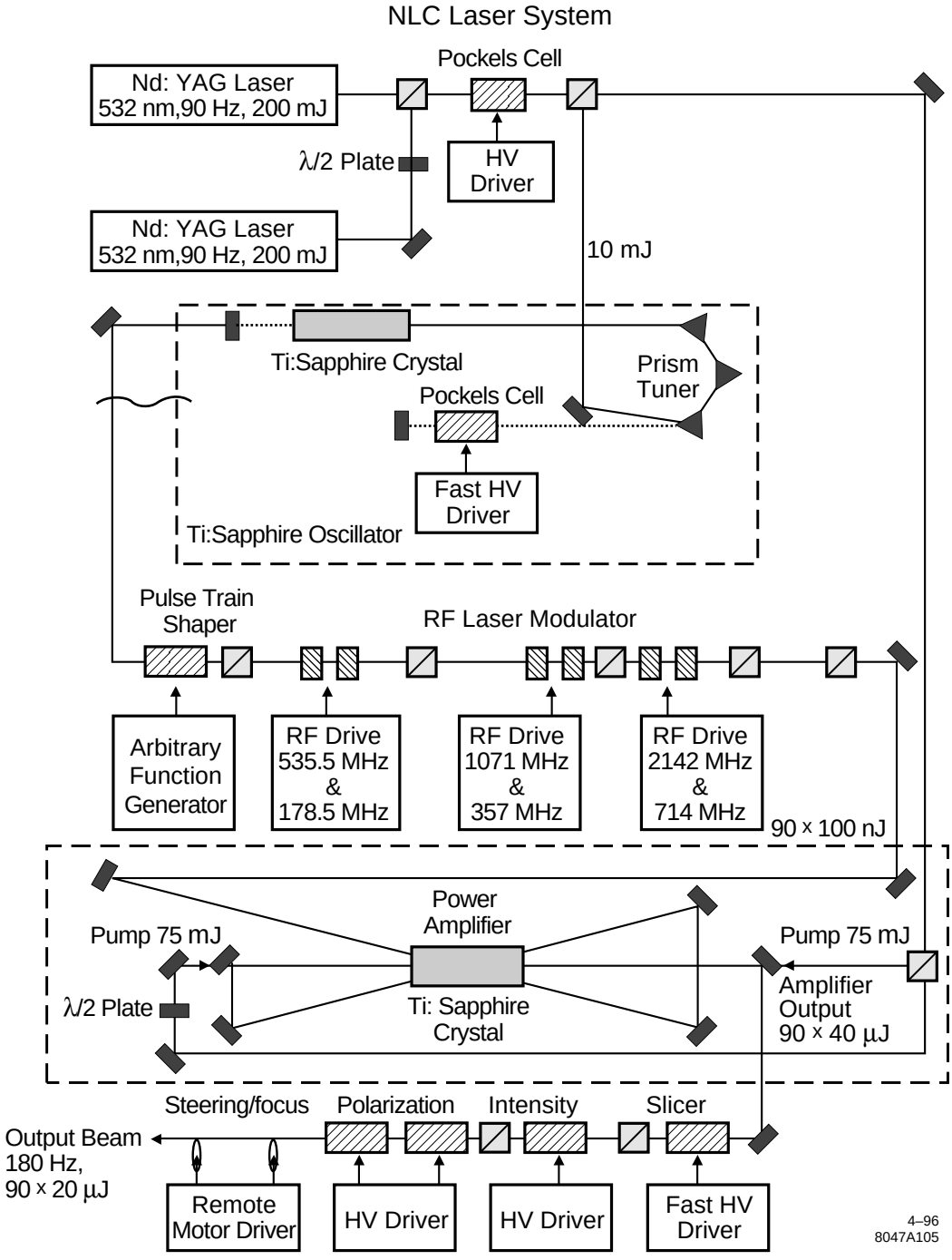


Figure 2-4. *Schematic diagram of the NLC polarized source laser system.*

pump laser intensity. We believe this technology can be improved to reduce output fluctuations to <0.5% rms (with expected pump fluctuations of <2% rms) and to improve the feedforward system.

All the oscillator components are commercially available. With the exception of the intensity stability, the required output parameters should be obtainable using conventional technology. The intensity stability should be obtainable with a modest extension of the feedforward technology used at SLAC.

2.3.3 Pulse Shaping

The pulse shaper uses an arbitrary function generator and a fast (10 ns) high-voltage (5 kV) amplifier driving a Pockels cell to shape the output pulse. The final electron beam pulse train shape will be measured and a software feedback will be used to control the function generator. The Pockels Cell are commercially available. The high-voltage driver is similar to units developed at SLAC for other laser projects.

2.3.4 RF Laser Modulator

The NLC requires a train of bunches about 1-ns wide, nearly square in shape, and 1.4-ns apart. In addition, the option of running with a 2.8 or 5.6-ns bunch separation is also desired. The laser system will use a series of rf-driven resonant Pockels cells to produce the required pulse shape. Two rf-driven Pockels cells at f_1 = 714 MHz and f_2 = 2142 MHz in series and a polarizer will produce an optical train according to the equation $\mathrm{OT}_{1.4} = \mathrm{A}\sin^2[\sin(\omega_1 t - \pi/4) + 0.2\sin(\omega_2 t - 3\pi/4) + \pi/4]$ which is a train of nearly square pulses 1.4-ns apart and 1-ns wide as illustrated in Figure 2-5. To produce the 2.8-ns pulse separation, this system of Pockels cells and polarizer is used in series by two other Pockels cells at $f_3 = 357$ MHz and $f_4 = 1071$ MHz and a polarizer, and for the 5.6-ns bunch separation, an additional pair of Pockels cells at $f_5 = 178.5$ and $f_6 = 535.5$ MHz and a polarizer is used. Thus, the equation for the train of bunches separated by 5.6 ns is the product of the output of each of the pair of Pockels cells and polarizer systems:

$$\begin{aligned}\mathrm{OT}_{5.6} = \quad & \mathrm{A}\sin^2[\sin(\omega_5 t + \pi/16) + 0.2\sin(\omega_6 t + 3\pi/4) + \pi/16]^* \\ & *\sin^2[\sin(\omega_3 t + \pi/8) + 0.2\sin(\omega_4 t + 3\pi/8) + \pi/4]^* \\ & *\sin^2[\sin(\omega_1 t - \pi/4) + 0.2\sin(\omega_2 t - 3\pi/4) + \pi/4]\end{aligned} \tag{2.1}$$

The laser intensity between the main pulses is almost completely quenched and the area of the residual pulse between the main pulses is 0.2, 0.5, and 1.4% of the main pulse for the 1.4, 2.8 and 5.6-ns separation cases, respectively. It is expected that most of this charge produced between the main pulses will be lost early in the injector beam line due to its poor bunching and matching into the lattice, as compared to the main bunch.

If standard KDP Pockels cells are used, a peak rf voltage of approximately 5 kV is required. With a resonant circuit Q=100, the required drive power is approximately 3 kW peak (<1 W average).

2.3.5 Ti:Sapphire Power Amplifier

The power amplifier operates in a 3-pass "bow tie" configuration with a gain of 10 each pass. A fairly conservative amplifier efficiency of 5% is assumed. Both ends of the laser rod are pumped to reduce the possibility of damage from the pump laser energy density. The required pump energy density of approximately 2 J/ cm^2 is within the limits for continuous operation. The average pump power to the crystal is approximately 30 W, well within the thermal fracture

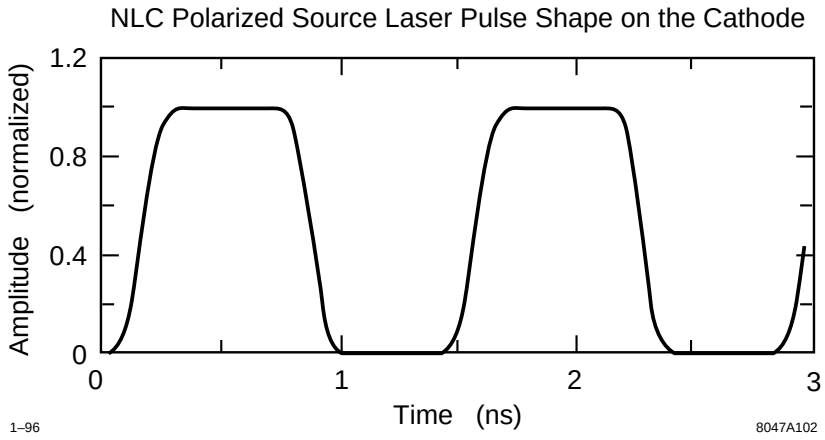


Figure 2-5. *NLC polarized source laser pulse shape at the cathode.*

limit for Ti:Sapphire. The Ti:Sapphire thermal fracture limit is 5 times more than for the YAG [Wanant 1994], the YAG limit is 30 W/ cm^2, [Koechner 1992], and the spot size on the cathode is 2 to 3 cm^2.

The feedforward system used in the oscillator can also compensate for changes in the gain of the amplifier. This should permit operation at the required 0.5%-rms output stability.

2.3.6 Pulse Train Length, Intensity, Spot Size, and Steering Control

The train of pulses from the amplifier is sliced to the desired overall length with a fast, high-voltage driver and a Pockels Cell. This system will probably have rise and fall times of approximately 5 ns. Faster pulses may be possible, but some high-voltage driver development would be required. The output intensity is controlled by a high-voltage amplifier driving a Pockels cell.

The output polarization is switched using two Pockels cells and high-voltage driver. Laser circular polarizations $>$99.7% are easily obtainable. Electron polarization measurements can be used to optimize the drive voltage on the Pockels cells to eliminate any birefringence in the transport line to the photocathode.

The beam spot size and position on the photocathode can be controlled with a remotely actuated telescope.

The propagation path and optics of the laser will be determined by the geometry of the physical layout of the injector with respect to the laser and is straightforward.

2.3.7 Overall Technical Risks

The most challenging requirement for the laser system is the output intensity stability. It is believed that this can be met with the use of feedforward system. The rf laser modulation is in principle straightforward, but some development may be required to obtain the required electrical Qs for the Pockels cells. The transverse mode structure from the output of the amplifier is difficult to predict. With the addition of spatial filtering, it should be possible to produce a

good transverse-mode beam. In general, development of this laser system is anticipated to require approximately three man years.

2.4 NLC Electron Injector Beam Dynamics

2.4.1 NLC Injector Baseline Design

As shown in Figure 2-1 and discussed in Section 2.1, the NLC injector is like the conventional injector used on the SLC with a DC high-voltage gun and subharmonic bunching system. In this section we describe the beam dynamics from the gun through the injector linac.

2.4.2 Electron Gun Optics

The plan is to use a modulated laser to extract an electron bunch train from the photocathode. The electron beam energy from the gun will be 120 keV, a comfortable operating point for the polarized electron source. The proposed gun for the NLC polarized electrons has a large, 3-cm^2 cathode. The electrodes are shaped to minimize emittance for operating bunch currents and reduce the electric field gradients near the cathode to avoid arcing in this region. We used EGUN to simulate the beam from the gun. Figure 2-6 shows the ray-trace of the beam from the cathode to 12-cm downstream. The normalized edge emittance is 5.6×10^{-6} m-rad and the beam envelope radius is 1.2 cm at the tip of the anode.

2.4.3 Intensity Jitter Limiting Aperture

A scenario for reducing the electron intensity jitter is to locate a fixed aperture in the 120-keV beam line at a location where the beam size is proportional to space charge. With this method, we hope to reduce the intensity jitter by scraping more charge when more charge is produced at the gun and less charge when less charge is produced at the gun. For the proposed NLC injector design, such a location exists 24-cm downstream of the 20° bend. Using a 1.25-cm aperture, we can reduce a 2% intensity jitter at the gun to be immeasurable in the bunched beam, while losing 17% of the charge in the aperture. Since this loss occurs after the gun vacuum isolation chamber and the 20° bend, it should not cause any damage to the photocathode based on various gun test experiences on the SLC polarized source. Figure 2-7 shows the correlation of charge per bunch in the downstream part of the injector with the charge produced at the gun. Figure 2-8 shows the charge in 18 ps at 80 MeV with and without the aperture. Notice that without the aperture the correlation has about a 45° slope, while with the appropriate aperture the slope is greatly reduced for the bunched beam intensity.

Studies were conducted to see how sensitive this scheme is to the location of the aperture along the beam line and to various nominal charge intensity scenarios. To study the effect of the criticality of the location of the aperture it was moved in simulations by ± 2 cm from the nominal location and no observable difference in the performance of the aperture was detected. Reducing the nominal operating intensity to 3/4 or 1/2 of the nominal charge has a measurable effect on the aperture size needed for reducing the intensity jitter. At 3/4 charge, an aperture 1.5-mm smaller in radius is needed and at 1/2 charge, another 1-mm smaller is needed. To accommodate the changes in the operating intensity

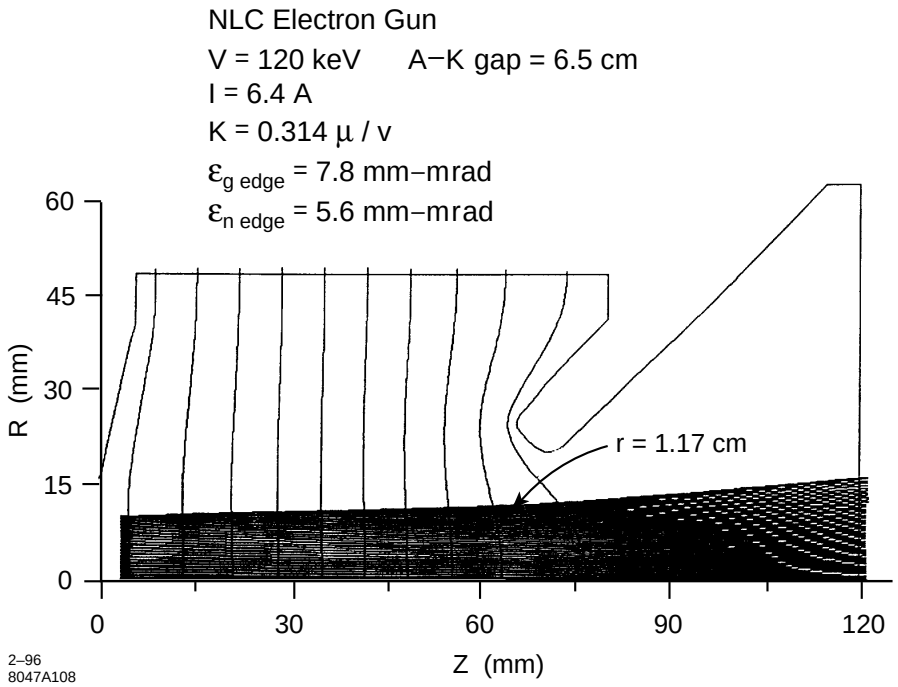


Figure 2-6. *Ray trace of the NLC electron beam from the gun.*

from the gun, the aperture would have to be mechanically designed such that its size can vary in small increments of less than 1 mm.

2.4.4 Bunching System

As shown in Figure 2-1, two 714-MHz standing wave subharmonic bunchers are located 74-cm apart downstream of the intensity jitter limiting aperture. The S-band buncher is a four-cavity, $\beta = 0.75$ traveling-wave section and is located 42-cm downstream of the second subharmonic buncher. Immediately following the S-band buncher is a $\beta = 1$, 3-m S-band accelerator section which acts as a further buncher in the first meter. This bunching system compresses the beam such that 83% of the in-coming charge is captured in 18° of S-band or 17.5 ps. Given that 17% of the beam is lost in the intensity limiting aperture, for $2.8 \times 10^{10} e^-$ per bunch at the gun we expect to have $1.96 \times 10^{10} e^-$ in 18 ps per bunch after bunching. PARMELA simulations were conducted from the gun to the end of the second accelerating section where the beam energy is 80 MeV±0.6%. Figure 2-9(a), (b), (c), and (d) show the single bunch charge distribution profile in time, transverse particle distribution, longitudinal particle distribution, and the energy spread profile, respectively, at the end of the second accelerator section.

At the end of the second accelerator section the beam-normalized rms emittance is 4.5×10^{-5} m-rad. Figure 2-10 shows the normalized rms emittance from the gun up to the end of the second accelerator section and Figure 2-11 shows the beam edge envelope.

Beam loading in the subharmonic bunchers is minimized by judiciously choosing a low $R/Q \sim 10\Omega$, a high $Qo \sim$10000, and a coupling coefficient of $\beta = 1$, we can force a low beam-induced voltage and a long filling time of the cavity,

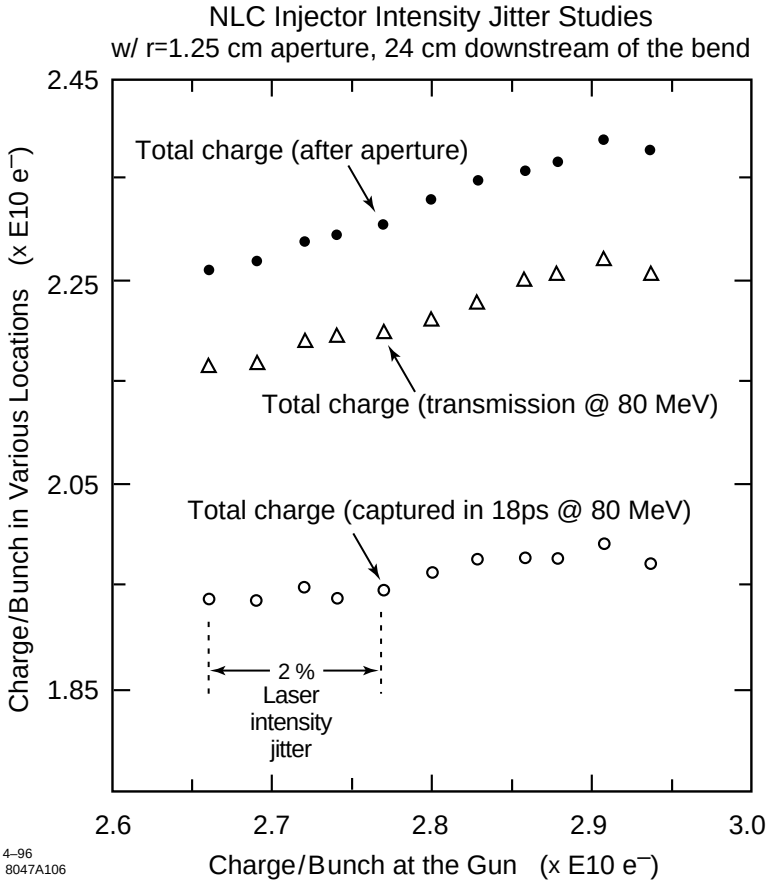


Figure 2-7. *NLC injector intensity at various locations as a function of charge from the gun going through a fixed aperture after the 20° bend near the gun.*

thus minimizing the effects of beam loading in it. The fill time in this case will be 2.2 μs, much greater than the 126-ns bunch train length.

Since 17% of the charge is intercepted by the intensity jitter limiting aperture for the highest charge operating scenario, with 4.5 nC from the gun there will be 3.7 nC at the first subharmonic buncher and with a 700-ps bunch length it will occupy 180° of the 714-MHz subharmonic buncher rf. The beam-induced voltage in the first subharmonic buncher by the time the ninetieth bunch goes through will be:

$$\mathrm{Vb1} = [0.5q(R/Q)\omega/\pi] \times 90 \text{ bunches} = 3.7\,\mathrm{nC} \times 10\omega \times 714\,\mathrm{MHz} \times 90 = 2.4\,\mathrm{kV} \quad . \tag{2.2}$$

For the second subharmonic buncher, the bunch length is more like a δ function and

$$\mathrm{Vb2} = [0.5(R/Q)\omega] \times 90 \text{ bunches} = 3.7\,\mathrm{nC} \times 10\omega \times \pi \times 714\,\mathrm{MHz} \times 90 = 7.5\,\mathrm{kV} \quad . \tag{2.3}$$

Bunching simulations indicate that we need 18 kV in the first subharmonic buncher and 34 kV in the second for optimum bunching for the highest charge scenario. Since the centroid of the bunch is at about 90° ahead of the crest, the induced phase change in the first subharmonic buncher will be arctan(2.4/18) ~8° and in the second arctan(7.5/34) ~12°. This phase shift is not negligible and needs to be compensated. We plan to drive the cavity with as much power as would be needed to compensate for the voltage induced by the beam in the steady state plus the voltage needed for

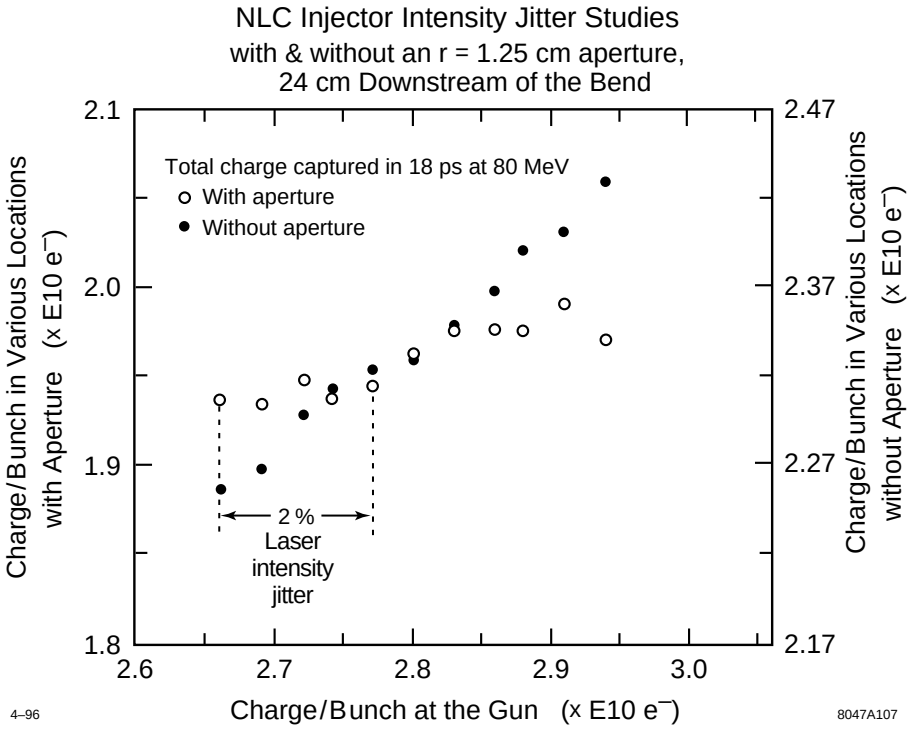


Figure 2-8. *NLC injector intensity after bunching as a function of charge from the gun with and without an aperture after the bend.*

optimum bunching, and then inject the beam during the fill time of the cavity when the voltage in the cavity reaches the voltage needed for optimum bunching. At the time of beam injection, the phase of the generator power must be shifted by the amount that would be induced by the beam in the steady state.

The average current in the bunch train is 3.7 nC/1.4 ns= 2.6 A. The steady-state beam loading in the first subharmonic buncher where the beam occupies 180° of the rf, is (1/2) $\times$ Ib $\times(R/Q) \times Qo/(1+\beta) = 65$ kV. The generator power needed for bunching and beam-loading compensation is 42.3 kW. For a beam which is at 90° ahead of the crest, the steady-state phase shift is $\arctan(\mathrm{Vbs}/\mathrm{Vg}) = \arctan[65/(18+65)] = 38°$, where Vbs is the steady-state beam induced voltage, and Vg is the generator voltage. Thus, at the time of the beam injection, the generator power is shifted by 38°, which is felt by the cavity at the fill rate. By the time the entire bunch train goes through the cavity, the generator voltage felt by the cavity is only 2.3 kV with an 8° phase shift, exactly what is needed to compensate for the beam loading.

For the second subharmonic buncher where the bunch width is a δ function compared to 714 MHz, the steady-state beam loading is Ib $\times(R/Q) \times Qo/(1+\beta)$ =130 kV. The generator power needed to compensate for beam loading plus the bunching voltage is 167.8 kW. The phase shift induced by the steady-state beam loading is $\arctan[135/(34+135)] = 39°$.

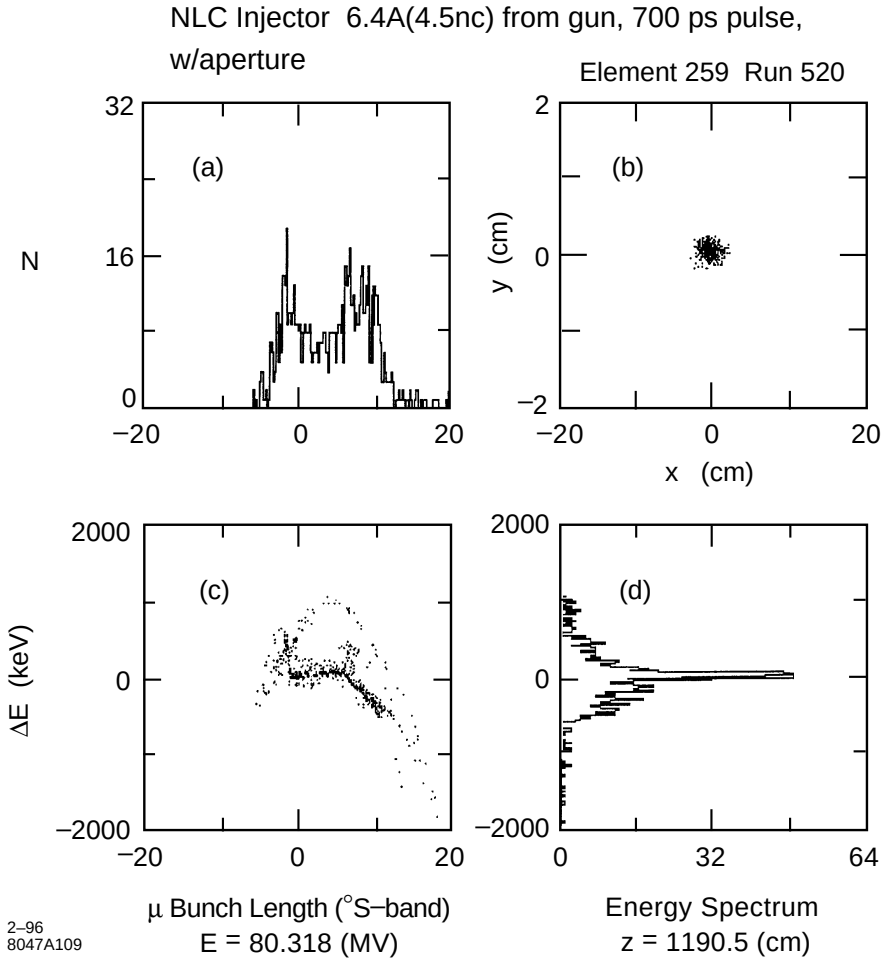


Figure 2-9. *PARMELA results of beam longitudinal and transverse distribution at the end of the second accelerator section. (a) bunch shape, (b) transverse distribution, (c) longitudinal distribution, (d) energy spread.*

2.4.5 Injector Linac

The Linac for accelerating the bunched electrons from 80 MeV to 2 GeV consists of 3-m S-band structures. The nominal rf frequency for the accelerator structures is 2856 MHz. Two SLC model 5045 klystrons drive one SLED1 cavity and the power out of the SLED drives four accelerator sections on one module. Figure 2-12 shows the layout of one module, and the injector for the collision electrons will employ 11 such modules. The beam loading compensation method using the Δt technique for the S-band linac is described in detail in Chapter 6. For the highest charge operating scenario, the maximum current in the injector linac will be 2.2 A. For 2.2 A in the injector linac the maximum possible energy gain per module is 208 MeV with 60-MW usable klystron power. in the accelerating structures. The central energy variation from bunch 1 to bunch 90 is about 0.005% at the end of each accelerator section, and the single bunch energy spread due to the bunch width and assuming the beam is on the crest will be $\pm$0.6% edge.

The lattice from 80 MeV to 2 GeV up to the linac-to-ring transition was designed using MAD. A 30° achromatic bend system at 80 MeV allows us to introduce a second injector line parallel to the first one up to 80 MeV. Two sets of

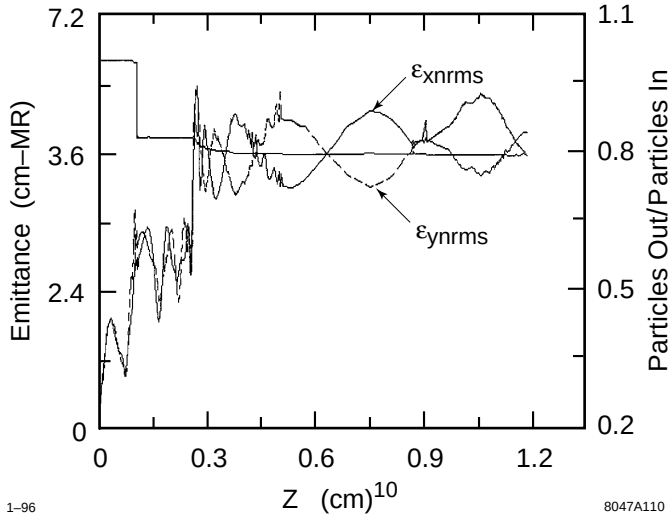


Figure 2-10. *Beam emittance along the injector from gun to the end of the second accelerator section (80 MeV).*

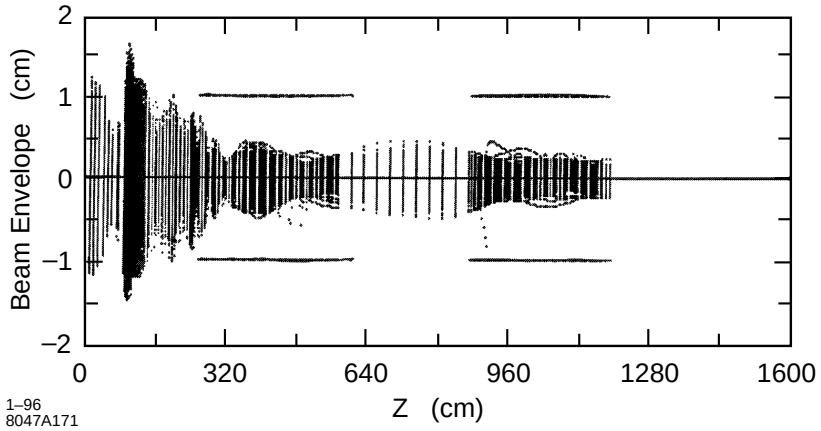


Figure 2-11. *Electron beam edge envelope from gun to end of the second accelerator section (80 MeV).*

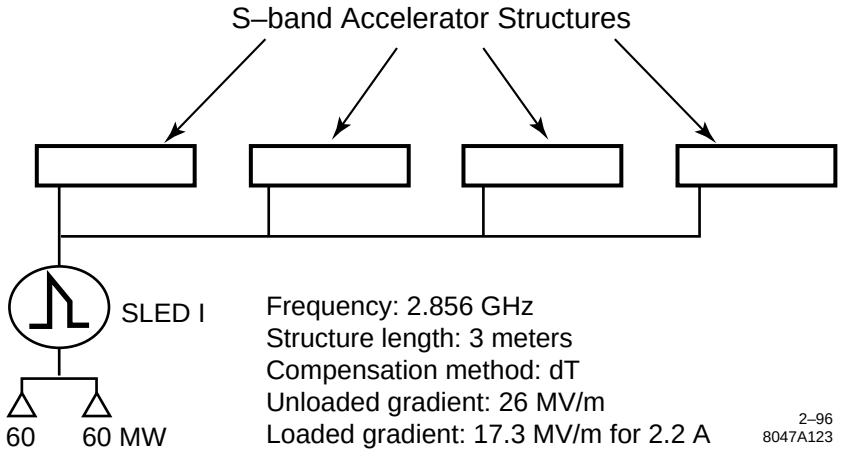


Figure 2-12. *A single module configuration for the 2 GeV electron injector linac.*

emittance measurement stations are included, one on each side of the bend to measure emittance coming out of the bunching system and the other for measuring the emittance into the 2-GeV accelerator. The transport line from the end of the second accelerating section to the beginning of the first injector linac module has been designed. The lattice is shown in Figure 2-13. The bend is an achromatic and isochronous system consisting of two rectangular dipole magnets which bend the beam by 30°. A FODO lattice with seven identical quads separated by 70 cm from center to center are used. The beam envelope for $\pm 2\%$ energy spread is less than 1 cm in radius at its maximum and comfortably fits in a 1.5-in pipe. The bunch length expansion is only 0.08 ps per 1% energy spread. The total length of the achromatic bend is 4.9 m.

The accelerator lattice is designed to take the $\beta = 1.4$ m beam achieved at the end of the second accelerator with PARMELA simulation and increase it by a factor of the square root of the energy up to 2 GeV. As shown in Figure 2-1, wrap-around quadrupoles are used around the first part of the accelerator section to maintain the small β function at the low energy end where the wakefield effects are more significant. We assumed that each module contributes 192 MeV for the lattice design to allow for some energy overhead. With this assumption, the beam reaches 2 GeV after 10 modules, thus there is an additional module for overhead. While the spare module will not accelerate the beam during normal operation it still needs at least one klystron to be operating for beam loading compensation in that module as described in more detail in Chapter 6. If a klystron fails in one of the accelerating modules then the spare module can be powered in the nominal way while the module with the problem klystron becomes the spare. Figure 2-14 shows the lattice for the accelerator and Figure 2-15 shows the lattice for the matching section from the end of the 2-GeV linac to the linac-to-ring transition. Another emittance measurement station is included in this portion of the injector. Figure 2-16 shows that the pole-tip field of all the quadrupoles in the 2-GeV injector is under 8 kGauss, which is a very comfortable operating range.

The energy spread from the beginning of the train to the end of the train due to beam loading is compensated for completely at the end of each accelerator section. Thus at the doublets between the acceleration sections, the energy spread is uniform from the first to the last bunch and is due only to the single-bunch width. However, at the wrap-around quads on the accelerator sections themselves, at the beginning of the accelerator there is energy variation in the energy of the individual bunches within the train. This is because the gradient at the quad location is different when the first bunch goes by verses when the last bunch goes by. We tracked the first, the middle and the last bunch of the train though the design lattice to see how this effects the emittance out of the injector for the various bunches. We assumed an input single-bunch energy profile which is similar to the energy profile at 80 MeV as calculated by

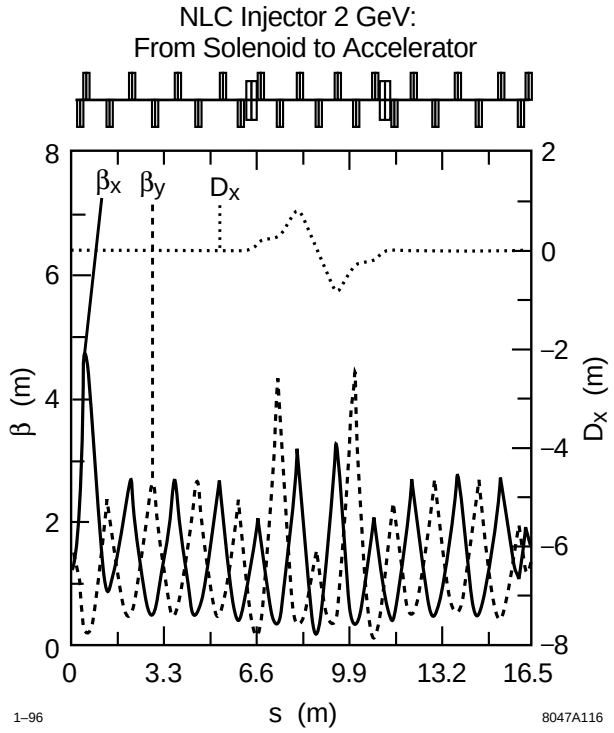


Figure 2-13. *Electron injector lattice to transport beam from the end of the solenoids to the injector accelerator.*

PARMELA, and central energy variation at the wrap-around quads based on the beam loading compensation scheme. Tracking simulations show that at the end of the 2-GeV linac there is only 3% emittance growth in the x plane and none in the y plane. The bends in the 80-MeV achromat are in the x plane.

The effects of the wakefield on the emittance were estimated assuming that the beam is offset by one sigma from the accelerator centerline and that the β grows as the square root of the energy monotonically starting with 1.5 m at 80 MeV. At 2 GeV the actual emittance of the beam increases by only 4% but the effective emittance growth, taking into account that the beam is not centered, is 44%. Even with this emittance growth the total emittance will be about 0.63×10^{-4} m-rad, well within the 1×10^{-4} m-rad threshold.

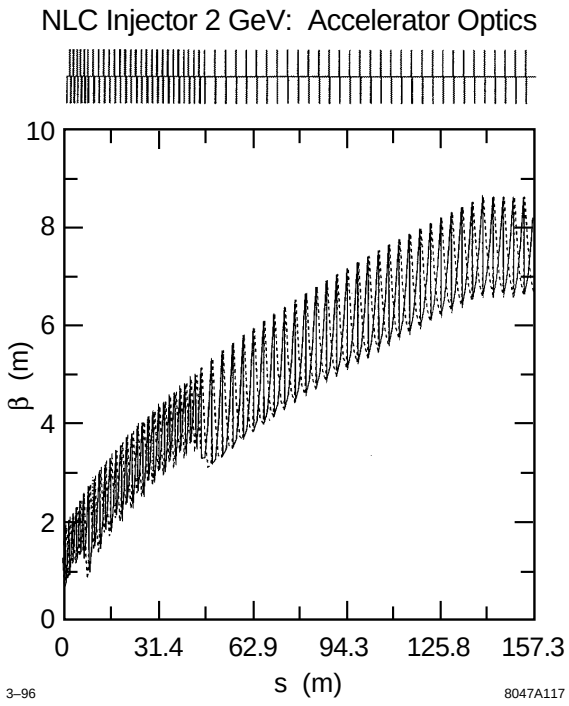


Figure 2-14. *Electron injector lattice over the 2-GeV accelerator.*

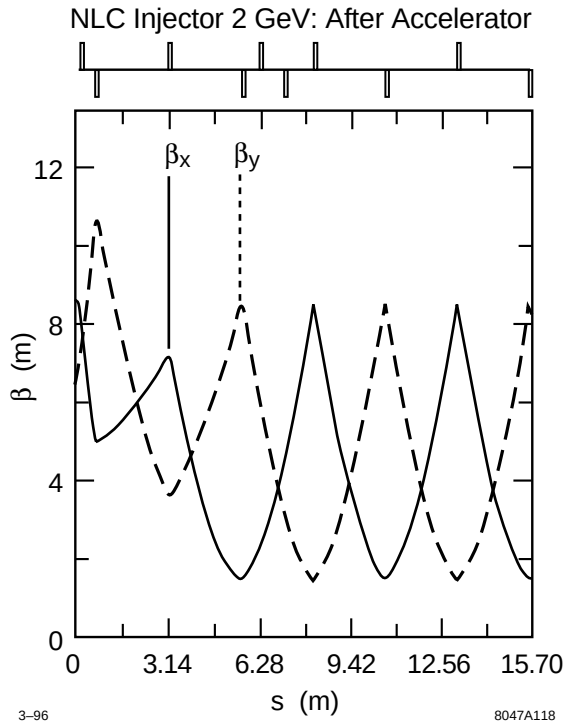


Figure 2-15. *Electron injector lattice from the 2-GeV accelerator up to the damping ring.*

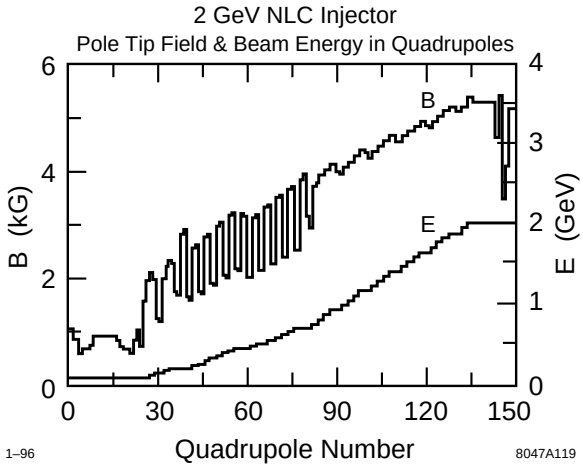


Figure 2-16. *Electron injector quadrupole pole-tip fields and electron beam energy from the end of the solenoids up to the damping ring.*

2.5 Buncher Cavities

There are two types of bunchers: Two standing wave 714-MHz subharmonic bunchers and an S-band traveling wave buncher with four cavities with a phase velocity of 0.75 c. The traveling wave buncher is similar to the one on the SLC and will be integrally brazed onto the first 3-m accelerator structure.

The subharmonic buncher will have the following parameters:

f =	714 MHz
R/Q	$\sim$ 10 W
Q	$\sim$ 10000
β	$\sim$ 1

Subharmonic structures with similar parameters need to be constructed for the NLCTA injector upgrade project, thus much experience is expected to be gained in this area.

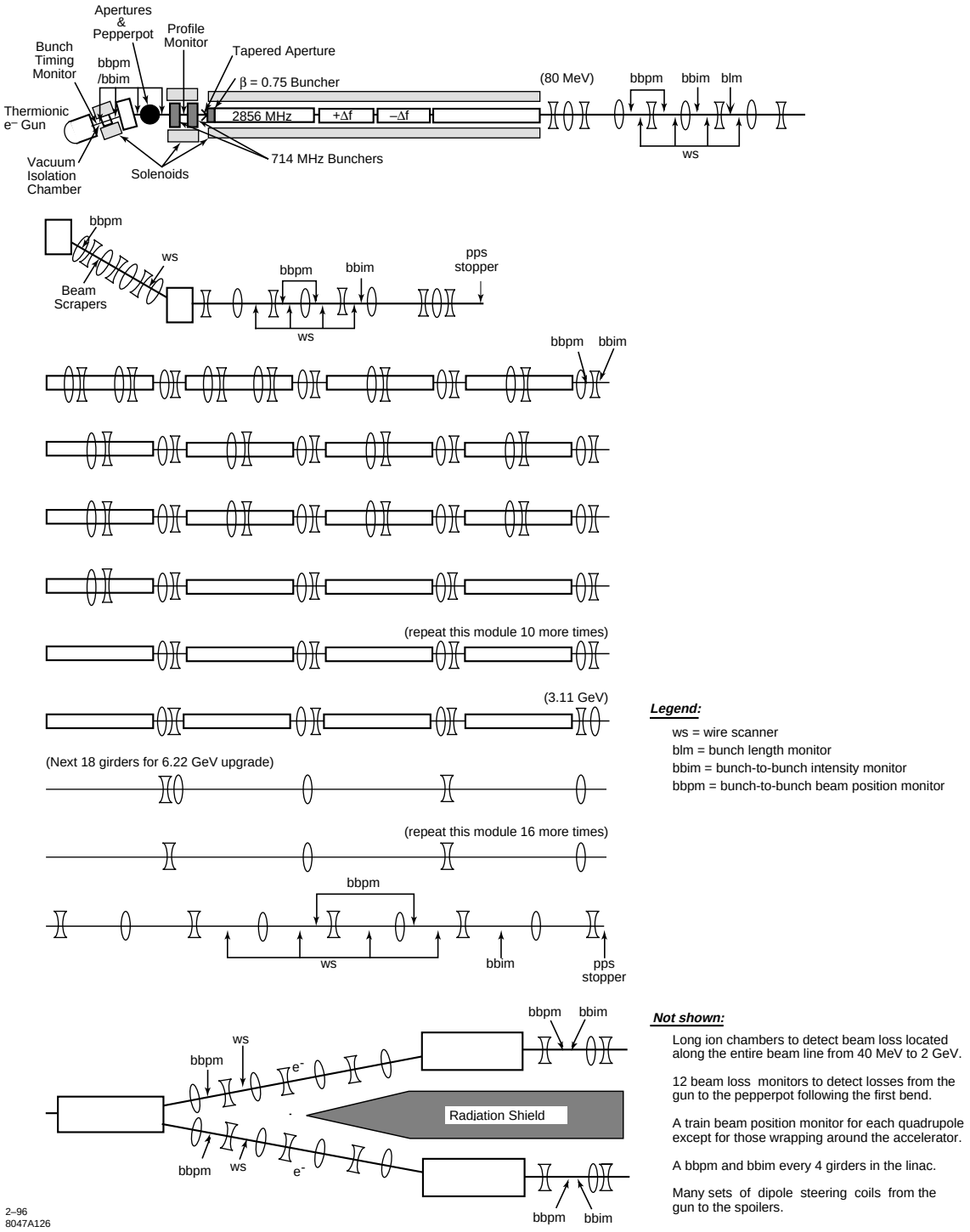
2.6 Positron Drive Linac

The linac for the positron drive beam is similar to the injector of the collision electron beam. Two differences are that the positron drive linac will use a thermionic gun and will extend to 3.11 GeV in NLC I and 6.22 GeV in NLC II with optics to transport the electron beam to one of the parallel positron production beam lines. Figure 2-17 shows the electron injector for the positron drive beam. Initially, accelerator sections will be installed to accommodate the 3.11-GeV injector with one module to spare, and room will be left to install accelerating sections to double the energy of the drive beam for NLC II.

The thermionic DC gun rather than a photocathode DC gun is chosen for the positron drive injector because thermionic guns are easier to construct and operate, and they do not have an expensive laser system associated with them. One major difference between the thermionic gun for the NLC and the thermionic gun which was used on the SLC is that the NLC requires a bunch train format right from the gun. The nominal bunch train consists of 90 bunches, each bunch 700 ps in full-width-half-maximum and 1.4-ns apart. In addition, we must be able to change the bunch separation to 2.8 and 5.6 ns. We believe it will be too difficult to achieve all three possible train structures with one pulser, but since the change in train format is not a routine operation and we would be willing to take the time to exchange pulsers, we feel that it would be possible to construct three different pulsers, one for each bunch separation scenario. The 1.4-ns bunch interval pulser can be a resonant pulser while the other two would have to be broad-band pulsers similar to the ones constructed for the KEK Advanced Test Facility thermionic gun [Naito 1994].

The positron drive injector beamline from the gun through 80 MeV is identical to the polarized e-beam injector. The lattice from 80 MeV to the spoiler in front of the positron target in either of the two parallel positron production lines has been designed. Figure 2-18 shows the β function from 80 MeV though the 3.11-GeV linac and the allocated space for additional accelerators for NLC II. Figure 2-19 shows the β function for the beam transport from the end of the accelerator to the spoiler including the 10° achromatic bend to switch the beam to either of the positron production lines. The beam size at the spoiler in front of the positron target is designed to be $\sigma = 0.5$ mm, about 1/3 of the size required at the target after the spoiler. Figure 2-20 shows the quadrupole pole-tip strengths for the positron drive beam linac which are at very reasonable values at less than 8 kGauss.

As mentioned before, the β function and the beam line up to 2 GeV is identical to the collision electron case but the β continues to grow as the square root of the energy up to 3.11 GeV in this lattice. Using these parameters, the wakefield



Legend:

ws = wire scanner
blm = bunch length monitor
bbim = bunch-to-bunch intensity monitor
bbpm = bunch-to-bunch beam position monitor

Not shown:

Long ion chambers to detect beam loss located along the entire beam line from 40 MeV to 2 GeV.

12 beam loss monitors to detect losses from the gun to the pepperpot following the first bend.

A train beam position monitor for each quadrupole except for those wrapping around the accelerator.

A bbpm and bbim every 4 girders in the linac.

Many sets of dipole steering coils from the gun to the spoilers.

Figure 2-17. *Electron linac for the positron drive beam.*

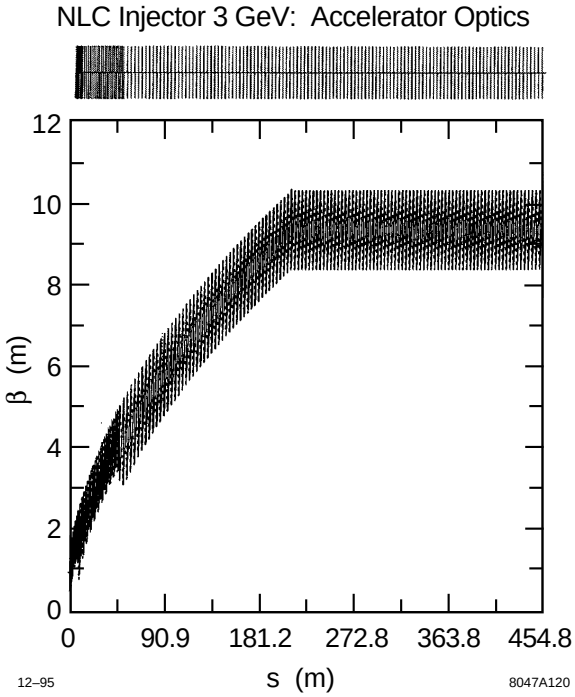


Figure 2-18. *Positron drive linac lattice to transport the electron beam from the end of the solenoids to the bend at the end of the drive linac.*

contribution to emittance growth was estimated for the 3.11-GeV linac in a similar way as for the 2-GeV collisions electron injector linac. The actual beam emittance of the last bunch increases by 11% at the end of the 3.11-GeV linac but the effective emittance increase due to the beam being offset by one sigma is 52%. This is still acceptable for the drive injector for positron production.

No special difficulties are expected with the positron drive injector which have not already been taken into consideration for the collision electrons.

2.7 Diagnostics

The purpose of the diagnostics system is several-fold: to aid in tuning the injector, to diagnose the beam parameters including jitter, and to be used for machine protection triggers. The diagnostics consist of charge monitors and beam position monitors some of which are capable of measuring the parameters for a small group of bunches in a portion of the train selected by the user, bunch-to-bunch timing monitor, aperture/pepperpot insert, beam profile monitors, bunch length monitor, emittance measurement and energy spread measurement stations after 80 MeV and 2 GeV. There is a rather high density of current monitors and position monitors in the beam line from the gun through 80 MeV. Every doublet in the injector linac has a beam position monitor in it. There will be microchannel-plate photomultiplier-tube beam loss monitors in the 120-keV region and a PLIC cable in the region from 40 MeV and beyond.

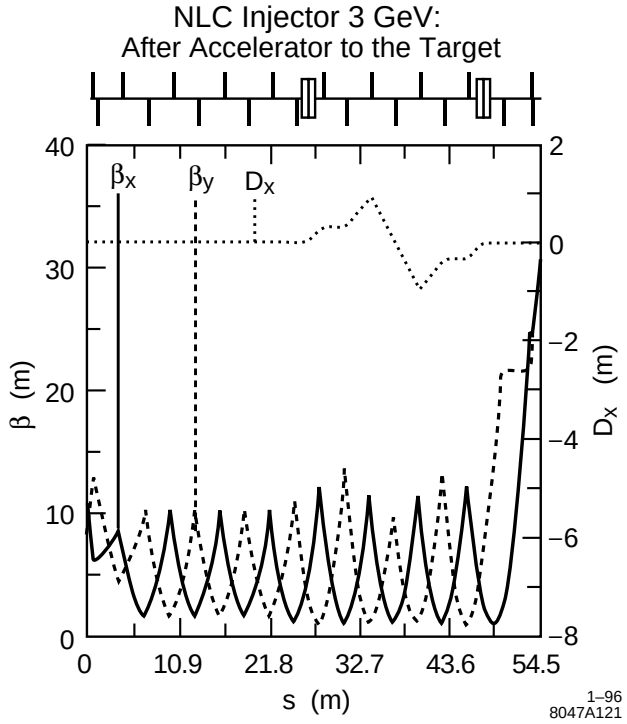


Figure 2-19. *Positron drive beam lattice from the end of the diagnostics station after the drive linac up to the spoiler.*

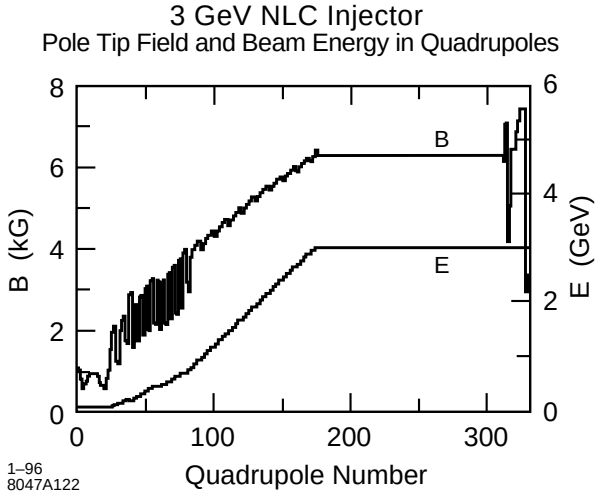


Figure 2-20. *3-GeV NLC injector. Pole-tip field and beam energy in quadrupoles.*

Diagnostic	Total Bunch Train		Bunch bundle[a]	
	Range	Resolution	Range per bunch	Resolution per bunch
Toroid charge mon.	45–250 $\times 10^{10}$ e^-	±0.2%	0.5–3 $\times 10^{10}$ e^-	±1%
Gap current mon.				
charge	45–250 $\times 10^{10}$ e^-	±3%	0.5–3 $\times 10^{10}$ e^-	±3%
pulse width	1–150 ns	±1 ns	∼200 ps	—
Streak Camera	1–150 ns	±1 ns	5–40 ps	1 ps
Beam loss mon.	0–10^8 e^-	10^7 e^-	—	—
Beam position mon.	0–1 cm	±20μ	0–1 cm	±20μ
Beam size mon.				
at the cathode[b]	R=0 - 1.5 cm edge	±1%	R=0 - 1.5 cm edge	±1%
at 120 keV	R=0 - 1.5 cm edge	±1%	R=0 - 1.5 cm edge	±1%
at 80 MeV	R=0.2 - 1 cm edge	±1%	R=0.2 - 1 cm edge	±1%
Energy	80 MeV	±0.2%	∼80 MeV	±0.5%
Energy spread at 80 MeV	∼4%	±0.2%	∼4%	±0.2%
Pepper pot	Qualitative diagnostic			
Bunch length monitor	Qualitative measurement for tuning by maximizing signal			

[a] Bunch bundle represents a group of bunches as close to a single bunch as possible. Some of the diagnostics may not be fast or sensitive enough to measure single bunch parameters but can measure a group of about five adjacent bunches and this would be sufficient. These diagnostics would require a timing gate width of about 3 ns, with a 0.5-ns rise time and 10-ps timing stability. The details of how to construct such diagnostics will be covered in Chapter 15.

[b] A profile of the laser at a screen located at the image point of the cathode.

Table 2-3. *Injector diagnostics and their specifications.*

Most of the monitors at the gun and at the end of the first accelerator section need to have the ability to measure variation from bunch-to-bunch within the same bunch train. Measuring the variation between a group of bunches in 5-ns intervals would be sufficient, that is, it is not necessary to measure the parameter of a single bunch.

Figure 2-1 shows the various kinds of monitors on the e^- injector beam line. The drive linac for the positrons will have these same diagnostics except the wire scanners for the linac emittance measurement will be located just upstream of the achromatic bends up to the two positron production lines. In addition there will be bunch-to-bunch beam position monitors, intensity monitors and wire scanners in the achromatic bends and just upstream of the spoilers as shown in Figure 2-17.

Table 2-3 shows the various diagnostics in the injector, and their specifications.

2.8 Operation

2.8.1 Initial Set-up and Beam Maintenance

Figure 2-1 shows a schematic of the injector beam line components from the gun anode up to 2 GeV. Each of the subharmonic bunchers, the S-band buncher and the S-band accelerator sections will have independent rf phase and amplitude control. There are beam steering coils at the gun and after the bend. Large steering coils are draped over

the solenoids from the subharmonic bunchers through the second accelerator section. There is a set of steering coils on each module in the injector linac.

The procedure for operating the NLC injector has several aspects. First of all it is necessary to set the focusing, bunching, and accelerating parameters of the injector to the optimum values for beam transport and quality. Then it is necessary to maintain the optimized beam quality achieved during startup over a long period of time. Finally it is necessary to establish guidelines for troubleshooting problems when the beam quality deteriorates. Additionally, it is necessary to establish good, clean configuration records to aid in recovering from various faults or down times.

To set up the injector parameters initially, it is necessary first of all to have good calibration data on all the beam diagnostics, the various power supplies, and rf drivers including the losses in cables and processing electronics. It is necessary to characterize the subharmonic bunchers and buncher and accelerating sections so that one can actually convert the power measured at the test points into electric fields as the beam sees it. Having calibrated the various components such that one can determine the state of the injector and the beam in it, we can then start the tuning procedure.

At first and when turning on the beam after long down times, it is necessary to establish a good quality, single-bunch beam all the way to the end of the injector before going to long-pulse operation.

The first order of business is to establish beam from the gun by shaping and pointing the laser on the photocathode and adjusting its power for the desired charge from the gun. Next, the beam is transported around the 20° bend to the subharmonic buncher by using the magnetic lens and solenoid strengths designated by the simulations and by steering. Once the beam reaches the subharmonic bunchers, it is necessary to synchronize the beam with the rf of the Subharmonic bunchers, the accelerators, the current monitor, and beam position monitor sampling cycles. Next, we need to bunch the beam and steer it to achieve at least 80% of the gun charge captured in 18 ps at 80 MeV. It might also be necessary to adjust the strengths of the axial magnetic field coils to achieve the simulation beam size at the profile monitor. Next, we steer the beam in the injector linac while phasing the klystrons to maximize the energy using the energy spread wire scanners both in the chicane at 80 MeV and 2 GeV. Once the desired quality beam is achieved at the end of the injector, all configurations and beam parameters should be archived for future use.

Now it is time to lengthen the train gradually without losing beam on the monitors all the way to the end of the injector. The subharmonic buncher timing and the phase of the compensating klystrons is adjusted slowly to maximize the signal on the bunch length monitor and to minimize the energy spread on the energy spread wire scanners. Once the long bunch train configuration is set, we can then adjust the intensity jitter limiting aperture for the optimal size. This process also has to be done slowly, adjusting the timing of the subharmonic buncher and the power in the compensating klystrons at the same time. Next we can use the scrapers in the 80-MeV chicane to clip away the tails. More adjustment of the power in the compensating klystrons will be necessary.

Configuration files showing the orbit and charge transmission of the beam throughout the injector and all the device settings will be saved. In addition, the bunch length, the energy and energy spread, the emittance of the beam at 80 MeV, and at the entrance to the damping ring should be measured. This information is important for expediting the accelerator setup in the future and maintaining the beam quality in a consistent way.

To maintain the beam quality day-to-day, various steering and energy feedback systems will be used. RF and magnet parameters, as well as beam parameters, will be compared to the parameters in the best saved configuration files and flagged if they are off by more than the allowed tolerance.

2.8.2 Troubleshooting

Some of the typical things that can go wrong in the injector are that the energy spread of the beam downstream, for example, at the damping ring is too large. If this spread cannot be reduced to an acceptable size by phasing the accelerators, most likely the phases of the bunching components have drifted away from optimum. Check their phase and amplitude against the saved optimized configuration and make the necessary changes. It is important to simultaneously maximize the signal at the bunch length monitor

Another problem could be the increase in energy jitter. This usually stems from rf-amplitude jitter in the accelerator or bunching rf sources and should be fixed by the rf technicians.

Another problem could be the increase in intensity jitter. This could stem from a number of reasons. Check the intensity jitter of the laser as it will be running on the edge of available technology as far as intensity jitter is concerned. Another cause for intensity jitter is when the beam is large compared to the various apertures and is not centered, thus is being scraped. One can check this with the various screens and PLIC signals. The most effective knobs for solving this problem are the small lenses near the gun and after the bend and the second most effective are the steering coils.

It is impossible to enumerate all the possible problems which can occur in the injector area, but we have addressed the most common ones.

2.9 Conclusion

The NLC polarized source injector will be a conventional injector employing subharmonic bunching to achieve the required bunch intensity and structure. Simulation results show that it is possible to capture 83% of the charge into the subharmonic buncher system into 18-ps bunches when a train of bunches are produced from the gun which are 700-ps wide and 1.4-ns apart. The total energy spread at 2 GeV is expected to be $\pm$0.6%, well under the $\pm$1% threshold. The emittance, including wakefield effects, is expected to be less than 0.6×10^{-4} m-rad, well within the about 1×10^{-4} m-rad threshold. The charge intensity jitter limiting aperture is able to reduce 2% incoming intensity jitter to an immeasurable amount at the cost of losing 17% of the charge from the gun. Using the Δt compensation technique in the injector linac, we can reduce the bunch train energy spread without any effect on the single-bunch energy spread at the end of the injector.

The injectors for both the collision and positron drive electron beams are almost identical up to 2 GeV with the exception that the positron drive injector will use a thermionic gun. This should allow us to reduce construction costs and operational complications.

The NLC injector scheme is similar to the current SLC injector with the added complication of working with a long bunch train instead of a single bunch. There are several issues regarding this added complication which deserve a hard look either in detailed simulations and or demonstrations in a laboratory environment. The R&D programs needed to ensure us of success with the NLC injector are: 1) The cathode charge limit as addressed in Appendix 2.B. 2) The beam-loading compensation in the bunching system. 3) Demonstrate the production of the 1.4-ns-separation bunch train in the thermionic gun. 4) Development of a reliable and stable polarized source rf gun to produce small emittance beams from the injector, thus simplifying the operation and maybe the design of the damping ring. Item 1) can be studied in the SLC polarized source laboratory. Item 2) and 3) will be demonstrated in the NLCTA upgrade. There are several options for demonstrating item 4). One of these options would be to use the rf gun test stand being constructed at SSRL.

Despite the differences between the SLC and NLC sources, experience with the SLC can be applied and extended for a reliable and stable operation of the NLC injector.

2.A Polarized e^- Beam Photocathode RF Gun Development for the NLC

2.A.1 Introduction

While the baseline injector design for the NLC polarized electron source is presently a conventional polarized DC gun with a subharmonic bunching system, the development of a low-emittance polarized electron beam rf gun in parallel, as a possible future upgrade, is important to simplify the damping ring design and/or operation.

Polarized e^- beam photocathode rf gun injectors are an attractive alternative to the conventional electron source for the NLC because of the possibility of achieving very low electron beam transverse emittance and thus making it possible to reduce the size of the damping ring or at least making it easier to operate. Simulation results show that it is possible to achieve normalized rms emittance of $\epsilon_{n,\mathrm{rms}} = 1$ mm-mr for a 1 nC bunch, 10 to 20 ps in width [Gallardo 1993, Palmer 1995b, Sheffield 1993]. Some experiments have been conducted at LANL and BNL to achieve beams which are very similar in quality to these simulation results. The low emittances are achieved due in part to the emittance compensation technique discussed in reference [Carlsten 1989]. The NLC requirements for the beam from the rf gun are about 4 nC in a 10 to 20-ps bunch 1.4-ns apart in a train of 126 ns. The normalized, rms emittance requirements at the input of the current NLC damping ring design is $< 10^{-4}$ mm-mr rms at the damping ring. The emittance damping time in the ring depends on the emittance of the incoming beam. For each order of magnitude reduction in the input beam emittance, the number of bunch trains to be simultaneously stored in the ring can be reduced by 1, thus also reducing the size of the ring. Even a factor of 2 or 3 reduction in emittance would be helpful in improving the operation of the damping ring even if its design is not simplified.

Photocathode rf guns have shown themselves to be a stable and reliable source for the FEL community [Travier 1994]. In the NLC, Cylindrically Symmetric Emittance-Compensated Polarized RF Gun Development program, the work of the FEL community would be extended to the development of a polarized source, emittance-compensated, ultra-high-vacuum rf gun. This new gun design would address the issues associated with polarized rf guns some of which are discussed in reference [Clendenin]. The main technical issues to resolve for polarized rf guns which are more difficult than for the DC guns are: 1) A stable and reliable pulsed laser system to produce the electron bunch train, 2) the beam dynamics associated with producing a low-emittance electron bunch train including the possible depolarization effects due to strong space charge and transverse magnetic fields near the cathode, 3) elimination of field asymmetries in the gun which will cause emittance growth, 4) the ultra-high-vacuum environment which is difficult to achieve in rf guns but is required by the GaAs cathodes for long lifetimes, 5) the effect of high-gradient fields on GaAs cathodes, 6) cathode issues such as promptness of the photoemission, cathode lifetimes and quantum efficiency, and 7) finally, the reliable and efficient operation as a injector system.

A development program for an asymmetric beam rf gun is also important, since the emittance required at the interaction point is two orders of magnitude smaller in the vertical plane than in the horizontal plane. An asymmetric beam could be produced at the cathode of the rf gun by the production of an asymmetric transverse laser pulse. This program would initially proceed with the unpolarized source, emittance-compensated rf gun, since the flat-beam emittance compensation has not yet been experimentally demonstrated.

2.A.2 Beam Dynamics

In this section we shall discuss two parallel R&D efforts that will elucidate two areas of accelerator physics and in the end allow the combining of these developments to produce a flat-beam polarized rf gun.

The study of negative electron affinity (NEA) GaAs in high-gradient fields, along with the development of a dedicated cylindrically symmetric, polarized-source rf gun facility is one of these development programs.

Intensive programs to study and demonstrate low-emittance, high-brightness beams from symmetric photocathode rf guns is ongoing at LANL, BNL, SLAC, UCLA and other laboratories abroad. The best available symmetric rf gun emerging from these studies should be used for studying the survivability of polarized e^- source cathodes in rf guns. This will allow for the study of vacuum conditions, field amplitudes, and cesiation processes that are necessary for the production of polarized electrons without the additional problems of flat-beam production.

In another development program for the flat electron beam production, the physics of emittance compensation of a flat-beam and rf field uniformity can be addressed without the added difficulty of polarized electron production.

Round Beam

PARMELA simulations of injectors for FELs show that it is possible to produce an electron beam with a transverse normalized emittance of $\epsilon_{n,\mathrm{rms}} = 1$ mm-mr. The beam parameters from these simulations are listed in Table 2-4 and shown in Figure 2-21. Simulation studies of L- and C- band systems are ongoing using the S-band system as a baseline [Kirk 1995].

The emittance compensation scheme realigns different "slices" [Sheffield 1992] of an electron bunch. Depending on the cathode spot size, total charge per bunch, and peak field at the cathode, the precise position of the compensation solenoid [Palmer 1995a] is critical. Figure 2-22 shows this for the parameter set in Table 2-4. In the physical construction of the gun and solenoid magnet, the waveguide physically limits the position of the magnet with respect to the cathode of the gun. This problem has been corrected in the proposed L-band rf gun being designed for the TESLA X-ray FEL [TESLA FEL]. It uses a "door-knob" rf coupler downstream of the full cell, eliminating the physical constraint of the waveguide and maintaining cylindrical symmetry. As a result, this scheme eliminates the higher-order spatial harmonics that can also cause emittance growth, and allows for the precise positioning of the compensation magnet. The tuners and vacuum connections to the gun will also physically limit the magnet position, but these problems can be eliminated by longitudinal connections to the gun body versus radial coupling used in the previous gun designs.

Asymmetric Emittance RF Photocathode Electron Sources for the NLC

While the expected performance of cylindrically symmetric rf gun photo-injector systems is well understood from theoretical analysis and simulation, which have been benchmarked by experiment, an extension to a fully three-dimensional approach has not until recently been undertaken by investigators at UCLA and Fermilab. This effort has been motivated by the demands of the superconducting TESLA linear collider design, in which the normalized emittances are $\epsilon_{x,y}$=20, 1 mm-mr at Q=8.3 nC. In this machine, it is possible to have asymmetric emittances at the injection point of an electron linac which meet the constraints set by the interaction point. While the emittances may not be low enough to eliminate the need for an electron damping ring in the NLC designs under consideration here, benefits in designing and operating the damping ring may be derived from injection of lower, asymmetric-emittance electron beams.

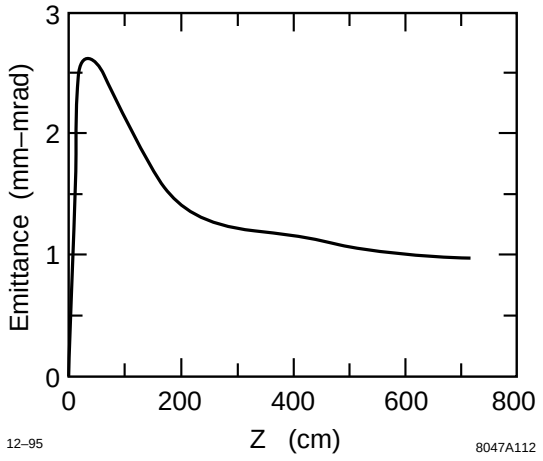


Figure 2-21. *PARMELA simulations of emittance compensation for the BNL/SLAC/UCLA S-band rf gun*

Total Charge	1 nC
Number of particles	10K
Cathode Spot Size	1 mm radius
Longitudinal Profile	Flat-top
Transverse Profile	Flat-top
Initial Cathode KE	.5 eV
Initial Thermal Emittance, ϵ_o	0 mm-mr
$E_{Full\ Cell}/E_{Half\ Cell}$	1.00
$E_o\ at\ Cathode$	160 MV/m

Table 2-4. *Electron bunch parameters used in PARMELA.*

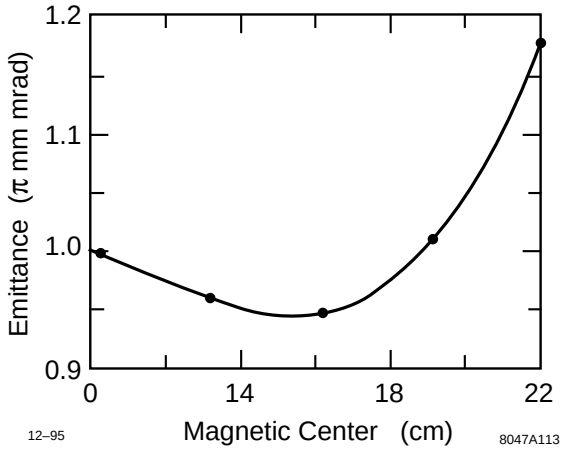


Figure 2-22. *Optimized emittance versus magnetic center.*

To state simply, the approach to design an asymmetric rf photoinjector source [Rosenzweig 1993], one uses the fact that the temperature of the emitted beam particles from a photocathode is only a function of cathode material and laser photon energy. This allows the creation of asymmetric emittances $\epsilon_x \gg \epsilon_y$ by illuminating the photocathode with a ribbon laser pulse ($\sigma_x \gg \sigma_y$). Typically, the dominant emittance growth mechanism at relevant current densities is space-charge, and this is conveniently mitigated by the ribbon geometry of the electron beam pulse. One must also pay attention to the potential for rf-induced emittance growth, especially in the horizontal dimension where the beam is large. In the spirit of the problem for cylindrically symmetric systems performed by K.J. Kim [Kim 1989], we have examined the scaling of emittances with respect to beam dimensions, accelerating gradient and phase, and beam charge. The "thermal" contribution to the emittance given by the finite spread in transverse momenta of photo-emitted electrons is, assuming 0.1 eV emission temperature (typical of semiconductors),

$$\epsilon_{x,y}^{th} \cong \sqrt{\frac{kT_\perp}{m_e c^2}}\sigma_{x,y} \cong 0.25\sigma_{x,y}\,(\mathbf{mm})\ \text{mm-mr} \tag{2.4}$$

For TESLA, this implies beam sizes of $\sigma_x < 4\,$cm, and $\sigma_y < 2\,$mm, while for the NLC requirements we must have $\sigma_x < 0.4\,$cm, and $\sigma_y < 0.2\,$mm. This is a potential challenge for the NLC cases, because a $Q = 4\,$nC beam of such a small cross-section implies an accelerating field needed at the cathode to overcome longitudinal space charge of E>72 MV/m at injection. Under the assumptions of a cylindrically symmetric rf cavity structure, it is easy to derive the minimum rf contribution to the rms emittances, which can be written as

$$\epsilon_{x,y}^{rf} \cong \frac{eE_{rf}}{\sqrt{8}m_e c^2}\sigma_{x,y}^2\left(\Delta\phi\right)^2 \tag{2.5}$$

where $\Delta\phi = \left(\frac{\omega}{c}\right)\sigma_z$ is the phase extent of the beam. Minimizing these quantities implies operation at low frequency, low accelerating gradient, and with small bunch sizes in all dimensions. For our case, we are considering a fairly large horizontal beam size as well as high gradient, and so this is an issue for the horizontal emittance.

There are a number of proposed ways to remove or minimize the emittance-diluting phase-dependent defocusing due to the transverse rf forces in the gun. Because of the use of emittance compensation to deal with the space-charge contribution to the emittance, it is perhaps wisest to use a scheme which mitigates the transverse kick in the x direction. This can be accomplished by removing as best as possible the dependence of the rf fields on x. It should be noted in this regard that the defocusing kick induced at the end of a cylindrically symmetric cavity is approximately equal parts electric and magnetic. The electric component of the kick is mainly due to the fringing fields near the iris at the exit of the gun. The fringe field can be made mainly vertical simply by making the iris opening a slit, long in the horizontal dimension.

The magnetic component, however, cannot be so easily diminished, as it is not dependent on the iris, but on the accelerating field in the interior of the cavity. The only way to diminish the magnetic force in the horizontal direction is to break the symmetry of the cavity outer wall. The simplest asymmetric structure is a rectangular box cavity, but this only diminishes (does not eliminate) the horizontal rf forces, and produces a sinusoidal dependence of the accelerating field on the x dimension. A better choice is to use an "H-shaped" structure (as suggested by R. Miller), which has been investigated in the context of sheet-beam klystron development at SLAC. In this structure, the accelerating field in the "bar" region of the H-cavity has no dependence on x and therefore has acceleration independent of transverse position, as in a symmetric structure. It should be noted that the vertical forces (which are due mainly to the backward speed-of-light space harmonic in this p-mode structure), in this case in the bar region, are now entirely in the vertical dimension, and are twice as large as the equivalent symmetric cavity. The function of the side regions is to allow the longitudinal field to go to zero sinusoidally, to satisfy the boundary condition, and choose (along with the vertical dimension) the cavity resonant frequency.

An analysis of the space-charge contribution to the emittance gives the following scaling, derived from both approximate analysis and PARMELA simulation [Rosenzweig 1995]:

$$\epsilon_{x,y}^{sc} \approx \frac{2N_b r_e}{7\sigma_{x,y}W}e^{-3\sqrt{W\sigma_y}}\sqrt{\frac{\sigma_y}{\sigma_z}} \tag{2.6}$$

$$W = \frac{eE_{rf}}{2m_e c^2} \sin(\phi_o) \tag{2.7}$$

Here, $\epsilon_x^{sc} \approx \frac{\sigma_x}{\sigma_y}\epsilon_y^{sc}$ is approximately obeyed; this is a consequence of the fact that for an ellipsoidal charge distribution, the normal electric field at any bunch boundary is approximately the same as any other. It should be noted that minimizing the space-charge contributions to the emittances implies operation at high accelerating gradient. It should also be emphasized at this point that these emittances, induced by the space charge at low energy, are mainly due to the different orientations of each "z-slice" of the beam. A scheme is discussed below, emittance compensation, which can remove these correlations, effectively lowering the emittances of the bunch. Before taking up this discussion, it should also be pointed out that one cannot arbitrarily reduce σ_y, because of the limit on surface charge density emitted from the cathode before the retarding space-charge cuts off emission ($\sigma_x\sigma_y > 2\frac{Q}{E_z}$).

This effect has been observed at UCLA. It is instructive to look at the product of the emittances, to see if there is some chance of achieving the desired emittances without the need the for emittance compensation

$$\epsilon_x^{sc}\epsilon_y^{sc} \approx \left[\frac{2N_b r_e}{7W}\right]^2 \frac{e^{-6\sqrt{W\sigma_y}}}{\sigma_x \sigma z} \tag{2.8}$$

The argument in the exponential in this expression is generally smaller than one for reasonable accelerating gradients, and so one can see that, in general, a beam large in the longitudinal and horizontal dimensions is needed to minimize the product of the space- charge emittances. Unfortunately, these are precisely the dimensions that the rf contributions are sensitive to, and thus one cannot consider making them arbitrarily large. As an example of a design where the rf contribution is not yet significant, we take $N_b = 5 * 10^{10}$, W=80, $\sigma_z = 2$ mm, $\sigma_x = 2$ cm, $\sigma_y = 1$ mm. In this case, we have $\epsilon_x^{sc}\epsilon_y^{sc} \approx 450$, as opposed to $\epsilon_x^{sc}\epsilon_y^{sc} \approx 20 \sim 50$ as needed by TESLA, and we miss the design goal by an minimum of three to four in both dimensions. For the NLC case, the degree to which the design goals will not be met without emittance correction is even larger.

It is apparent that further efforts must be made to achieve the emittances needed for linear collider designs; that is, one must employ an emittance compensation scheme for sheet beams accelerated in asymmetric rf structures. Emittance compensation is essentially a process by which the correlations (mainly as a function of the longitudinal position in the bunch) in the beam's transverse phase space that are induced by space-charge forces operating at low energy (in the photocathode gun), are removed by focusing the beam, reversing the direction of additional correlations. When the orientation of the transverse phase space ellipse of each "z-slice" has the same angle, the linear component of the emittance due to this effect is removed. Since the space-charge forces must still be large enough after the focusing lens to allow for compensation, this scheme tends to work better at lower accelerating gradients for convenient focusing geometries. This happy direction in the design of the system allows us to mitigate concerns about the rf contribution to the emittance, and about peak and average rf power (source and heat dissipation problems), in both the symmetric and asymmetric cases. In the asymmetric case, we must have the rf contribution to the transverse phase space trajectories be small, or else the space-charge compensation will suffer from interference. This condition is achieved by the choice of large aspect ratio (sheet) beam profile, and the asymmetric structure described above.

While the serious design calculations must be performed with simulation programs which include as many real experimental effects as possible, these are very time consuming. One needs to have a model with which to predict the approximate behavior of the system before proceeding to simulation. This has been developed; it is a program which integrates the envelope equations for a number of different "z-slices," at positions. The focusing can arise from conventional quadrupole focusing, or from the first order transient kicks It is apparent that further efforts must be experienced by electrons at the exit and the entrance of the rf cavities, and the second order (alternating gradient focusing) in the interior of the rf cavities. Solenoidal focusing is not allowed in this device, because of a coupling of the x and y phase planes due to an $\vec{E} \times \vec{B}$ rotation which is, of course, dependent on longitudinal position in the beam. The emittance terms included in the envelope calculations are only the thermal components.

The results of a typical calculation are shown in Figure 2-23. In this case, the vertical emittance is well compensated, while there is essentially no improvement in the horizontal emittance. This is because the effective defocusing strength associated with the space charge is different in the two dimensions, as can be seen from the envelope equation

$$K_{x,y}^{sc} = \frac{4I\left(z_i\right)}{I_o\gamma^3\left(\sigma_x+\sigma_y\right)\sigma_{x,y}} \quad . \tag{2.9}$$

Because of this, the compensation process generally proceeds much faster in the small dimension, and thus it is difficult to design a system which simultaneously compensates in the vertical and horizontal dimensions.

The solution to this difficulty is straightforward: since the beam near the cathode is very much larger in x than in y or z, one can effectively remove the dynamics in the x direction (and the dependence on x of the y and z components of the space-charge forces), as is already done for the rf contributions, from consideration by making the beam distribution uniform in x. The effect of this is illuminated by Figure 2-24, which displays the x component of the force along the x axis of the bunch, and the y component of the force at one sigma in y, in the y-z plane, for two cases: (a) where the beam distribution is Gaussian in all three dimensions, and (b) where it is Gaussian in y and z, but uniform up to a hard boundary in x. It can be seen that for the tri-Gaussian beam that the horizontal forces rise approximately linearly to a large value at about one sigma in x. For the uniform distribution in x, however, there is very little horizontal field over about 85 to 90% of the beam, and the field rises steeply and nonlinearly near the beam boundary. As was noted before, the maximum field is nearly the same, but the region of the beam which is affected is much smaller. Note that the vertical field is also nearly uniform over almost all of the beam, again degrading near the beam edge. In practice, one uses this final 10 to 15% of the beam as "guard charge" during the compensation to remove the horizontal field and homogenize the vertical field, then removes it by collimation after the linac section. In this way, one effectively removes the x dimension from the compensation problem, reducing the problem to two dimensions, as in the (by now well understood) cylindrically symmetric case. It should be noted in this regard that the multiple-envelope model for emittance compensation dynamics has been compared extensively to simulation as well as extended analytically by Serafini and Rosenzwieg [Serafini 1995], and essentially validated for cylindrically symmetric beams. Benchmarking of the models in the asymmetric case awaits further refinements of three-dimensional computational tools.

The expected emittances in the case of a 2856-MHz, 1.5 or 4-nC can be scaled from the previously investigated 1300-MHz, 10-nC, 35-MV/m rf gun TESLA case, where the emittances are $\epsilon_{x,y} = 30, 2$ mm-mr. The beam dimensions, charge, electromagnetic fields, and emittances all scale linearly with rf wavelength, while scaling of charge at a given wavelength scales the emittances approximately linearly with the charge. Thus we expect that a 77-MV/m rf gun at 2856 MHz should produce emittances of approximately $\epsilon_{x,y} = 4.5, 0.3$ mm-mr at 1.5 nC and $\epsilon_{x,y} = 12, 0.8$ mm-mr at 4 nC.

2.A.3 Material and RF Considerations

Vacuum

It has been the SLC polarized-source experience that for successful and reliable operation, the vacuum in the vicinity of the cathode needs to be extremely low, on the order of 10^{-11} Torr and the partial pressure of some diatomic gases such as CO must be maintained in the mid 10^{-12} range [Schultz 1992]. The ultra-high-vacuum system is a critical part of making the polarized-electron rf gun work, not only because of the cathode but also to prevent any possible rf arcing in the gun, thus damaging the cathode. Some great progress has been made in Japan in selecting appropriate copper and treating it with pressurized ultrapure water to achieve vacuum levels of 2×10^{-10} Torr in an S-band standing-wave cavity. The cavity was excited with a 2856-MHz klystron, and surface gradients of 337 MV/m were achieved

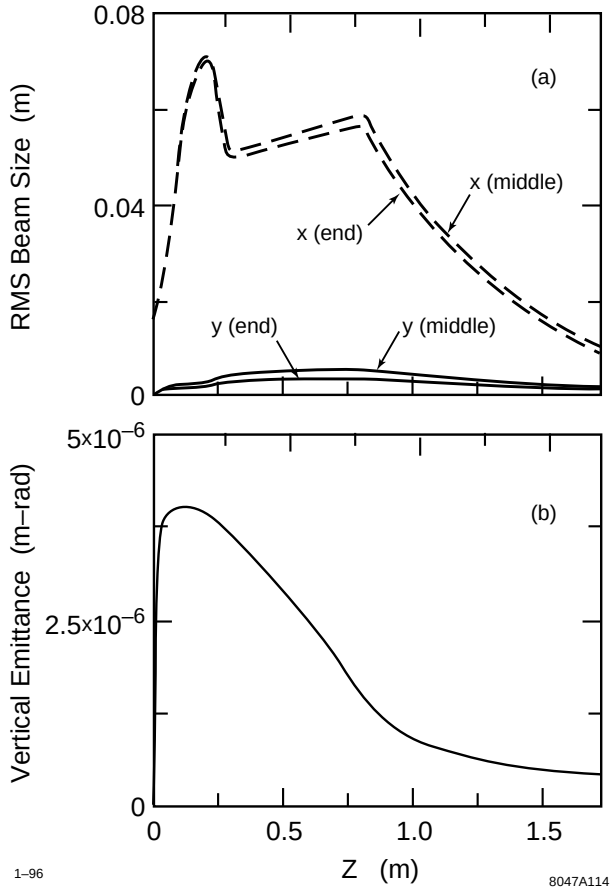


Figure 2-23. *(a) Rms beam sizes for two beam slices in multiple-slice model calculation of emittance compensation for 10-nC beam. A vertically focusing quadrupole is placed symmetrically over the cathode plane, followed by a horizontally focusing quadrupole, another doublet, a drift, and a TESLA cavity linac. (b) Vertical emittance evolution in this case.*

without arcing or any rise in the pressure in the presence of rf [Yoshioka 1994, Matsumoto 1994]. Demonstration of the survivability of the polarized cathode in the rf gun is needed.

It has been proposed that the new 1.6-cell rf gun designed by the BNL/SLAC/UCLA collaboration be constructed out of High Isostatic Pressure(HIP) processed Oxygen Free High Conductive (OFHC) Cu [Palmer priv] to minimize the vacuum problems associated with a high-gradient gun. This gun has been symmetrized to remove the dipole-like spatial harmonic of E_z. It has been proposed that future guns of the BNL/SLAC/UCLA collaboration have their quadrapole-type fields symmetrized thereby eliminating the next higher-order mode emittance growth factor. This would allow for more pumping on the gun and thereby reduce the ultimate vacuum limit.

Assuming that the out-gassing rate of copper is 10^{-12} Torr $\ell/$ s cm^2, an estimate of the needed pumping speed can be calculated. The total surface area of a state-of-the-art, BNL-type, S-band rf gun is on the order of 420 cm^2. Therefore this type of rf gun will need 42 $\ell/$ s of total pumping speed to maintain the gun vacuum on the order of 10^{-11} Torr. This amount of pumping does not take into account the conductance-limiting effects of the waveguide to full-cell coupling iris.

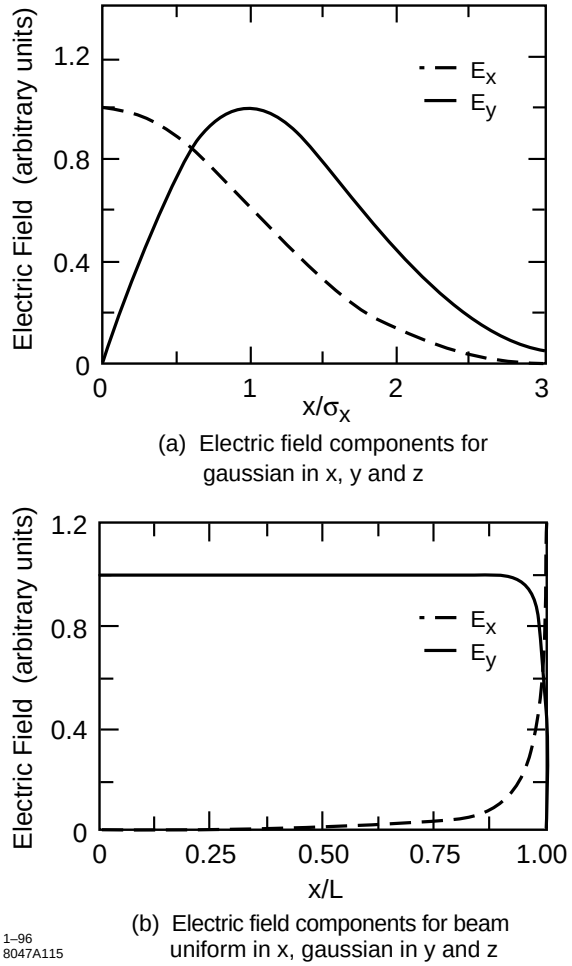


Figure 2-24. *Electric field components for (a) uniform, as opposed to (b) horizontally-Gaussian density profile.*

The polarized source cathode will need to be recesiated periodically and activated less occasionally. Both of these procedures are a threat to the required ultra-high-vacuum environment. A scheme involving a "loadlock" where the cathode can be retrieved and isolated from the gun vacuum environment to carry out these procedures needs to be incorporated into the gun design. Another possible option might be to mask the cathode with a "shroud" which is part of the cesiation or activation mechanism lowered directly into the gun. Further design studies to pin down the optimum way to carry out the cesiation and activation procedures is essential.

High-gradient Effects

The effects of high gradients in the vicinity of the GaAs cathode are best studied with a specially designed HIP copper rf gun whose design allows for a 42 ℓ/s pumping speed. This design should incorporate an rf feed that will eliminate the dipole-field asymmetry and allow for the mechanical positioning of the emittance compensation magnet in its optimal position. The properties of the various cathode materials will detune the half-cell resonant frequency and thereby effect the field balance of the gun. Symmetric tuners in both the full and half cell will be needed to maintain

the $f_\pi = 2856$ MHz with the interchange of different cathode material under ultra-high-vacuum conditions. Also, the field levels in the full and half cell will need to be monitored during operation. The mechanical and rf design of these coupling probes should not induce dipole asymmetries in the cavities which would cause emittance growth in the beam.

To maintain the high gradient in the gun necessary for the emittance compensation at S-band, the mechanical design should take into account the scheme discussed in Section 2.A.3, "Vacuum", for the recesiation process of NEA photocathode materials.

There is a plan to install a GaAs cathode from SLAC in the rf gun at CERN and to study survivability in the presence of rf. The success of this test would be important in building confidence in the feasibility of polarized-source rf guns, but failure would not necessarily rule out the possibility of using a polarized rf gun for the NLC since the CERN Linear Collider Test Facility (CLIC) gun is not specifically designed for this test.

It is important to pursue a dedicated effort to carry out the full measure of high-gradient studies at SLAC. This would entail the design and construction of a new rf gun which includes the necessary upgrades discussed above to make the present BNL/SLAC/UCLA 1.6-cell S-band rf gun a polarized electron source.

Photocathode Response Time

In order for an S-band rf gun to produce a beam with a bunch length that meets the NLC injector requirement, *i.e.,* not to exceed 20 ps in full pulse length, the response time of the photocathode must be $\leq$10 ps. For a thin strained GaAs photocathode whose surface is treated to have negative electron affinity, the photoemission response time is determined by the thickness of the active layer, which is typically 100 nm, and by the electron diffusion coefficient. With a doping density in the mid-10^{18}-cm^{-3} range, the diffusion coefficient in GaAs is about 35 cm^2/ at room temperature [Sze 1981], Thus, in the absence of multiple emission attempts, the response time for a 100-nm-thick cathode is estimated to be approximately 3 ps, which adequately meets the requirement for response time. However, this estimate ignores the likely scenario in which an excited electron may require several attempts to be emitted as it may fail to escape the first time or even the first few times it reaches the cathode surface and, as such, must be regarded as a lower limit on the response time. Electrons that are emitted after multiple escape attempts will ultimately lead to a longer response time and may prevent negative electron affinity photocathode from being incorporated into an S-band rf gun.

Measurements of the response time for GaAs photocathodes have been made only with bulk samples [Aleksandrov 1995, Hartmann 1995a]. Aleksandrov *et al.,* placed an upper limit of 40 ps on the response time, whereas Hartmann *et al.,* found it to be about 30 ps. However, both measurements were done under the condition that the cathodes' quantum efficiencies were more than an order of magnitude lower than that of an optimally activated bulk GaAs, implying that the essential property of negative electron affinity is or close to be lost. Therefore, the measured response time is actually that of a thin active emitting layer at the surface of the bulk GaAs cathode, whose thickness is probably on the order of 100 nm due to the lack of negative electron affinity, rather than that of the bulk GaAs cathode. In this sense, these two studies may represent an experimental approximation for measuring the response time of a 100 nm strained GaAs cathode. Of course, since the response time depends quadratically on the thickness of the active emitting layer and it is difficult to estimate its exact value given a bulk GaAs cathode with a low quantum efficiency, the measured response time should be regarded only as an order of magnitude estimate for a 100-nm strained GaAs cathode. Ultimately, the exact response time for such a cathode must be determined experimentally under a high quantum efficiency condition appropriate for high beam intensity operation. At present, it appears that the response time of such cathodes may marginally satisfy the requirement of an S-band rf gun.

Photocathode Charge Limit

For a thin strained GaAs photocathode to be used in an rf gun for NLC, it must be capable of producing up to 3 nC of charge for an approximately 20-ps-long bunch, which implies that the charge limit of the cathode for a 20 ps charge pulse needs to be $\geq$3 nC. Earlier charge limit studies at SLAC has shown that the cathode charge limit in a 2-ns pulse depends almost linearly on the cathode bias or equivalently the extraction electric field at the surface within a range from 0.15 to 1.8 MV/m [Tang 1994a]. In an rf gun, the extraction field at the cathode surface is on the order of 100 MV/m, which is about two orders of magnitude higher than that in an SLC polarized electron gun operated at 120 kV. If the linear scaling with respect to extraction field remains valid up to orders of magnitude higher fields, then, the charge limit of a similar cathode in an rf gun will be scaled higher by the same factor. The maximum observed charge limit for a 2-ns charge pulse from a 100-nm strained GaAs cathode with an 80% polarization in an SLC gun was about 9.5 nC/cm^2 at an extraction field of 1.8 MV/m. Scaling to 100 MV/m yields the charge limit for a 2-ns pulse in the rf gun to be about 530 nC/cm^2. Based on the fact that the typical size of the laser beam used to illuminate the cathode in an S-band rf gun is 1-mm in radius, the maximum extractable charge from such a cathode for a 2-ns pulse is estimated to be approximately 17 nC.

It remains to estimate the charge limit in the rf gun for a 20-ps pulse. Unfortunately, no systematic studies on the effect of the pulse length, especially in the picosecond regime, on the charge limit have been made. Observation made during earlier R&D experiments of the SLAC polarized electron source has shown that the charge limit does depend, albeit not strongly, on the pulse length. For example, the charge limit for a 200-ns pulse was observed to be more than five times greater than that of a 2-ns pulse in an SLC gun. Physics consideration suggests that from 2 ns to shorter time scales the charge limit should depend on the pulse length more weakly. Thus, given the estimated charge limit of 17 nC in an rf gun for a 2-ns pulse, the prospect of achieving $\geq$3 nC in a 20-ps pulse appears reasonable. Plans have been made to systematically study this dependence for pulse lengths ranging from several nanoseconds to sub-nanosecond scales to further our understanding of this important issue. Finally, it is worth noting that the extremely high extraction field in an rf gun will greatly mitigate the multi-bunch effect of the charge limit phenomenon.

2.A.4 Laser System

Energy Requirements

The laser system for an rf gun for NLC, which is shown in Figure 2-25, must produce a train of 90 pulses separated by 1.4 ns, corresponding to a frequency of 714 MHz. This 126-ns pulse train (the macropulse) repeats at 180 Hz. Each of these 90 micropulses in the train must have a width adjustable between 5 and 20 ps, so that the emittance can be minimized. This optimization will also involve adjusting the temporal pulse shape (to be discussed below). Each pulse must be sufficiently intense to generate a charge of up to 4 nC (2.5×10^{10}) electrons. To obtain the high polarization (80%) now achieved with the SLC's DC cathode, a strained GaAs cathode would again be used, with an excitation wavelength near 850 nm. For this material and wavelength, a quantum efficiency of 0.1% has been measured in the DC gun. In an rf gun, the efficiency could be higher due to enhancement by the high rf electric field on the cathode, or lower if the rf causes degradation of the surface. Assuming $QE = 0.1\%$, the laser must produce a train of pulses on the cathode with at least six μJ/micropulse and 0.5 mJ for the macropulse. With a safety factor of eight to allow for losses in the optics and declines in quantum efficiency, the laser must be capable of delivering up to 4 mJ at the output of the laser pulse compressor over the 90-pulse train.

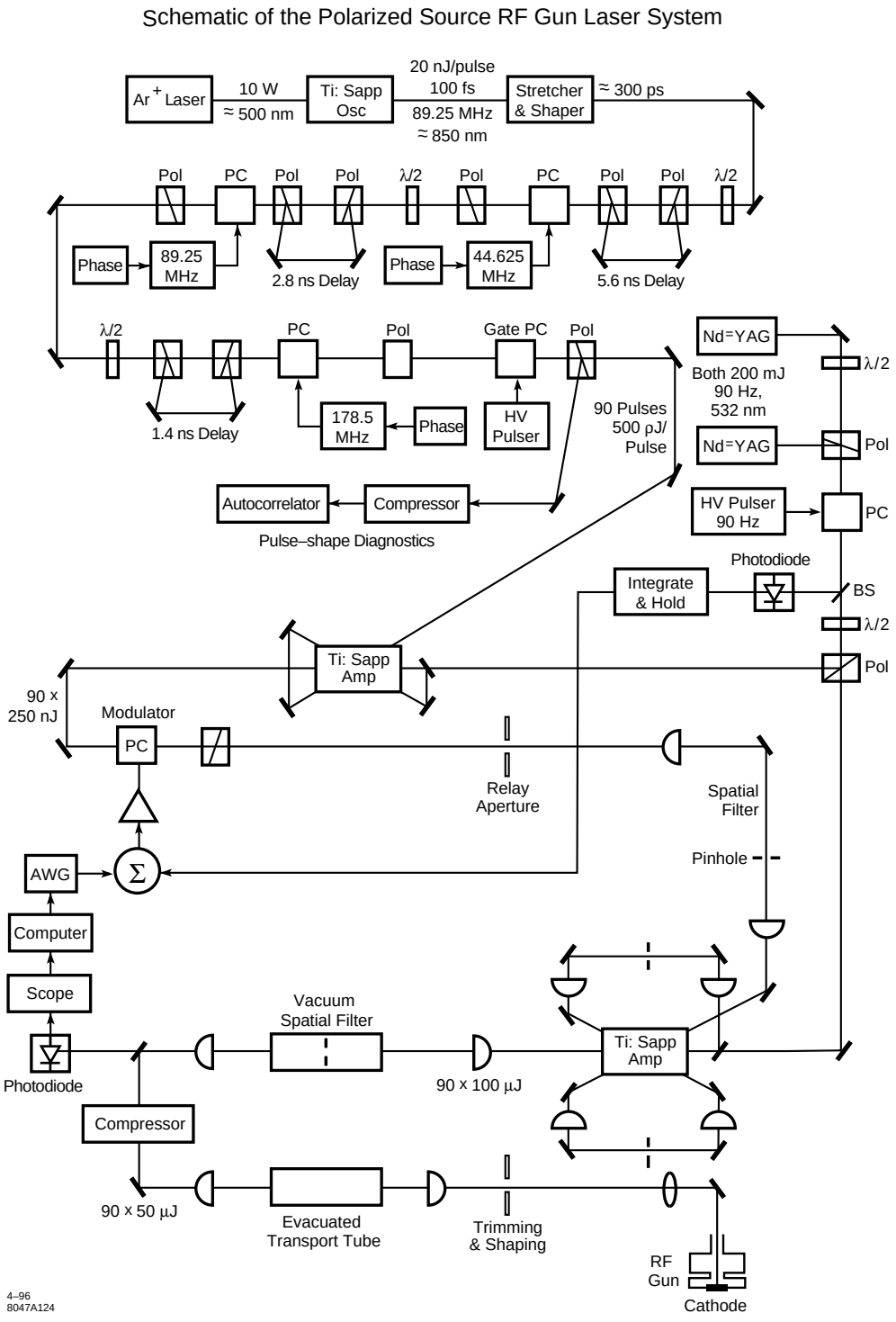


Figure 2-25. *Schematic diagram of the polarized-source rf gun laser system.*

Stability Requirements

A high degree of stability in several laser parameters is essential for the NLC. Variations in the energy per pulse, both within one macropulse and from macropulse to macropulse, should be limited so that the charge per pulse is constant. Charge fluctuations vary the beam loading in the accelerator sections, and hence introduce additional energy spread in the beam. Allowable intensity jitter budgets are discussed in Chapter 1. The allowable macropulse charge intensity fluctuation needs to be less than 0.5% rms. Thus for the NLC, we require a variation of no more than 0.5% rms in the macropulse laser energy and in the overall flatness of the envelope of the micropulses in each macropulse. Fluctuations of 2% are tolerable from one micropulse to the next.

In an rf gun, the laser pulse must occupy a small fraction of the rf period, with typical widths of 10 to 15 ps. The laser's repetition frequency must be tightly locked to a subharmonic of the rf frequency. Pulse-to-pulse timing jitter, with respect to the phase of the gun's rf, leads to energy changes and also affects the beam's transverse phase space. This jitter must be kept below 1 ps rms.

The goals also include 1% variations in the diameter of the laser spot on the photocathode, with a centroid location that varies by no more than 1% of the diameter. The micropulse duration should also fluctuate by no more than 1%. Techniques for achieving this stability will be discussed below.

Oscillator

A CW actively mode-locked oscillator is needed to obtain picosecond pulses synchronized with the rf. As with the DC gun's laser, titanium sapphire is a natural choice for this wavelength range. Commercial mode-locked Ti:sapphire oscillators are available with wide tunability (from 700 to 1060 nm), 1 to 2 W of output power, and, in some types, with pulse widths selectable from 0.1 to 60 ps. A 10-W CW argon-ion laser provides the pumping, and can be stabilized with feedback for both pump power and pointing. Compared to other oscillators (*e.g.,* Nd:YLF or Nd:YAG), it is difficult to lock Ti:sapphire oscillators to an external rf source, since they tend to self-mode-lock due to inherent nonlinearities, without any reference to an external drive. However, carefully timed feedback to adjust the cavity length can control both the mode-locking frequency and phase relative to the reference signal. The commercial products claim a timing jitter below 2-ps rms, but measurements of 1 ps have been made, approaching the best values obtained with Nd:YLF. However, these units support pulse trains at frequencies of about 80 to 100 MHz. A new design with an extremely short cavity would be needed to mode-lock at the 714 MHz required for the NLC.

It is possible to provide a higher pulse-train frequency while making use of a standard oscillator. Start with a linearly polarized laser operating in the commercial range, at 89.25 MHz, which is 1/8 of the desired frequency and corresponds to a pulse spacing of 11.2 ns. Split the beam in two with a half-wave plate and polarizer. Delay one beam by half the pulse spacing, 5.6 ns, which requires an additional path of 1.7 m. Recombine the beams with a second polarizer, then use a Pockels cell driven by rf at 44.6 MHz to restore all pulses in the train to the original polarization. By repeating this process two more times with delays of 2.8 and 1.4 ns, and with Pockels cells at 89.25 and 178.5 MHz, we can prepare a complete pulse train with the desired 1.4-ns spacing. At times, the NLC may be operated with micropulses spaced by 2.8 or 5.6 ns. The number of pulses would be reduced by the same factor, so that the macropulse duration remains fixed. These spacings are readily accomplished by rotating the waveplates at the third (and second, for the 5.6-ns case) splitting, so that no splitting takes place.

Efforts are also underway to develop a mode-locked Ti:sapphire oscillator with a pulse train repeating at a very high rate. Hartmann and colleagues [Hartmann 1995b] have built a 76.54-MHz Kerr-lens-mode-locked Ti:sapphire oscillator synchronized with an external rf source. They have presented preliminary results from a second oscillator with a repetition rate of 1.039 GHz; this frequency will next be raised to 1.225 GHz, and the eventual goal is 2.45 GHz.

Even the demonstrated operating frequency is fast enough for the NLC, and such a laser would avoid the elaborate split-and-delay procedure of the previous paragraph.

Diode lasers may offer a promising alternative that can provide even lower jitter. Their small size makes them well suited for high repetition rates. They can be actively mode-locked with minimal jitter, since the laser gain can be directly modulated by applying an rf signal to the diode, without the need for an acousto-optic device to modulate the cavity Q. Diode lasers combining active and passive mode-locking have been developed [Delfyett 1992] at frequencies from 302 to 960 MHz, at wavelengths near 830 nm, and with bandwidths of 4–10 nm. When followed by another diode acting as an amplifier, and after removing the chirp, Fourier-transform-limited pulses of 0.46 ps with a jitter of 0.4 ps were obtained. These pulses can be dispersively stretched to any desired width while maintaining the low jitter. Filtering in the compressor removes unwanted spontaneous emission. However, compared to the Ti:sapphire oscillator, the output power is low, typically 10 mW, and so more amplification would be needed. There is also considerably less tunability. This second drawback may not be too serious, since a diode laser could be prepared to operate at whatever wavelength proves best for the polarized rf gun. Wider micropulse spacings can be implemented by driving the diode laser directly at 357 or 178.5 MHz, so that the gain is high only at multiples of the fundamental cavity spacing. With further development, this source may be well suited to providing the high repetition rate and low jitter needed for NLC.

A4.4 Amplifier

With either oscillator, the amplifier is similar to that proposed for the laser system for the NLC's DC gun—a multiple-pass Ti:sapphire amplifier pumped by two *Infinity* Nd:YAG lasers from Coherent Inc., each providing 200 mJ at 532 nm in alternation at 90 Hz. However, the pulsed oscillator used with the DC gun produces 100 nJ in each micropulse, after allowing for the losses in the Pockels cells that shape its train of square, 1-ns-wide pulses. This is more energy than is produced by either type of CW mode-locked oscillator. Consequently, we need more amplifier gain, which can be obtained with two amplifier stages, with the pump light split between them. With the Ti:sapphire oscillator, the 3-pass bowtie configuration with a gain of 8 to 10 per pass, would be suitable for both stages. For the diode laser, where the energy extraction would be very low in the first pass, a 5-pass first stage and a 3-pass second stage would provide sufficient gain.

Fourier-Relay Optics

A technique known as Fourier relay optics can address the conflict between two important goals: better energy extraction by filling the rod more completely, while maintaining a clean transverse mode. Initially, the oscillator beam is trimmed in an aperture. The technique combines relay imaging, in which lenses relay an image of this aperture to each pass through an amplifier rod, and filtering of the beam's spatial Fourier transform. At each step, a lens of focal length f_1 is placed a distance f_1 after one of the image planes. The Fourier transform is formed at the focus, f_1 beyond the lens, where higher spatial harmonics are removed by a pinhole. A second lens with focal length f_2 then recollimates the beam (with expansion $\frac{f_2}{f_1}$) and forms the relay image at a distance $f_1 + f_2$ from the previous image plane. An image of the aperture is finally relayed to the photocathode, to define the area of photoemission.

Shaping the Beam in Space

We usually consider laser pulses that are Gaussian in time and space. However, it would be useful to investigate ways of making electron pulses that are more square, with steeper sides and flatter tops. Simulations of Brookhaven's rf electron gun have shown that electrons emitted during the temporal tails of a Gaussian laser pulse occupy a different

region of phase space than those emitted near the middle. With a flatter laser pulse, the electrons are more tightly clustered, with a lower emittance over the full pulse (although the instantaneous "slice" emittance at times within the pulse is similar). Calculations also show that flattops in both space and time are best for emittance compensation using solenoidal focusing after the gun.

To shape the pulse in space, a position-dependent attenuation could be applied to the beam [Van Wonterghem 1993]. Relay imaging should be used after the flattening, to limit diffraction. Stability of the laser's position and diameter on the photocathode can be achieved by trimming the edge of the beam with an aperture on the final relay plane; this aperture is then imaged onto the photocathode. While a Gaussian beam could still have fluctuations in the position of the centroid within the aperture, with a flattop beam, pointing jitter does not cause any change in cathode illumination.

An rf gun that directly produces a spatially flat electron beam is under consideration. The flat-beam gun would have a vertical beam size and emittance substantially smaller than the horizontal, as ultimately required at the interaction point. The required flat laser pulse can readily be produced with a small modification of the final relay stage: a pair of cylindrical lenses could provide different magnifications in the two transverse directions to produce an ellipse on the cathode.

Stretching, Compressing, and Shaping the Micropulses in Time

In amplifiers for picosecond and especially subpicosecond pulses, the peak power must be limited to avoid optical damage and nonlinearities. The Ti:sapphire oscillator can produce a 0.1-ps pulse with a large bandwidth. The dispersion of a grating pair can then stretch the pulse to hundreds of picoseconds, so that the peak power is reduced before amplification. The stretching results from having different wavelengths in the pulse take different optical paths, and so correlating time, space, and wavelength. After amplification, the process can be reversed to compress the pulse to the original or any greater width. This technique is known as chirped-pulse amplification [Maine 1988].

During the stretching, temporal pulse shaping is readily accomplished in the dispersive region between the gratings, by using a spatially varying filter to selectively attenuate the wavelengths corresponding to different times [Skupsky 1993]. (Filtering could also be done in the compressor, but the power levels are higher.) More sophisticated shaping, including the production of square picosecond pulses, has been achieved using both amplitude and phase masks to manipulate the pulse's Fourier transform [Weiner 1988]. The rise time and flatness of the resulting pulse are limited by the bandwidth of the input pulse, determined by the oscillator's pulse width. Thus the shaped pulse cannot rise faster than the oscillator.

For lower-resolution shaping, an interesting possibility is the use of a linear array of liquid crystals to form a voltage-controlled, spatially varying, amplitude mask, allowing active control of the pulse shape. Guided by measurements of the amplified pulse shape, a computer could provide a feed forward, adaptive control system.

Pulse-width stability is important for the NLC's performance. If the bandwidth of the oscillator pulse is wider than that transmitted by the phase and amplitude masks, so that the masks are illuminated almost uniformly, then fluctuations in the oscillator width do not affect the final pulse width or shape, which is determined only by the masks and the pulse compressor following the amplifiers.

Feedback and Feed-Forward Corrections

A high-extinction (>1000), fast-rise, variable-width, Pockels-cell gate after the oscillator chops the pulse train to the desired length. Over the course of the macropulse, a substantial fraction of the energy stored in the amplifiers by the pump laser is extracted. To compensate for the resulting drop in gain for the later micropulses, a second Pockels

cell after the oscillator or between the amplifiers modulates the macropulse, attenuating the early pulses to flatten the envelope of the pulse train. An arbitrary waveform generator drives the Pockels cell with a waveform that results in a flat pulse train after amplification. A computer controls the process in a feed-forward loop: the computer monitors the envelope with a photodiode and continually reprograms the arbitrary waveform generator for subsequent macropulses to maintain the desired envelope.

For modulation, it is convenient to use a cell that imposes a transverse field across a crystal such as Lithium Tantalate (LTA). Although the extinction is not as high as that of the more common, longitudinal-field, Dereterated Potassium Dihydrogen Phosphate (KD*P) cell used for the gate, the switching voltage is much lower—about 150 V rather than 8000 V for half-wave rotation. The difficulty of making a high-voltage, high-speed, linear amplifier is avoided; instead, a transverse-field cell is available from Conoptics matched to a 120-MHz solid-state linear amplifier. The high bandwidth of the driver, and its 1-V input, allow for sophisticated control of the macropulse envelope.

Feed-forward should compensate well for droop and slow drifts in the laser's performance. Random, shot-to-shot fluctuations will also arise, largely due to fluctuations in the flash lamps of the pump laser. Feedback within each macropulse is needed to correct for shot-to-shot changes in the envelope. An error signal for feedback could be made by comparing the measured envelope to a square pulse. This difference could then be subtracted from the arbitrary waveform generator output. However, cable delays and the amplifier's internal propagation delay limit the feedback bandwidth to about 3 MHz. In principle, the amplified pulses could be measured, then delayed with several meters of optical path while the voltage on a Pockels cell is adjusted. Fast and precise electronics would be required to achieve the desired accuracy. Such a correction has been tried briefly on the SLC's gun laser, but noise in the electronics limited the correction to about 2%.

Other approaches are more practical. We could monitor the pump energy and derive a correction added to the arbitrary waveform generator output over the macropulse. Since the *Infinity* pump laser has a pulse duration of 3.5 ns and titanium-sapphire has an upper-state lifetime of 3 μs, there is sufficient time between pumping and extracting to make this correction. Alternatively, we could widen the gate, then use the modulator Pockels cell to shape the pulses into two trains separated by about 1 μs. After amplification, the earlier one, occurring after the rods are pumped but before the gun's rf starts, would be measured to sense the system gain. The later pulse could then be corrected and sent to the gun.

2.A.5 Integrated System

Once the various physics issues and technical difficulties have been resolved for the polarized source photocathode rf guns, an integrated system including the laser and accelerator beam line up to 80 MeV needs to be constructed and demonstrated to operate in a stable, reliable, and efficient way, for an extended period of time. When the system integration tests show that a polarized photocathode rf gun can operate with a 99% availability, it will become the most attractive option as an injector the NLC.

2.A.6 Conclusions

The polarized source photocathode rf gun is a significant option for the NLC injector because of the potential of producing low emittance bunch trains from the injector region, thus simplifying the design of the damping ring or at least simplifying the operation of the damping ring. If the various physics issues concerning the beam dynamics and the technology issues of reliable operation of a polarized source rf gun are solved in the laboratory, it would become the most attractive option for the polarized electron source for the NLC.

2.B Charge Limit and its Implications on High-Polarization Long-Pulse Charge Production

2.B.1 Introduction

The charge limit (CL) phenomenon refers to the suppressed emission of excited electrons in the conduction band from a p-type negative electron affinity (NEA) or nearly NEA semiconductor cathode due to an increase in the surface work function caused by those electrons that fail to escape and eventually become trapped at the surface [Woods 1993]. This happens because among the largely thermalized conduction band electrons that reach the surface, only a fraction of them may successfully escape due to limited escape probability as determined by the surface NEA property. The electrons trapped at the surface are removed mainly by combining with holes that tunnel to the surface through the band-bending potential barrier at a rate critically dependent on the doping density [Tang 1994b]. Charge limit (CL), or suppressed emission, occurs only if the excitation laser intensity is sufficiently high such that the rate of electrons getting trapped at the surface considerably exceeds the rate of removal and, therefore, leads to an appreciable buildup of electrons at the surface within the laser pulse duration. CL is not a total charge limit, as was originally believed. In a long-pulse mode, CL will likely manifest itself as a current limit. CL depends on the extraction electric field, in this respect bearing a vague resemblance to the space-charge limit, as well as on the cathode's quantum efficiency (QE) [Tang 1993]. The significance of the CL effect with respect to high-current long-pulse charge production is well addressed in reference [Tang 1994c].

In the summer of 1993, a systematic experimental study on the CL phenomenon using a series of strained layer GaAs cathodes was conducted. The cathodes used in that study included 100-nm and 300-nm high- and medium-doping (2×10^{19} and $5 \times 10^{18}\,\mathrm{cm}^{-3}$, respectively) strained GaAs, etc.. It is fair to state that from this study most of the important properties of the CL phenomenon for 2-ns pulses were learned. These results may be used to project the CL behavior for long-pulse ($\sim$100-ns) operations before actual long-pulse CL data become available. Reference [Tang 1994c] furnishes a quantitative analysis of the prospects of producing the required charge from a polarized electron photocathode for the NLC using the 300-nm medium- and high-doping strained GaAs cathodes used in the 1993 study. The conclusion was that charge production for the NLC appears possible only with the high-doping cathode. The fundamental deficiency associated with the medium-doping cathode was its long relaxation time of the bunch-bunch effect, which leads to a substantial increase in the surface work function and, therefore, strongly suppressed emission in the high-current long-pulse operation required by the NLC.

2.B.2 Generalization of CL Effect to Long-Pulse Operation

In the following, a similar long-pulse CL analysis as presented in reference [Tang 1994c] shall be given for a 100-nm high-doping strained GaAs cathode, also used in the 1993 study. When excited at the band-gap threshold, that is, at 866 nm, this cathode yielded a maximum polarization for emitted electrons of about 65%. The relevant experimental 2-ns CL data from this cathode are shown in Figures 2-26 and 2-27. For the sake of facilitating our quantitative analysis, it will be assumed that a 866-nm laser pulse with a sufficiently long, flattop pulse length ($>$126 ns) is used to illuminate the cathode. Under the illumination of the long laser pulse, the cathode response will first undergo a transient period during which the photocurrent yield will vary (or more exactly, decrease) with time. The duration of the transient period will be determined by the relaxation time of the bunch-bunch effect and the instantaneous laser power within the pulse. A steady emission state will be realized following this initial transient period. It is this steady-state emission current that we will try to evaluate based on the available 2-ns charge pulse data.

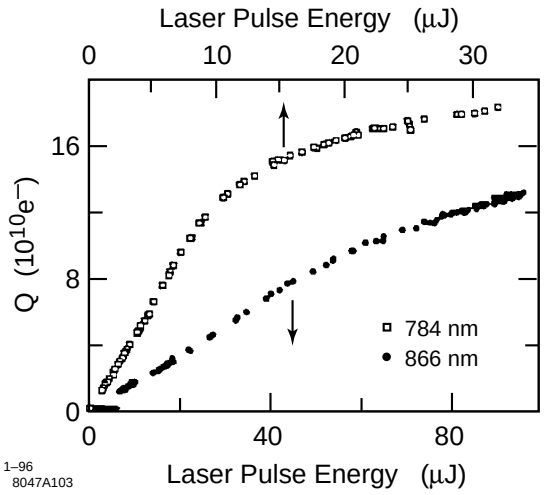


Figure 2-26. *Change vs. laser pulse energy for a 100-nm high-doping strained GaAs cathode with $QE = 22\%$ measure at 833 nm. The cathode is biased at 120 kV. Note that charge saturation was not reached at 866 nm due to insufficient laser energy.*

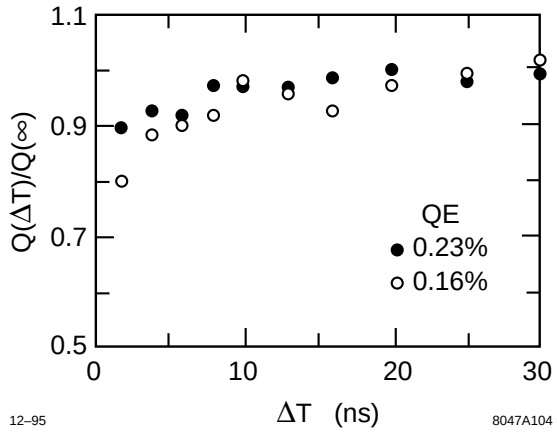


Figure 2-27. *Bunch-bunch effect for a 100-nm 2×10^{19} cm^{-3} doped strained GaAs sample for two different QE conditions (measured at 833 nm). Bunch 2 is at peak polarization with a laser pulse energy of 95 μJ at 866 nm. Laser pulse 1 is at 784 nm and has 32 μJ.*

Based on the earlier assumption that the beam from the gun is a long 126-ns square pulse, we need a steady-state emission of roughly 4 A from the cathode to meet the NLC Phase I requirement. We may view the long pulse as consisting of many back-to-back 2-ns pulses of 5×10^{10} electrons each. As shown in Figure 2-27, the bunch-bunch effect is almost completely diminished after 10 ns. Therefore, the emission of an arbitrary 2-ns pulse can be affected only by the five preceding 2-ns pulses. From Figure 2-26 we can see that a laser pulse of about 30 μJ at 866 nm, or about 5 μJ at 784 nm, is needed to produce a single stand-alone 2-ns pulse of 5×10^{10} electrons at $QE(833\,\text{nm}) = 0.22\%$ (slightly below the cathode's maximum QE of 0.23%). Neglecting the inter bunch effect, the total laser energy for five such pulses is 150 μJ at 866 nm or 25 μJ at 784 nm, which is less than the Bunch 1 laser energy used for taking the inter-bunch effect data shown in Figure 2-27 Here, the fact that Bunch 1 is at 786 nm hardly matters at all because the first bunch serves to pump a large number of electrons into the conduction band, thus saturating the cathode surface with electrons, and photoexcitation at 786 nm is about six times as efficient as at 866 nm for the cathode under study. Due to the accumulated electrons at the surface, the emission factor for the next pulse following these five pulses is reduced to about 0.9, corresponding to the first data point in Figure 2-27 for which Bunch 2 immediately follows Bunch 1 with a QE of 0.23%. This is a conservative estimate since it has been assumed that practically none of the electrons trapped at the surface during the five preceding pulses has been removed. Due to this interbunch CL effect, the actual laser energy required to produce a long pulse of intensity 5×10^{10} electrons per 2 ns needs to be about 10% higher. We can carry out the above sequence of analysis all over again with the upward-adjusted laser energy. Clearly, with a laser power of 15 kW(= 30 μJ/2 ns) at 866 nm this cathode appears capable of producing the required 4-A-long pulse, with considerable headroom by noting that increasing the laser power would lead to an increased current up to a certain limit. Of course, overdriving the cathode with too much laser power would result in a decreased emission current due to overly suppressed emission. Even when the QE drops to 0.16%, this cathode still appears promising for generating the required 4-A steady-emission charge pulse.

In the NLC polarized electron source design, the long laser pulse that illuminates the cathode is actually modulated at a duty factor of 50% with a period of 1.4 ns. With such a modulated pulse, the charge buildup at the cathode surface that causes the charge limit effect will be reduced by a factor of two compared with a square pulse as assumed for the above analysis. This would allow us to double the instantaneous laser power in the modulated pulse while keeping the emission factor at a similar value and, therefore, lead to a factor of two increase in the instantaneous emission current in the steady emission state. Based on the above analysis, the 100-nm high-doping strained GaAs cathode appears capable of generating an 8-A instantaneous current in a 50% modulated pulse, which corresponds to a microbunch charge of 3.5×10^{10} electrons. Such charge performance should even be adequate for the 1.5 GeV NLC.

It must be stressed that the favorable conclusions obtained from this analysis and from reference [Tang 1994c] do not mean that high-doping cathodes are actually capable of generating NLC-type charge pulses. These analysis should be viewed as merely a logical and physically reasonable generalization of the CL effect from a nanosecond scale to a time scale of an order (or two orders) of magnitude greater. It is possible that new unknown CL properties may exist on such a long time scale which may adversely affect emission. Therefore, the ultimate proof must come from experiments.

2.B.3 Where We Are

Even if high-doping strained GaAs cathodes are indeed adequate with respect to charge production, they still suffer from the deficiency of lower polarization than their medium-doping counterparts. The highest polarization from a high-doping strained cathode is about 65% (with a 100-nm-thick active layer), which is substantially lower than the 80% polarization specified for the NLC. While medium-doping 100-nm strained GaAs cathodes, such as the ones used for E-143 and SLC94-95, meet the NLC polarization requirement, their long CL relaxation times severely limit their long-pulse charge production capability. Recent long-pulse charge tests using such a 100-nm strained GaAs cathode showed that these cathodes are capable of sustaining an emission current of about 0.7 A with an 80% polarization in

an SLC gun. If, as proposed in the ZDR, a modulated laser pulse with a 1.4-ns period and a 50% duty factor within the pulse duration is used instead of the square laser pulse used in the test, the steady-state emission current within a microbunch should exceed 1 A, yielding a microbunch charge of $\geq 0.45 \times 10^{10}$ electrons. This leaves us about a factor of 4 (or 7) away from realizing an NLC-I (or NLC-II) polarized electron source. While the difference is still substantial, an NLC source is certainly within reach.

2.B.4 Cathode Improvements and Outlook

The key to realizing an NLC source is improved cathode performance. Strained GaAs cathodes with a thickness on the order of 100 nm will be our primary choice because of their proven superiority in polarization performance. The 100-nm medium-doping strained GaAs currently used at SLAC may be improved via the following approaches:

1. Modulated cathode doping scheme: High doping ($2 \times 10^{19}\,\mathrm{cm}^{-3}$) in a thin, typically 10-nm, surface layer and low doping ($\leq 1 \times 10^{18}\,\mathrm{cm}^{-3}$) in the rest of the active layer. The high doping at the surface helps minimize the CL relaxation time, thereby enhancing the long-pulse current, whereas the low doping in the interior maximizes the electron polarization. The presently employed heat-cleaning method for NEA activation may not be compatible with such a cathode as it may destroy its modulated doping profile. Alternative cathode cleaning techniques, notably atomic hydrogen assisted low-temperature cleaning, will be investigated.
2. Cathode surface protection: As (Arsenic) capping. By growing an As cap layer of about 1-μm thick on strained GaAs before exposing to atmosphere, the active emission surface may be kept atomically clean. Such a clean surface may substantially improve the quality of an NEA activation. As a result, the cathode's quantum efficiency may be improved by about a factor of two. Such an improvement is expected to greatly suppress the CL effect, thereby maximizing the charge performance. As-capping is particularly desirable on modulated-doping cathodes since surface cleaning may be adequately performed at 400°C, which would leave the modulated doping profile intact, instead of at 600°C necessary for an uncapped cathode. The reduced heat cleaning temperature should also significantly improve the operational reliability of the loadlock system.

Other R&D areas on cathodes include tailoring the composition of both the substrate and the strained active layer to optimize the polarization and charge performances, and developing other types of cathode structures such as unstrained or strained superlattice cathodes.

At present, the Gun Test Lab and the Cathode Test Lab at SLAC are adequately equipped for polarization and 2-ns pulse charge tests. As new cathode materials arrive, such tests may be performed and the results may be used to project their long-pulse behavior. For actual NLC-type long-pulse tests, a high-powered long-pulse laser, such as the one discussed in Section 2.3 or a Q-switched flashlamp-pumped Ti:sapphire, must be developed.

In conclusion, the CL effect is the fundamental limiting factor for developing an NLC polarized electron source. The present SLC source is already within an order of magnitude in terms of charge performance from the ultimate NLC-III source. With the numerous possibilities of cathode improvements, however, the prospects for realizing an NLC source are very good. We should all be encouraged by the superb charge performance demonstrated by an As-capped KEK superlattice cathode [Kurihara 1995].

References

[Aleksandrov 1995] A.V. Aleksandrov *et al., Phys. Rev. E* **51**, 1449 (1995).

[Alley 1994] R. Alley *et al.,* "The Stanford Accelerator Polarized Electron Source", SLAC–PUB–95–6489 (1994).

[Alley 1995] R. Alley *et al.,* "The Stanford Linear Accelerator Polarized Electron Source", NIM in Physics Research, A365, 1–27 (1995).

[Carlsten 1989] B.E. Carlsten, *Nucl. Instr. and Methods* **A285**, 313 (1989).

[Clendenin] J. Clendenin *et al.,* "Prospects for Generating Polarized Electron Beams for a Linear Collider using an RF Gun", SLAC–PUB–6172.

[Delfyett 1992] P.J. Delfyett *et al., IEEE J. Quantum Electron.* **28**, 2203 (1992)

[Gallardo 1993] J.C. Gallardo and H.Kirk, "An Injection Scheme for the Brookhaven ATF Utilizing Space-Charge Emittance Growth Compensation", *Proc. 1993 Part. Acc. Conf.,* 3615–3617 (1993).

[Hartmann 1995a] P. Hartmann *et al.,* "Generation of short electron bunches from GaAs photocathodes", presented at the Workshop on Experiences with Polarized Electron Sources using GaAs Technology, Bonn, Germany (1995).

[Hartmann 1995b] P. Hartmann *et al.,* "Experience with Polarized Electron Sources using GaAs Technology", poster presented at Univ. of Bonn, Germany (June 10, 1995).

[Herrera-Gomez 1993] A. Herrera-Gomez, and W.E. Spicer, "Physics of high-intensity nanosecond electron source", *SPIE Proceeding Series*, **2022**, 51 (1993).

[James 1981] M. B. James, R. H. Miller, "A High Current Injector for the Proposed SLAC Linear Collider", *IEEE Trans. Nucl. Sci.* NS-28, (3), 3461, (1981).

[Kim 1989] K.J. Kim, *Nucl. Instr. and Methods* **A275**, 201-218 (1989).

[Kirk 1995] H. Kirk *et al., Proc. 1995 Part. Acc. Conf.,* 265 (1995).

[Koechner 1992] Walter Koechner, *Solid State Laser Engineering*, Springer Series, NY, 1992.

[Kurihara 1995] Y. Kurihara *et al.,* "A high polarization and high quantum efficiency photocathode using a GaAs-AlGaAs superlattice", *Jpn. J. Appl. Phys.,* **32**, 1837 (1995).

[Maine 1988] P. Maine *et al., IEEE J. Quantum Electron.* **QE-24**, 398 (1988).

[Matsumoto 1994] H. Matsumoto *et al.,* "High Power test of a High Gradient S-band Accelerator Unit for the Accelerator Test Facility", *Proc. LINAC 94*, 302–304 (1994).

[Naito 1994] T. Naito *et al.,* "Multi-Bunch Beam with Thermionic Gun for ATF", *Proc. LINAC 94*, 375 (1994).

[Palmer priv] D.T. Palmer, private communication.

[Palmer 1995a] D.T. Palmer *et al.,* TESLA FEL Group Meeting, August 1995.

[Palmer 1995b] D.T. Palmer *et al.,* "Simulation of the BNL/SLAC/UCLA 1.6 Cell Emittance Compensated Photocathode RF Gun Low Energy Beam Line", *Proc. 1993 Part. Acc. Conf.* (1995).

[Rosenzweig 1993] J.B. Rosenzweig and E. Colby, "Design of a High Duty Cycle, Asymmetric Emittance RF Photocathode Injector for Linear Collider Applications", *Proc. 1993 Part. Acc. Conf.*, 3021–3023 (1993).

[Rosenzweig 1995] J.B. Rosenzweig and E. Colby, "Charge and Wavelength Scaling of RF Photo Injector Designs", Advanced Accelerator Concepts (AIP Conf. Proc. 335) 724 (AIP, NY, 1995).

[Sheffield 1992] R.L. Sheffield *et al., Nucl. Instr. and Methods* **A318**, 282–289 (1992).

[Sheffield 1993] R.L. Sheffield *et al.,* "Operation of the High Brightness LINAC for the Advanced Free-Electron Laser Initiative at Los Alamos", *Proc. 1993 Part. Acc. Confl,* 2970–2972 (1993).

[Schultz 1992] D. Schultz *et al.,* "The Polarized Electron Source Performance in 1992 for SLC-SLD", presented at the Tenth Int. Sym. on H.E. Spin Physics, Nagoya (Nov. 9–14, 1992).

[Serafini 1995] L. Serafini and J.B. Rosenzweig, "Envelope Analysis of Intense Relativistic Quasi-laminar Beams in High Gradient Linacs", Submitted to *Phys. Rev. E* (1995).

[Skupsky 1993] S. Skupsky *et al., J. Appl. Phys.* **73**, 2678 (1993).

[Sze 1981] S.M. Sze, *Physics of Semiconductor Devices*, 2nd edition (John Wiley & Sons, NY, 1981).

[Tang 1993] H. Tang *et al.,* "Study of non-linear photoemission effects in III-V semiconductors", *Proc. 1993 Part. Acc. Conf.*, Washington DC, 3036 (1993).

[Tang 1994a] H. Tang *et al., Proc. Fourth European Part. Acc. Conf.*, 46 (1994).

[Tang 1994b] H. Tang *et al.,* "Experimental studies of the charge limit phenomenon in NEA GaAs Photocathodes", *Proc. Fourth European Part. Acc. Conf.*, London, England, 46 (1994).

[Tang 1994c] H. Tang *et al.,* "Prospects for a polarized electron source for next generation linear colliders based on a SLC-type gun", Proc. of the 11th Intern. Symp. on High Energy Spin Phys. (Bloomington, IN, 1994, in print).

[TESLA FEL] TESLA FEL Collaboration, TESLA-FEL 95-03.

[Travier 1994] C. Travier, "Review of Electron Guns", *Proc. EPAC94*, 317–321 (1994).

[Van Wonterghem 1993] B.M. Van Wonterghem *et al.,* Proc. Conf. Laser Coherence Control, SPIE Vol. **1870**, 64 (21–22 Jan. 1993).

[Wanant 1994] R. Waynant, M. Ediger, *Electro-optics Handbook*, Optical and Electro-optical Engineering Series, 34 (McGraw-Hill, 1994).

[Weiner 1988] A.M. Weiner *et al., J. Opt. Soc. Am.* **B5**, 1563 (1988).

[Woods 1993] M. Woods *et al.,* "Observation of a charge limit for semiconductor photocathodes", *J. Appl. Phys.* **73**, 8531 (1993).

[Yoshioka 1994] M. Yoshioka *et al.,* "High Gradient Studies on UHV Room Temperature Cavities at S-Band for Linear Colliders", *Proc. LINAC 94*, 302–304 (1994).

Contributors

- Ray Alley
- David Burke
- Jym Clendenin
- David Farkas
- Allan Fisher
- Joseph Frisch
- Zhenghai Li
- Roger Miller
- Yuri Nosochkov
- Dennis Palmer
- Tor Raubenheimer
- Louis Rinolfi
- James Rosenzweig
- Huan Tang
- Kathy Thompson
- James Turner
- A. Dian Yeremian
- Dieter Walz
- Juwen Wang

3

Positron Source

Contents

3.1 Introduction

3.1.1 Overview

The baseline positron source for the NLC is a conventional source based on an electromagnetic shower created by high-energy electrons impinging on a thick, high-Z target. Its design draws heavily on that of the SLC positron source, which has demonstrated excellent reliability over many years of operation. The source and its associated 2-GeV linac, initially built for the 500-GeV center-of-mass machine (NLC-I), must be upgradeable by simply doubling the energy of the incident electrons to produce the charge per pulse that meets the specification of the 1-TeV center-of-mass machine (NLC-II). In particular, NLC-II requires a train of 90 bunches with up to 1.25×10^{10} positrons/bunch in each accelerator pulse, at a pulse repetition rate of 120 Hz. This is a charge per pulse of about 23 times the design intensity of the SLC positron source. Target tests performed during the SLC R&D indicated a small (about 40%) margin of safety between the design and the destruction of the target from pulse heating. It appears wise to keep the pulse energy deposition density in the target below the SLC design value. Thus, the positron beam pulse intensity may be increased only by increasing the size of the drive electron beam to allow for increased pulse energy deposition in the target and by increasing the admittance of the positron capture system.

Therefore, the following strategy has been adopted in designing the positron source for NLC-II: (1) double the radius of the incident electron beam which allows the pulse energy deposition in the target, and hence the pulse positron production, to be quadrupled; (2) accelerate the positrons at an L-band frequency of 1428 MHz in a structure with an aperture slightly more than twice larger than the SLC source while maintaining the magnetic field of the uniform-field solenoid at the same value as in the SLC source. Doubling the beam radius on the target roughly quadruples the 4-D emittance of the outgoing positrons. However, the use of an L-band accelerator not only increases the 4-D transverse admittance by a factor of 16 but also doubles the longitudinal phase space admittance, though the benefit of the latter is not fully realized due to increased debunching caused by the more divergent particles. As a result, we find that the positron yield per electron per GeV is more than quadrupled in comparison with the SLC source. For NLC-I, a 50% increase in the radius of the incident electron beam is adequate. This leads to an improvement in the positron yield per electron per GeV by almost 40% over that of NLC-II. The net result of these scalings is that with incident electrons at 3.11 GeV and 6.22 GeV for NLC-I and NLC-II, respectively, the NLC source produces a positron beam with more than twice the intensity required at the interaction point. Such a safety margin is necessary for a conservatively designed positron source, as large beam losses may occur in transport through the booster linac and through the two damping rings. Assuming losses of 20% each in the booster linac and in the two damping rings and an additional 5% loss in the main linac and the final focus system, the source still boasts a 50% margin in beam intensity.

With the proposed scaling, the density of the energy deposited in the target for both machines is slightly below the SLC design, whereas the average deposited power is about a factor of 2.5 greater for both NLC-I and NLC-II. The higher average power is not a major problem, but it does exacerbate the problem of radioactivity which also has significant bearings on source maintainability. It will necessitate a larger-diameter rotating or trolling target and larger cooling water flow.

3.1.2 System Description

The proposed NLC positron source shares in design substantial similarity to its SLC counterpart. The conventional technique of producing positrons from $e^{\pm}$ pair production by bombarding a conversion target with high-energy electrons will be used. Three subsystems comprise the NLC positron source: an accelerator for the drive electron

beam, a positron production and collection system, and a positron booster linac. Figures 3-1 and 3-2 show a schematic layout of the NLC positron source. A brief description of each subsystem follows.

A DC electron gun, coupled to two subharmonic bunchers and an S-band buncher, produces an electron beam of the desired NLC multibunch structure, *i.e.,* with an inter-bunch spacing at 1.4 ns and a total of 90 bunches/pulse. The electron beam is then accelerated in an S-band linac to 3.11 GeV for the 500-GeV machine or 6.22 GeV for the 1-TeV upgrade before impinging on a positron production target. The drive beam intensity at the target is 1.5×10^{10} electrons/bunch.

The positrons emerging from the production target as a result of the electromagnetic shower cascade are captured by a 240-MeV L-band capture accelerator embedded in a uniform solenoidal field following an adiabatic phase-space matching device consisting of a flux concentrator and a tapered-field solenoid. The longitudinally varying magnetic field from the adiabatic matching device transforms the transverse phase space of the positron beam along the axis in such a way that

$$Br^2 = \text{const}, \tag{3.1}$$

and

$$B/p_\perp^2 = \text{const}, \tag{3.2}$$

where B is the magnetic field, r is the radial displacement, and $p_\perp$ is the transverse momentum. The choice for an L-band capture accelerator aims to boost mainly the 4-D transverse phase space acceptance by about a factor of 16 and the longitudinal phase space acceptance as well over a similarly designed S-band capture accelerator. The design employs a parallel dual source configuration for improved reliability. Electrons are dumped after the first bending magnet. Beam scrapers installed in between the bending doublet is used to to collimate the positron beam in energy as well as in transverse phase space prior to its injection into the booster linac.

The collimated positron beam, with the same bunch structure as the electron drive beam, is accelerated in an L-band booster linac to 2 GeV for emittance damping in a pre-damping ring followed by a main damping ring. The optics for the positron booster linac is a FODO array with its beta function scaled as E, which must suppress multibunch beam blow-up due to wakefields and limit chromaticity induced emittance growth. A large number of large-aperture quadrupole magnets surrounding the L-band accelerating sections and smaller-aperture quadrupole magnets elsewhere are used to form the lattice.

3.1.3 Parameters

The important parameters of the NLC positron source for both the 500-GeV and the 1-TeV machines are summarized in Table 3-1. The positron beam intensities are obtained from numerical simulation after optimizing various design parameters. As a reference, the operating and design parameters of the SLC positron source are also included.

3.2 Positron Yield Simulation

Positron yield simulations have been performed for both drive beam energies, *i.e.,* 3.11 and 6.22 GeV for NLC-I and NLC-II, respectively. The results are already presented in Table 3-1. In this section we shall discuss in detail the simulation results using a 6.22-GeV drive beam.

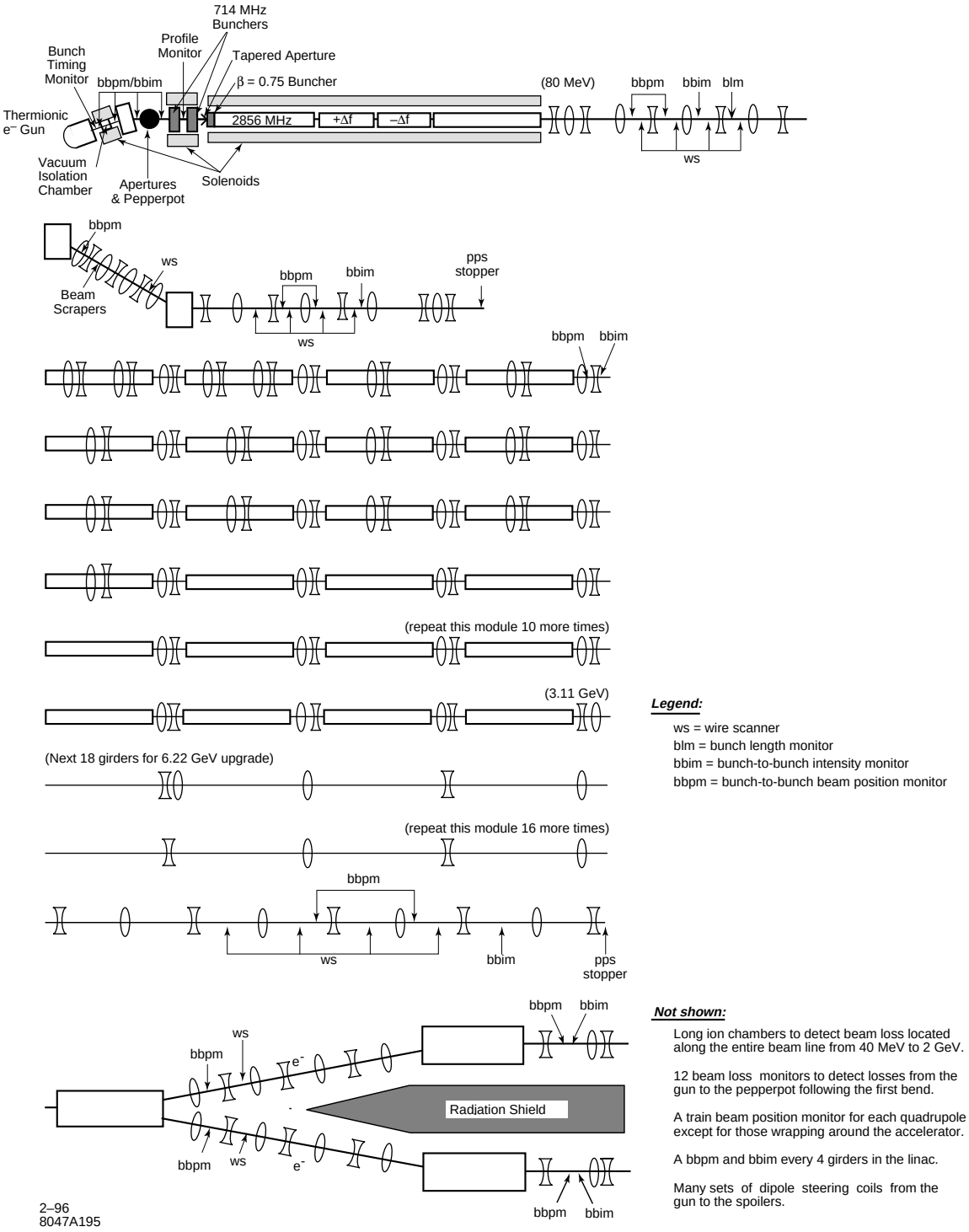
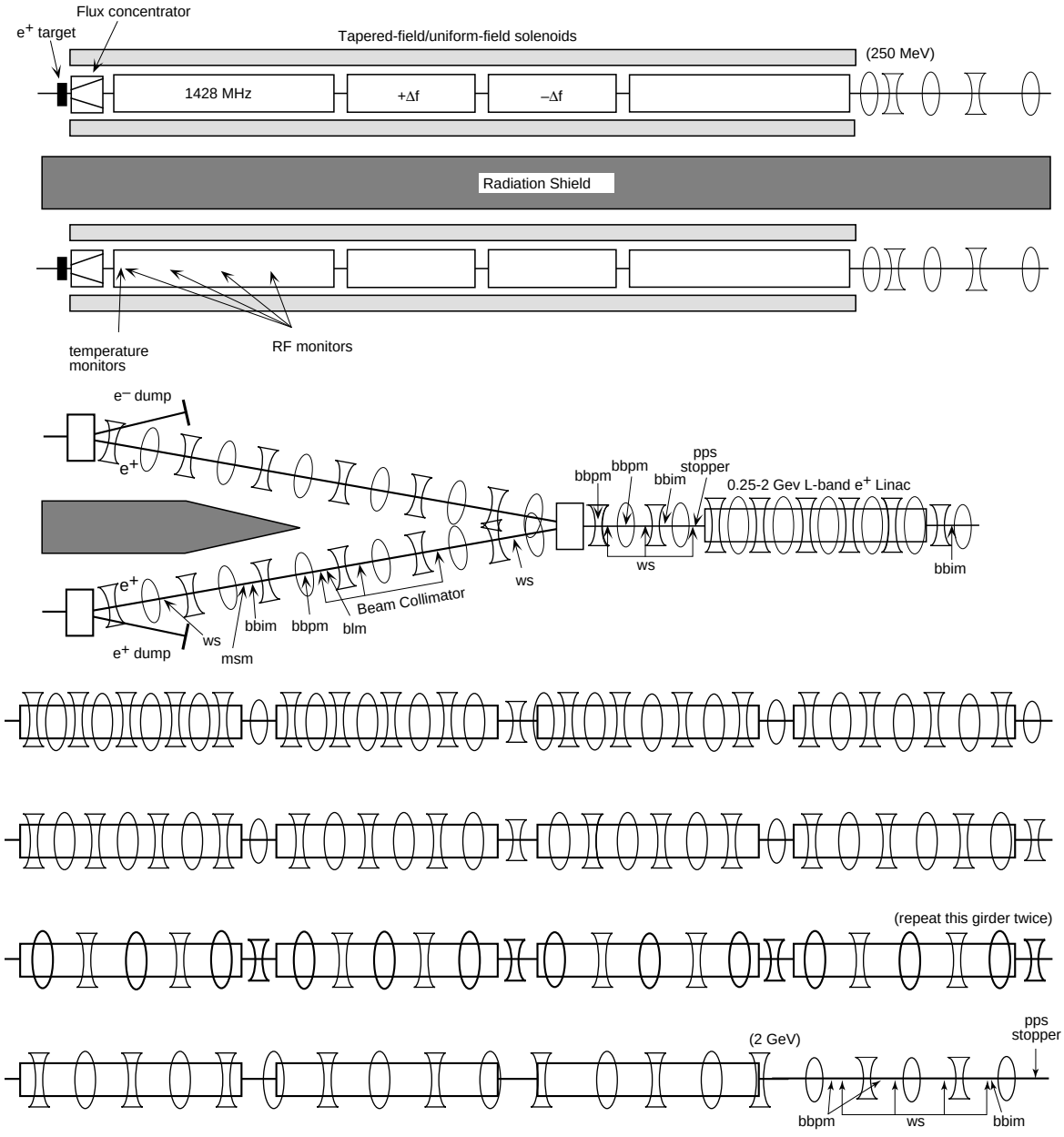


Figure 3-1. *Schematic layout of the NLC positron source (Part A): drive electron accelerator.*



Legend:

ws = wire scanner
msm = multistrip monitor
blm = bunch length monitor
bbim = bunch-to-bunch intensity monitor
bbpm = bunch-to-bunch beam position monitor

Not shown:

Long ion chambers to detect beam loss located along the entire beam line from 250 MeV on up in energy.

Beam position monitors at each quadrupole except those wrapping around the accelerator.

Dipole steering coils strategically placed along the L-band linac from 250 MeV on.

4–96
8047A196

Figure 3-2. *Schematic layout of the NLC positron source (part B): dual positron production/capture systems and 1.75-GeV booster linac.*

Parameter	Unit	SLC 94	SLC max design	NLC-I (500 GeV)	NLC-II (1 TeV)
General Parameters:					
Ne^+ per pulse at IP	[10^{10}]	3.5	5	76.5	112.5
# of bunches per pulse		1	1	90	90
Pulse duration	(ns)	0.003	0.003	126	126
Bunch spacing	(ns)	-	-	1.4	1.4
Repetition frequency	(hz)	120	180	180	120
Drive Electron Beam:					
Energy	(GeV)	30	30	3.11	6.22
Ne^- per bunch	(10^{10})	3.5	5	1.5	1.5
Ne^- per pulse	(10^{10})	3.5	5	135	135
Beam power	(kW)	20.2	47	121	161
RMS beam radius	(mm)	0.8	0.6	1.2	1.6
Beam energy density per pulse	(GeV/mm^2)	5.2×10^{11}	13.3×10^{11}	9.3×10^{11}	10.4×10^{11}
Positron Production Target:					
Material		$W_{75}R_{25}$	$W_{75}R_{25}$	$W_{75}R_{25}$	$W_{75}R_{25}$
Thickness	(R.L.)	6	6	4	4
Energy deposition per pulse	(J)	37	53	126	188
Power deposition	(kW)	4.4	9	23	23
Steady-state temperature	(°C)	100	200	400	400
Positron Collection:					
Accel. rf	(MHz)	2856	2856	1428	1428
Accel. gradient	(MV/m)	30	30	25	25
Minimum iris radius	(mm)	9	9	20	20
Edge emittance (allowing for 2 mm clearance)	(m-rad)	0.01	0.01	0.06	0.06
Positron yield per e^-		2.5[a]	2.5[a]	1.4	2.05
Ne^+ per bunch	(10^{10})	8.7	12.5	2.1	3.1
Ne^+ per pulse	(10^{10})	8.7	12.5	189	279

[a] The actual yield immediately following the capture section is 4, but decreases to 2.5 after two 180° bends and a 2-km transport line.

Table 3-1. *NLC Positron Source Parameters.*

3.2.1 Target Yield

The target yield for positrons has been calculated using the EGS program [Nelson 1985] for 6.22-GeV electrons impacting a $W_{75}Re_{25}$ target of thicknesses ranging from 4 to 6 radiation lengths. The optimal positron yield is obtained for a target thickness of about 5 radiation lengths. However, at this thickness about 23% of the beam energy is deposited in the target, as opposed to the 18% energy deposition in a 6-radiation-length-thick target in the SLC positron target. The increased volume density of energy deposition in the target is not desirable as it causes increased thermal stress which may lead to target failure. Therefore, a better choice would be to use a target of 4-radiation-lengths thick, for which the energy deposition is reduced significantly to about 14% while still retaining a respectable positron conversion yield. The total positron yield, defined as the number of positrons generated per incident drive electron, from such a target is about 12.5, whereas the total electron yield is about 15.8.

3.2.2 Ray Tracing through Capture Accelerator

The positron rays obtained from the EGS simulation have been traced through the flux concentrator and the capture accelerator up to the nominal 250-MeV point at the end of the capture region using the ETRANS program [Lynch 1989]. ETRANS is a ray tracing program developed at SLAC that integrates particle trajectories through static magnetic and rf fields while ignoring the effect of space charge and wakefields. The capture accelerator consists of two 5-m L-band (1428 MHz) accelerating sections with a loaded gradient of 25 MV/m and two off-frequency (1428 ± 1.4 MHz) 3-m sections for beam-loading compensation sandwiched in between the two accelerating sections. All four sections are embedded in a 0.5-T uniform solenoidal field. The rf phases of the accelerating fields in the capture region were varied to optimize the positron yield at the 250-MeV point. The best yield at the 250-MeV point, after applying energy and time cuts of ± 10 MeV and ± 30 ps, respectively, is found to be 2.05. With 1.5×10^{10} electrons/bunch in the drive beam, therefore, there will be approximately 3.1×10^{10} positrons/bunch in the L-band linac, which is a factor of 2.5 higher than the bunch intensity required at the interaction point. Figures 3-3 and 3-4 show the transverse and longitudinal phase space plots of the positron beam at the nominal 250-MeV point after the six-dimensional phase space cuts.

The electron rays emerging from the target were also traced through the capture accelerator. The electrons are approximately 65% more abundant than the positrons at the end of the capture system and have a most probable energy of about 230 MeV. These results are useful to evaluate the total beam loading in the capture accelerator and the power of the electron beam to be dumped after the first bending magnet.

3.3 Drive Electron Accelerator

3.3.1 Drive Electron Source

The baseline choice for the drive electron source is a DC thermionic cathode gun to be operated at 120 kV. A DC photocathode gun employing a bulk GaAs cathode and driven by a Ti:sapphire laser is considered as an option. The electron gun is required to deliver a very stable beam, with a pulse-to-pulse intensity jitter below 0.5% and a bunch-to-bunch intensity jitter below 2%. It is also required to have the capability to produce electron beams with variable bunch train structure to facilitate normal NLC operation as well as initial commissioning.

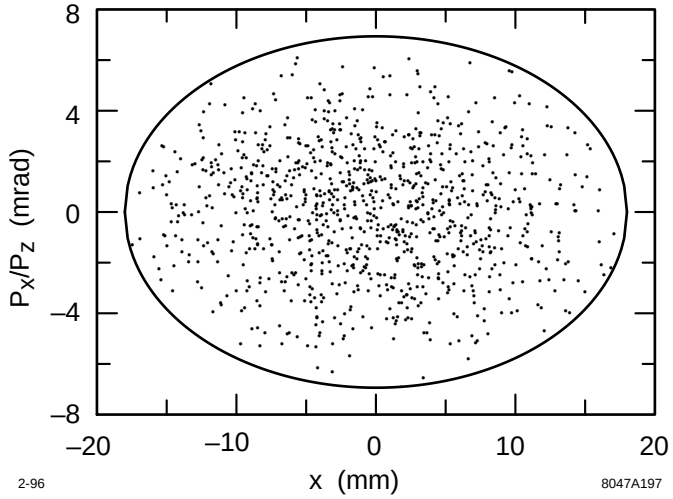


Figure 3-3. *Transverse phase space plot of the positron beam at the the exit of the capture accelerator.*

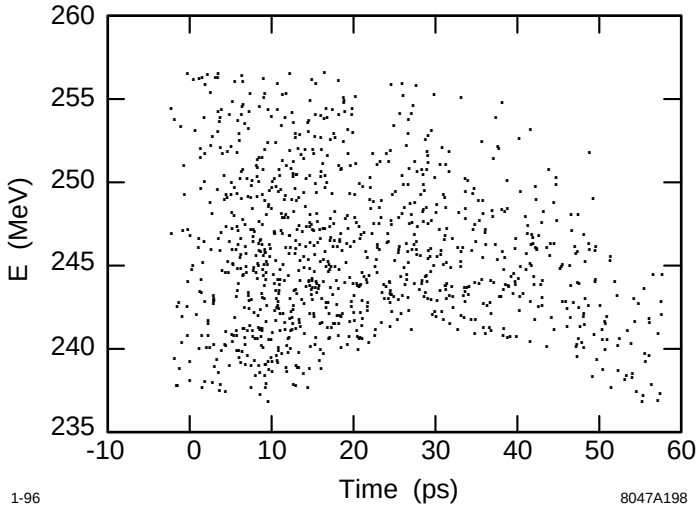


Figure 3-4. *Longitudinal phase space plot of the positron beam at the exit of the capture accelerator.*

The injector design for the drive electron beam is identical to that of the NLC electron source (see Chapter 2). To summarize briefly, it consists of two 714-MHz standing wave subharmonic bunchers and a 2856-MHz traveling wave ($\beta = 0.75$) S-band buncher, followed immediately by an 80-MeV S-band accelerator.

3.3.2 Drive Linac

The drive linac is also identical to the electron linac used in the NLC electron source except that it will be longer to accelerate the electrons to the higher energies required for efficient positron production. For NLC-I, the linac will end at 3.11 GeV, followed by a long drift section which will later be replaced by more accelerating sections to boost the energy to 6.22 GeV for the 1-TeV upgrade. Beam loading compensation in the linac will be accomplished by using the Δt approach. See Chapters 6 and 2 for details on the rf and lattice design issues.

3.4 Positron Production Target

The design of the NLC positron production target is modified significantly from that of the SLC target to accommodate the significantly higher beam power delivered to the NLC target. In addition, the target motion mechanism is modified from that of trolling as used in the SLC source to that of rotating to eliminate intensity modulations in the positron beam as experienced in the SLC. The periodic trolling motion of the target and its drive mechanism in the intense magnetic field of the tapered-field solenoid induces Eddy current in the moving parts. The Eddy current in turn generates a periodically varying magnetic field, leading to a small intensity modulation in the captured positron beam with the same frequency as that of the trolling motion. With a rotating target, however, the induced Eddy current should remain roughly constant, thus eliminating the source of intensity modulation.

3.4.1 Review of Target Test Data

Thermal stress from short pulse (or single pulse) heating is the underlying mechanism that causes material failure when a target is bombarded by a high-intensity, high-energy electron beam. As in the SLC positron source, $\mathrm{W_{75}Re_{25}}$ will be used as the target material for the NLC source because of its high-Z characteristic and its excellent thermal and mechanical properties. Material failure tests were conducted at SLAC in the early 1980s for W-Re targets with the Re concentration ranging from 18.6% to 27.6% [Ecklund 1981]. For these tests, the energy of the drive beam was in the range of 20–25 GeV, and the target thickness varied from 5–7 radiation lengths. The material failure threshold in terms of the maximum allowable beam energy density on the target per beam pulse was found to be approximately

$$\rho_{\max} = \frac{N\_E\_}{\pi\sigma^2} = 2 \times 10^{12}\,\mathrm{GeV/\,mm^2}\ , \tag{3.3}$$

with about 20% of the beam energy deposited in the target. In the above expression, $N\_$ is the number of electrons per pulse, $E\_$ the electron beam energy, and σ the rms radius of the electron beam on the target.

In the above, the failure threshold is expressed in an area beam energy density per pulse. While this is a convenient quantity to characterize the drive beam, it must be emphasized that the most critical parameter for causing material failure is the volume density of energy deposition per pulse in the target. However, the volume energy deposition density depends strongly on the longitudinal position across the target thickness, and it is the maximum density that matters the most for target failure considerations. In practice, it is also useful to evaluate the average energy deposition

density in the target. In terms of an average volume density, then, the failure threshold for $W_{75}Re_{25}$ is approximately 7×10^{10} GeV/ mm^2 per radiation length per pulse.

3.4.2 Electron Beam Size

For NLC-II, the drive beam has an energy of 6.22 GeV with a pulse intensity of $1.35 \times 10^{12} (= 90 \times 1.5 \times 10^{10})$ electrons. In order to avoid damaging the target, *i.e.,* to keep the energy density per beam pulse on the target comfortably below the threshold of 2×10^{12} GeV/ mm^2, the rms beam size of the incident electrons must be increased to 1.6 mm, twice the SLC drive beam size or four times the area. This yields a pulse beam energy density of 1.04×10^{12} GeV/ mm^2 on the target. For NLC-I, the beam radius may be reduced to 1.2 mm, yielding a pulse beam energy density on the target of about 0.93×10^{12} GeV/ mm^2. Since the electron beam size at the end of the 3.11-GeV or 6.22-GeV drive linac is typically much smaller than 1 mm (the normalized transverse emittance is less than 1×10^{-4} m·rad), a thin low-Z scatterer serving as an emittance spoiler must be used a few meters upstream of the target to blow up the beam size. For NLC-I and NLC-II, with 19% and 14% of the beam energy deposited in the 4-radiation-length thick target, the average volume density of energy deposition in the target is about 40% and 50% below the failure threshold, respectively.

3.4.3 Beam Power and Target Size

The average beam power deposited in the target amounts to approximately 23 kW for both NLC-I and NLC-II (Table 3-1). This power is about a factor of 2.5 and 5 larger than the SLC design and operating parameters, respectively. As the SLC positron production system has demonstrated good reliability, it is desirable to maintain a similar target size to beam power ratio for the NLC. Thus, by simply scaling the SLC target up a factor of 4, we find that the NLC target needs to have a diameter of about 20 cm. The target rotation frequency needs to be sufficiently high so that the areas of beam pulse impacts for two successive pulses do not overlap. Taking into account the increased beam size and the increased repetition rate (for NLC-I), the rotation frequency is chosen to be around 2 Hz.

3.4.4 Target Engineering Issues

By adopting a rotating target design, realizing a $\leq 1 \times 10^{-7}$-Torr vacuum in the target chamber in an environment of very high radiation levels becomes a challenging issue due to the lack of suitable materials for use as a leak-free seal around the rotating target shaft. Instead of pursuing a perfect vacuum seal, the present design seeks to achieve the vacuum goal by using a combination of conductance limiting seals and several stages of differential vacuum pumping.

Target Motion Mechanism, Monitoring and Control

The target wheel will be connected by way of a drive shaft and driven by an inline electric motor (Figure 3-5). The shaft will have two sets of outboard bearings such that the shaft is cantilevered into the target vacuum chamber. Similar to the SLC system, a stepping motor will be used to drive the target shaft, and an angular position sensor will be used for shaft angle and speed monitoring.

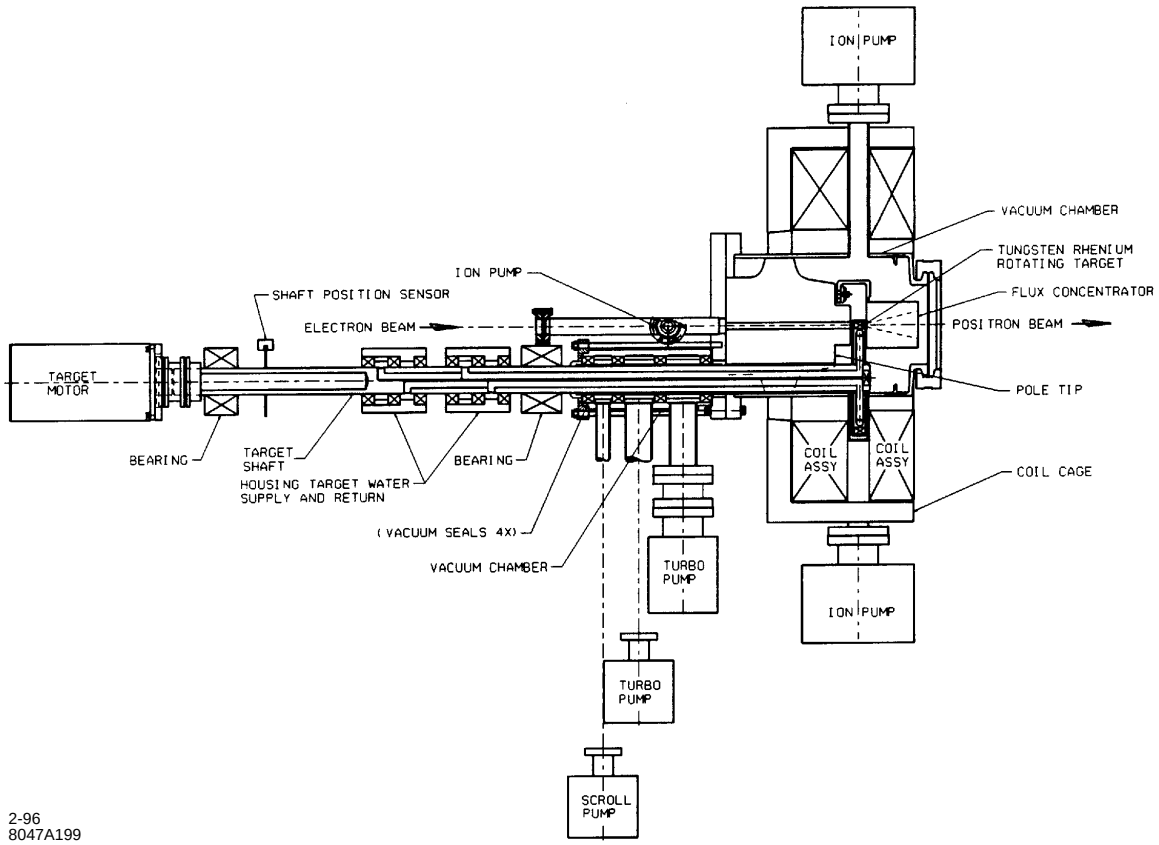


Figure 3-5. *Schematic of the NLC positron target system.*

Vacuum

The drive shaft will pass into the target vacuum chamber using a combination of radiation resistant seals and several stages of differential vacuum pumping along the length of the drive shaft (Figure 3-5). These seals involve a carbon and/or silver impregnated carbon element which rubs against a hard facing material such as tungsten carbide. These seals generally have tight clearances ($< 15\ \mu$m). Seal designs would be selected with leak path lengths of 25 mm or more. The leak rate goal for the first stage seal is on the order of 1 Torr-ℓ/s or less. Seals reviewed in the study included axial face seals, radial face seals, radial labyrinth seals, axial labyrinth seals, and magnetic face seals.

At least three stages and possibly four stages of differential vacuum pumping will be used along the drive shaft to isolate the rotating wheel target in the 1×10^{-7}-Torr vacuum from atmosphere. The first stage and possibly the second stage could use a dry scroll pump such as the Varian 600DS. This pump is completely oil free and has a base pressure in the 1×10^{-3}-Torr range. Differential pumping stages #2 and #3 or #3 and #4 (depending on the total number of stages used in the pumping system) could each use a turbomolecular pump such as the Varian Turbo-V70D. The Varian Turbo-V70D is a completely sealed, maintenance-free unit and is available in radiation-hardened configurations. The base pressure on these turbo pumps is 8×10^{-10} Torr. Each turbo pump will be backed with an oil free dry scroll pump. The vacuum objectives for the stages of differential pumping system in a three-stage system are: stage #1,

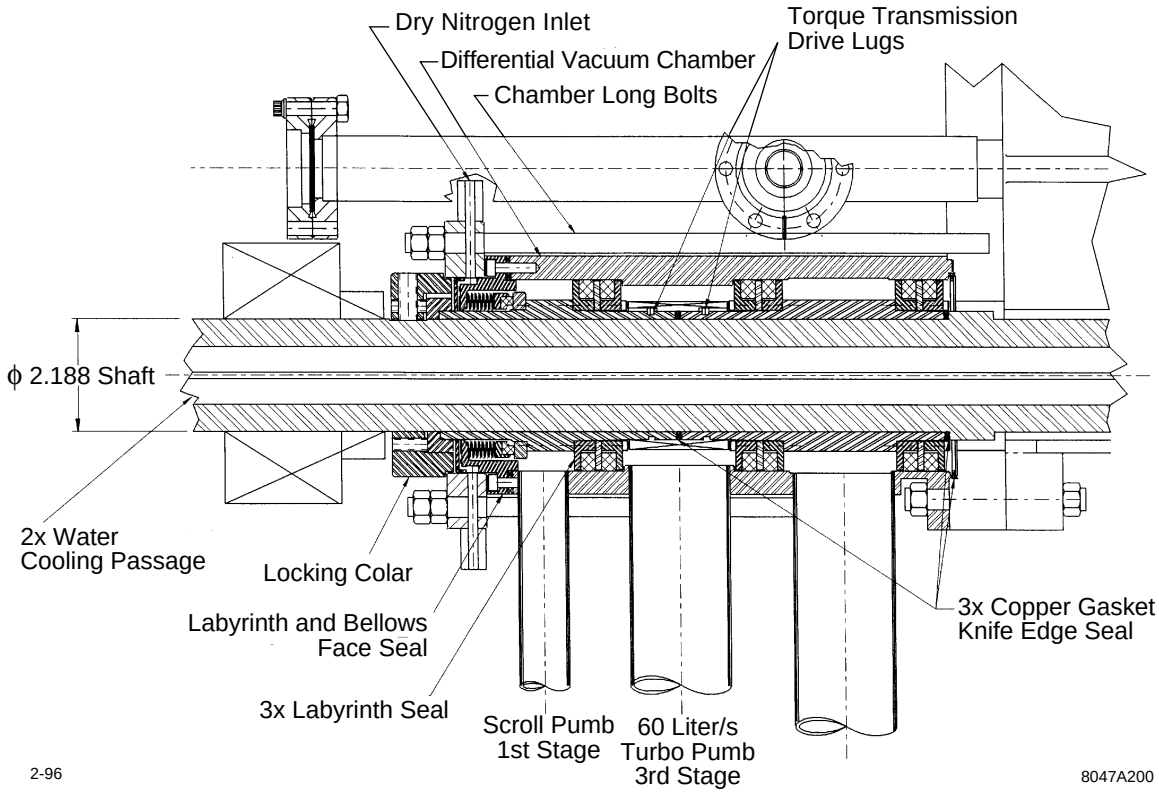


Figure 3-6. *NLC rotating target differential vacuum vacuum chamber and seals consept.*

1–3$\times 10^{-1}$ Torr; stage #2, 1–3$\times 10^{-4}$ Torr; stage #3, 1–3$\times 10^{-7}$ Torr. Vacuum in the target chamber will be about 1×10^{-7} Torr and provided by 30 ℓ/s or 60 ℓ/s ion pumps.

Figure 3-6 is a concept drawing showing a three-stage differential vacuum pumping system and support structure around the target drive shaft. This concept utilizes sets of radial labyrinth seals between vacuum pumping stages #1 and #2, stages #2 and #3, and again after stage #3. A bellows face seal with carbon contacting ring is used as the seal between atmosphere and stage #1.

A model two-stage differential pumping system with a rotating shaft will be built and experimented to prove the principle of this design and also to select the best seal materials.

Target Cooling

Figure 3-7 depicts a schematic cross-section view of the target wheel and drive shaft. Water supply and return lines will enter the shaft outside the vacuum space and pass through the center of the shaft to the target wheel assembly. The anticipated heat load on the wheel is about 23 kW for both NLC-I and NLC-II. The $W_{75}Re_{25}$ target element at the periphery of the wheel will be approximately 1.4-cm thick along the axial direction, and 1-cm thick in the radial direction. Cooling tubes will be located in a silver or copper casting adjacent to the $W_{75}Re_{25}$ target material to maximize the rates of heat conduction from the target to the cooling tube walls.

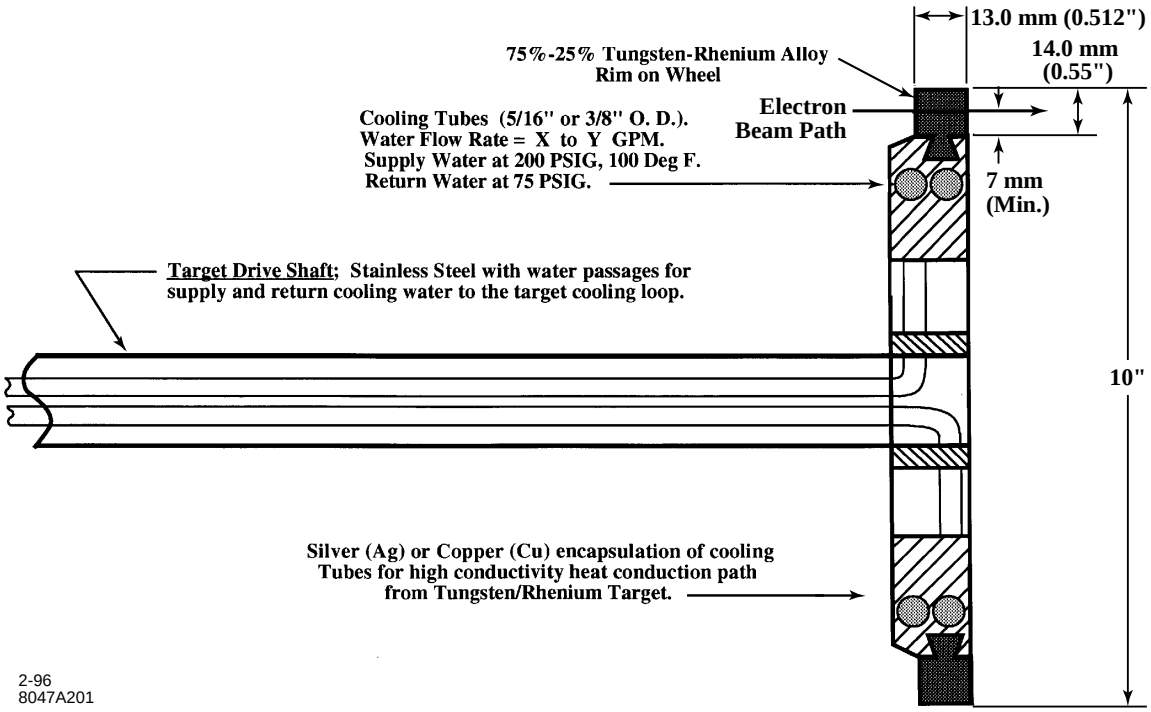


Figure 3-7. *NLC positron source target wheel schematic section.*

A high water velocity will be maintained in the cooling tubes (9–10 m/s) in order to yield a water side heat transfer film coefficient in the range of 17–20 kW/ m^2-°C. In order to produce the high cooling water velocity in the tubes, a water supply pressure of 200 psi (gauge) and return pressure of 75 psi (gauge) is assumed. Cooling water flow rate will be in a range of 80–120 ℓ/min by using cooling tubes with a 3/8-in or 5/16-in outer diameter.

For both NLC-I and NLC-II, temperature drops across the various components from the center of the beam impact point on the target to the cooling water are estimated to be as follows: ∼40°C across the water film inside the cooling tubes; ∼30°C across the stainless steel tubing wall; ∼50°C through the silver or copper casting around the cooling tubes from the $W_{75}Re_{25}$ target interface; and ∼220°C from the center of the beam impact point through the $W_{75}Re_{25}$ to the silver or copper interface. Thus, during operation, the maximum steady-state temperature in the target will be about 400°C which is quite comfortable for $W_{75}Re_{25}$. The steady-state temperature may be lowered substantially by modifications to the target design, such as reducing the radial thickness of the target to improve heat conductance.

The water supply and return lines to the target drive shaft will use radiation-hardened seals probably made from carbon and/or silver-filled carbon or metal composite.

3.4.5 Backup Power

The lack of a true vacuum seal in the positron target chamber means that, in the event of a power outage, the target chamber along with the capture accelerator will be vented due to lost pumping in the various stages. This will result in extended machine downtime as processing the accelerating sections following an air vent may take a long time.

Therefore, it is essential to have a reliable emergency backup power system. The backup system may include an uninterruptible power source as well as a power generator, both commercially available. In addition, the atmosphere side of the first stage seal will be flooded with dry N_2 gas. This is designed to provide further insurance that even in the event of a vent through the multi-stage seals the target chamber and the capture accelerator will be filled with dry N_2, which is relatively easy to pump out and processing the accelerator after venting to dry N_2 should be relatively quick. In addition, the use of a N_2 buffer gas also minimizes the amount of undesirable air gas species, such as H_2O, in the various differential pumping stages and the target chamber during normal operation.

3.4.6 Integration with Positron Collection System

The positron target system including its vacuum chamber must be closely mated to the flux concentrator and the subsequent capture accelerator to allow for efficient capture of positrons. As shown in Figure 3-5, the flux concentrator shares the same vacuum chamber as the target and is positioned close to it. The spacing between the flux concentrator and the first capture accelerating section must be kept at a minimum. Again, the SLC positron source serves as a design basis for these components and their integration.

3.5 Positron Collection System

The positrons emerging from the target have small spatial and temporal but large angular and energy distributions. Therefore, the use of a large-bandwidth phase-space transformer is essential to yield good capture efficiency of the positron beam into the capture accelerator which is embedded in a long solenoid magnet. As in the SLC positron source, a pulsed flux concentrator and a DC tapered-field solenoid [SLC 1984] will be used immediately following the conversion target to provide the adiabatically-varying longitudinal magnetic field which is essential for realizing the phase-space transformation [Helm 1962].

3.5.1 Flux Concentrator

The design and fabrication of the flux concentrator will follow exactly the SLC version [Kulikov 1991]. It is designed to produce a 5.8-T peak field along its axis. The minimum radius of the internal cone of the flux concentrator needs to be increased from 3.5 mm in the SLC version to 4.5 mm for the NLC to accommodate the increased radial extent of the emerging positron beam as a result of the increased incident electron beam size. The flux concentrator will be machined from a single block of Cu. The details of the fabrication process are described in the paper by Kulikov *et al.*, [Kulikov 1991]. The main advantage of using a flux concentrator is that it boosts the positron capture efficiency by a factor of 2–3 compared to a capture system utilizing a 1.2-T tapered field solenoid only.

3.5.2 High Gradient L-Band Capture Accelerator

For high capture efficiency, the positrons entering the capture accelerator must be accelerated to relativistic energies as quickly as possible to minimize bunch lengthening. Thus, the accelerating gradient needs to be as high as practically possible. The present design calls for the maximum unloaded gradient in the L-band capture accelerator, consisting of two 5-m accelerating sections and two 3-m beam loading compensation sections, to be 28 mV/m. To achieve

the desired gradient and also to facilitate beam loading compensation, each section will be powered by two 75-MW klystrons with SLED-I pulse compression. As the total beam loading current in the bunch train, including contributions from electrons as well as positrons both within and outside of the six-dimensional phase space acceptance window, can be up to 14 A for NLC-II, satisfactory beam loading compensation is realized only through the combined use of both Δt and Δf schemes. The positron beam energy at the end of the capture accelerator will be about 250 MeV, with a full energy spread of about $\pm 6\%$. The minimum iris radius of the disks will be 20 mm, and the outer radius of the disc-loaded waveguide will be about 11 cm. Chapter 6 contains more details on the rf design of the capture accelerator.

3.5.3 Tapered-Field and Uniform-Field Solenoids

A DC tapered-field solenoid producing a peak field of 1.2 T will be used in combination with the pulsed flux concentrator to serve as the phase space transformer, in which the magnetic field varies adiabatically from 7 T to 0.5 T. A 0.5-T uniform-field solenoid that encloses all four L-band accelerating sections will be used to provide transverse focusing in the capture accelerator. The design for the two solenoids is straightforward, as both have identical field specifications as those in the SLC positron source.

3.5.4 Space Charge

In this design, both the positrons and the electrons produced from the target will be collected and accelerated to about 250 and 230 MeV, respectively. The electrons will then be separated from the positrons after a bending magnet and dumped. In the region between the target and the accelerator, the electron bunch and the positron bunch overlap in space. Thus, wakefield and space charge forces will likely be insignificant. Once entering the capture accelerator, however, the positrons and the electrons will be quickly separated longitudinally, and space charge and wakefield forces may become important. However, experience from operating the SLC positron source suggests that the ETRANS simulation without taking into account space charge and wakefield forces is adequate.

3.6 Beam Dynamics and Transport

3.6.1 Control of Multibunch Beam Blow-up in Positron Linac

The normalized emittance of the bunches coming into the L-band linac just upstream of the positron pre-damping ring is $\gamma\varepsilon = 0.06\,\mathrm{m{\cdot}rad}$. The initial energy is 250 MeV and the final energy is that of the damping ring, 2 GeV. Other parameters of this linac assumed for the present simulations are shown in Table 3-2.

Calculations were performed using the program LINACBBU [Thompson 1990] which assumes smooth focusing scaling as an arbitrary power of energy. The effect of the focusing lattice was approximated by fitting an average beta function scaling as a power of energy. We obtained a beta function of 1.93 m at the beginning of the linac, and scaled it along the linac as $\beta = (E/E_i)\beta_i$. The resulting beam radius at the beginning of the linac is 1.8 cm. It will probably be desirable to strengthen the focusing at the beginning even more, to keep the beam size significantly smaller than the iris size.

Parameters	Values
Accelerating frequency	1428 MHz
Linac length	140 m
Initial energy	250 MeV
Final energy	2 GeV
Ne^+ per bunch	2.5×10^{10}
Bunch spacing	1.4 ns (2 rf buckets)
Initial beta function β_i (scales as E)	1.93 m
# bunches per train	90
Normalized emittance $\gamma\epsilon$	6×10^{-2} m·rad

Table 3-2. *Parameters of positron booster linac.*

The wake function was scaled from the SLC S-band linac, keeping the lowest 50 transverse dipole modes and assuming an uncoupled model. We examined the cases of: (1) a linear detuning with 4% total frequency spread; and (2) a Gaussian detuning of 10% total spread, where the truncation of the Gaussian distribution frequencies is at ± 2 sigma. There is about a maximum of 60% growth in the transverse offsets, for the case of a 4% linear spread. There is essentially no growth (approximately 4% maximum) in the transverse offsets, for the case of a 10% Gaussian detuning. In neither case did we include any damping beyond that of copper (we took the Qs of the modes to be 18400).

Effects of misalignments and frequency errors were not included in the results reported here, but we do not expect them to change our conclusion that the 10% Gaussian detuning should be fully adequate to control the multibunch emittance growth.

3.6.2 Aperture and Beam Optics System Parameter

The aperture of the positron linac is determined by the minimum aperture of the disk-loaded L-band accelerating structure, which will be 20 mm in radius. To allow for inevitable alignment and steering errors, the realistic transverse aperture for the beam is conservatively chosen to be 18 mm. This aperture, along with the 0.5-T longitudinal field for transverse focusing in the 240-MeV capture accelerator, defines the normalized edge emittance of the captured positron beam to be about 0.06 m·rad.

A preliminary beam optics design for transporting the 0.06-m· edge emittance positron beam through the two magnetic bends and in the common L-band linac to the positron pre-damping ring is shown in Figures 3-8 and 3-9. The design uses a scaled FODO lattice consisting of a dense array of quadrupole magnets. The quadrupole spacing is kept constant on the first accelerating section, and scaled as $\sqrt{E}$ on the remaining 23 sections. The quadrupole strengths are all kept nearly the same. The phase advance is 60° at the beginning of the lattice and gradually decreases along the linac. This design leads to a quasi-linear E scaling of the maximum beta function. The choice of relatively small phase advances is necessary to minimize the chromaticity-induced emittance growth, which has an initial full energy spread of $\pm 6\%$ at the 250-MeV point and decreases to about $\pm 2\%$ at the end of the L-band linac.

As shown in Figure 3-9, the full beam radius has shrunk to < 16 mm in the first 5-m section. From there on, the beam size is further reduced by the strong-focusing lattice to a final value of about 13 mm at the 2-GeV point. These results, obtained from first-order TRANSPORT runs, illustrate that the positron beam size can be controlled satisfactorily in the L-band linac. However, second-order TRANSPORT calculation shows that beam size control in the L-band linac is not as easy as the first order calculation indicates, mainly due to the inherent single-bunch energy spread of

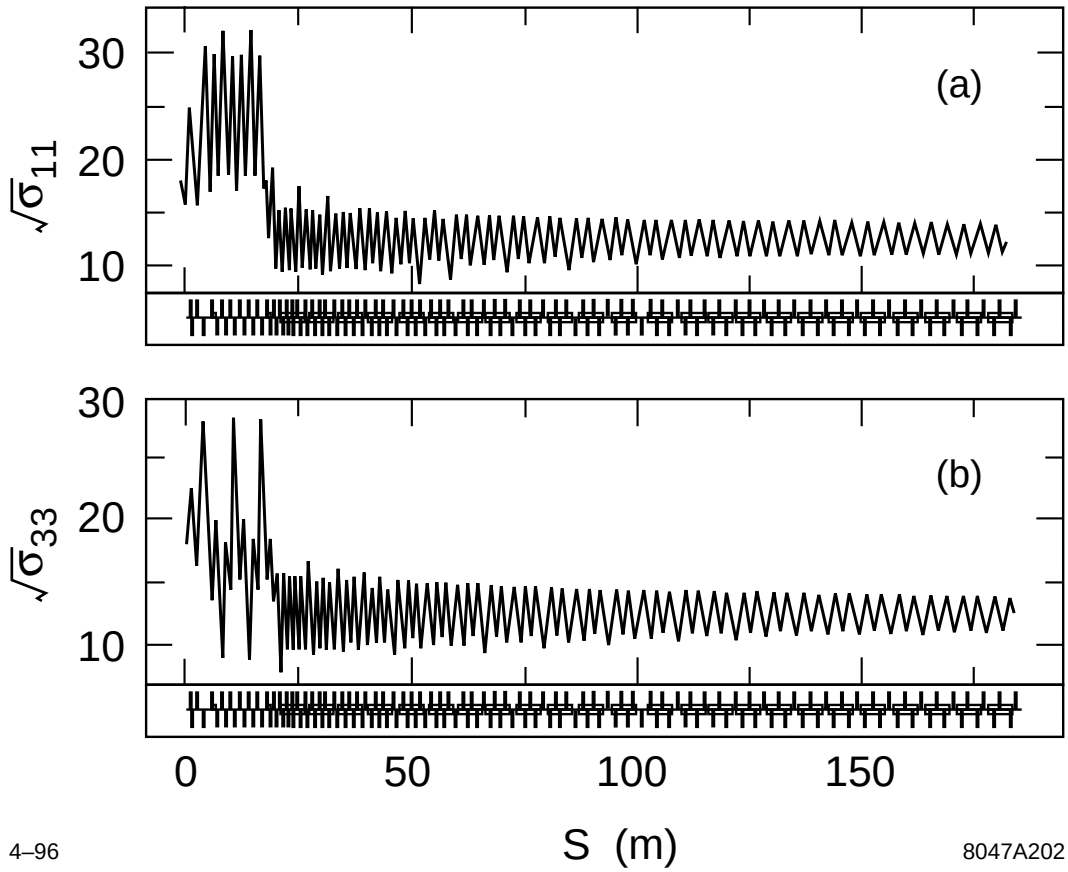


Figure 3-8. *Transverse size (edge radius) of the positron beam from the exit of the capture accelerator to the end of the L-band booster linac.*

the positron beam. This difficulty becomes more pronounced if the Δf beam-loading compensation scheme is used, which introduces large single-bunch energy spread along the linac. This is the main reason that the Δt scheme is chosen over the Δf scheme for beam loading compensation in the L-band positron linac.

3.6.3 Beam Position and Emittance Control

Due to the large transverse size of the positron beam and the limited aperture of the accelerating structure, reasonably good control (to within 0.2 mm) on the beam position is required throughout the 1.75-GeV L-band linac. A set of beam position monitors and X/Y steering correctors will be used to launch the positron beam into the linac. A beam position monitor and a pair of X/Y correctors will be installed in every gap between successive 5-m sections to maintain an optimal orbit in the L-band linac. It is also being considered to build beam position monitoring capabilities into the 5-m L-band structures by using the dipole signal so that the beam orbit in each 5-m section can be maintained at an optimum.

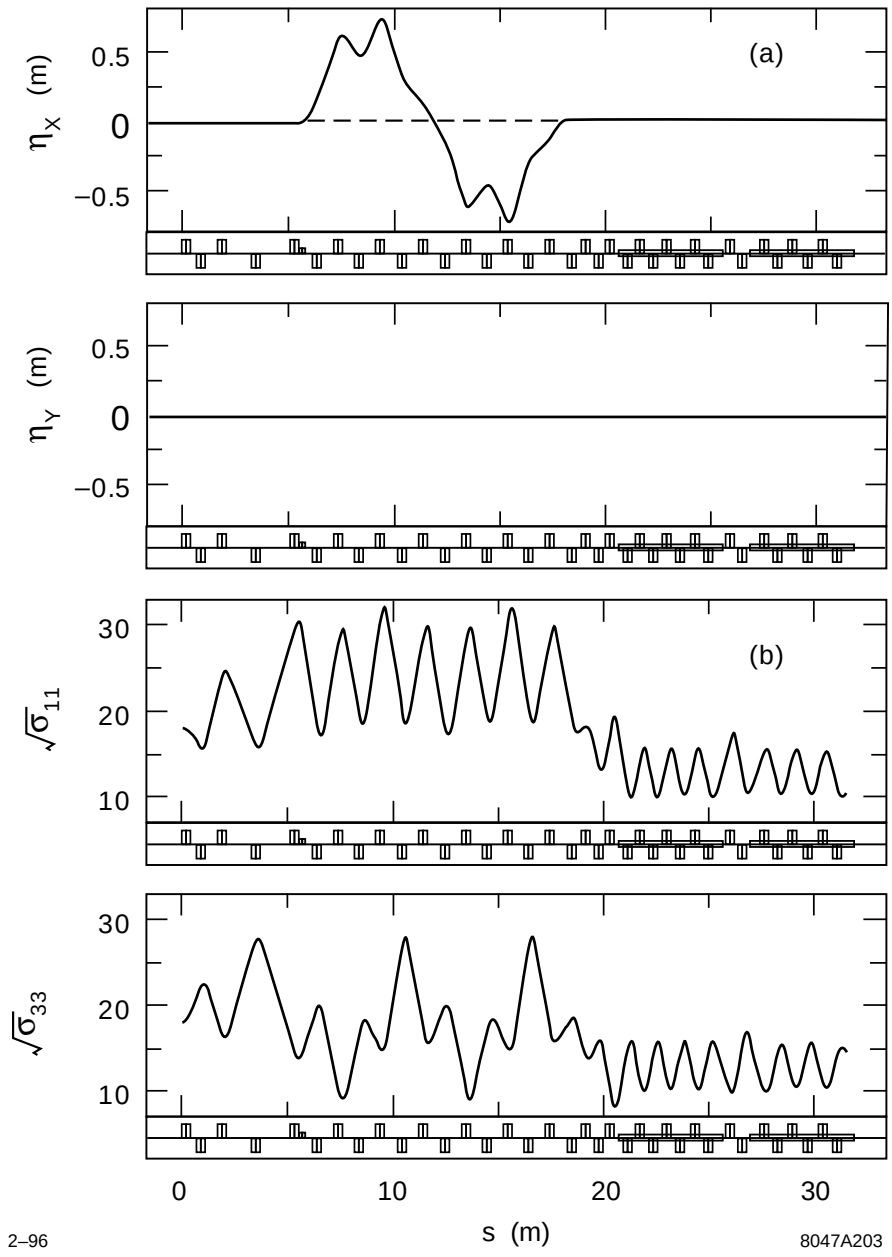


Figure 3-9. *Dispersion function and transverse beam size through the the double-bend achromat and in the first two 5-m L-band sections.*

Efficient injection into the pre-damping ring depends critically on the preservation of the beam emittance in the L-band linac. Emittance growth due to chromaticity, wakefields, beam line and magnets misalignment, etc., must be minimized.

3.7 Positron Linac

The L-band positron linac will consist of 12 accelerating modules. Each module contains two 5-m accelerating sections, which will be powered by two 75-MW L-band klystrons feeding a single SLED-I cavity. Beam loading compensation will be accomplished by using the Δt method, which offers the advantage of complete beam loading compensation at the end of each section without introducing single-bunch energy spread. The maximum attainable loaded energy gain for each module, assuming a beam current of 2.75 A across the bunch train, will be about 173 MeV. With an input beam energy of 250 MeV, the 12-module booster linac provides a large energy headroom—about 17% over the 2-GeV energy target.

3.8 Radiation Control Issues

3.8.1 Design Plan for Maintenance

Due to the high radiation activity in the areas around the target and the nominally 10-kW low-energy electron beam dump, access to these radiation-hot areas during a high-energy physics run must be delayed until the radiation activity drops to an acceptable level. As such cooling periods can be as long as several months, any maintenance work in these areas means extended downtime for the machine. A logical approach to improve the efficiency of the positron source is to add redundancy. In the present design, the positron source will have two identical positron vaults containing the positron production and collection systems, *i.e.*, from the target to the nominal 250-MeV point, adequately shielded from each other such that access to one vault is permitted while the other is in operation (Figure 3-10). The input electron beam can be directed to either system via a pair of bending dipoles which form an achromatic and isochronous beam line section. Likewise, the 250-MeV positron beam after the capture accelerator from either system can be directed into the 1.75-GeV L-band booster linac. If one system in use develops a problem during a run, we may quickly switch to the other system and continue the physics run. In the meantime, we can wait for the radiation level in the first positron vault to drop and then make an entry to repair or replace the broken components.

In addition to the dual positron vault design, efforts will be made to make the system components for positron production and collection as modular as possible to facilitate quick maintenance work. Use of materials that have long radiation decay times will be avoided.

3.8.2 Radiation Shielding

Primary Beam at Target

Figure 3-11 gives the calculated dose rate at 90° to a target-dump for various thicknesses of concrete shielding assuming a 150-kW beam of 6-GeV electrons impacting, conservatively chosen, a 30-cm-long and 10-cm-diameter

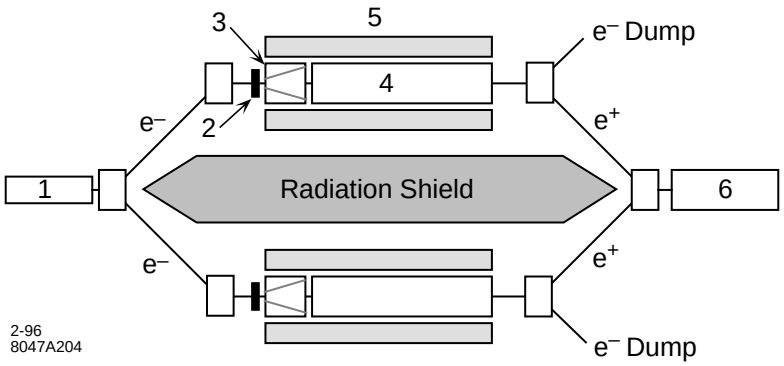


Figure 3-10. *An artist's view of the NLC positron source with two side-by-side positron production and capture systems: 1 – drive electron accelerator, 2 – positron target, 3 – flux concentrator, 4 – L-band capture accelerator, 5 – tapered-field and uniform-field solenoids, 6 – positron booster linac.*

iron target. The distance from the target to the shield is fixed at 1 m and the dose rate is determined at the outside surface of the shield. High-energy neutrons are the dominant component after about 2 m.

The integrated shield design limit is normally 500 mrem per year, which corresponds to 0.5 mrem/h for 1000 hours of operation in a year. This level, and the concrete thickness required to attain it (6 m), are indicated by the dotted lines in the figure. However, there is no need to require the radiation level to be as low as 0.5 mrem/h since the induced activity from the target and other beam components will most likely limit the occupancy time and will likely be at the level around 5 mrem/h. Thus, a 5-m concrete wall would be adequate for shielding purpose. High-Z shielding can also be used near the target to reduce the amount of concrete. For example, Fe is roughly twice as effective for shielding high energy neutrons as concrete.

Low Energy Electron Dump

After the first bending magnet, the electrons that are captured and accelerated to about 230 MeV in the capture accelerator will be bent away from the positron beam trajectory and dumped. The average power of this electron beam is on the order of 10 kW, which necessitates a water-cooled beam dump. Two such dumps, made of a material such as copper, will be built on each side.

3.8.3 Radiation Hard Components

All components in the areas of high-radiation levels, particularly near the target and the low-energy electron dump, must be made of materials highly resistant to radiation. Particular consideration must be given to the the target chamber including vacuum and cooling water seals, the tapered-field and uniform-field solenoids, all electrical cabling, various magnets, and other diagnostic instrumentation immediately downstream of the first bending dipole.

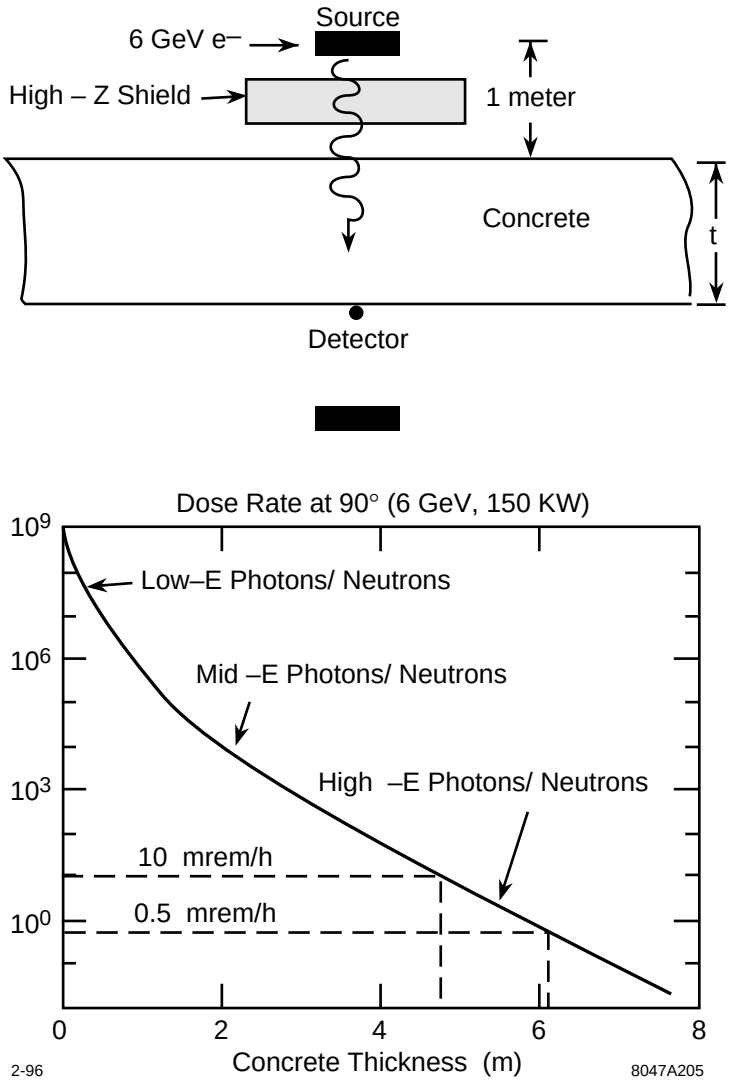


Figure 3-11. *Calculated radiation dose rate versus thickness of concrete shielding.*

3.9 Magnets

3.9.1 Solenoids

The tapered-field solenoid and the uniform-field solenoid are scaled-up versions of the SLC solenoids. Their inner radii need to be increased to approximately 11 cm and the total length to about 17 m to accommodate the four L-band accelerating sections, Their outer radii will also need to be increased accordingly to produce the required 1.2-T field at the front end which, over about 15–20 cm, tapers down to a field of 0.5 T and remains constant for the length of the uniform-field solenoid. The total power consumption of this solenoid will be on the order of 1.5 MW.

3.9.2 Quadrupoles

Two types of quadrupole magnets, one large aperture type that wraps around the L-band accelerating structures and the other regular type that wraps around regular beam pipes, will be used to construct the lattice for the 1.75-GeV L-band linac. Over 100 large-aperture quadrupoles are needed. They will have a pole-tip radius of approximately 11 cm, a length of 38 cm, and a pole-tip field of about 6 kg. Their power consumption is on the order of 15 kW each.

Another 50 or so smaller aperture quadrupoles will be used throughout the entire L-band linac, including approximately 20 for matching the positron beam from the solenoid-field region into the common L-band booster linac while maintaining an achromatic nature for this beam line section, and approximately another 20 for use in between accelerating sections and in the instrumentation section following the booster linac. This type of quadrupoles will have a pole-tip radius of 6.8 cm and a length of 36 cm, with a pole-tip field of about 6 kg.

3.9.3 Bending Dipoles

Six bending magnets, three before and three after the positron production and capture systems (Figure 3-10). are needed for directing the drive beam into either positron vault and directing the captured positron beam into the L-band booster linac. The latter three dipoles will bend the positron beam by 16°, and will have a 5.175-kg field, a 13.6-cm gap, and a length of 36 cm. The former three need to bend the 6.22-GeV electron beam by 10° each.

3.10 Diagnostics and Instrumentation

The bulk of instrumentation for the positron source, including the drive electron beam accelerator, is for routine optimization and diagnosis. While most of the instrumentation will be similar to what is used in the SLC, new instrumentation will be developed to diagnose the multibunch NLC beam. To maintain the beam quality, energy, energy spread, and bunch length will require continuous real-time monitoring before and after the target.

Many of these diagnostics and instrumentation will be used to address intensity jitter requirements. Measurements of intensity, energy, and position across the bunches of a train will need to be done to diagnose pulse-to-pulse variations in intensity out of the pre-damping ring. This is particularly important with beam intensity and energy since changes in these two attributes will lead directly to intensity fluctuations in the pre-damping ring.

Diagnostic	Bunch Train		Bunch Bundle	
	Range	Resolution	Range (per bunch)	Resolution (per bunch)
Toroid charge mon.	50–300$\times 10^{10} e^+$	$\pm$.2%	0.5–3.5$\times 10^{10} e^+$	$\pm$1%
Beam position mon.	0–2 cm	$\pm 20\mu$	0–2 cm	$\pm 20\mu$
Beam size mon.	$R = 0.5$–2 cm (edge)	$\pm$1%	$R = 0.5$–2 cm (edge)	$\pm$1%
Bunch length mon.	1–150 ns	$\pm$1 ns	10–100 ps	2 ps
Energy	250 MeV	$\pm$0.2%	250 MeV	$\pm$0.5%
ΔE at 250 MeV	$\sim$12%	$\pm$0.5%	$\sim$12%	$\pm$0.2%
Multi-strip mon.	Qualitative diagnostic for beam tuning			

Table 3-3. *Positron beam diagnostics and their specifications.*

3.10.1 Specifications

Beam diagnostic devices and their specifications for the drive electron beam have already been discussed in Chapter 2. Most of these diagnostics will also be used for the positron beam, but their specifications need to be modified slightly. Table 3-3 shows a list of the positron beam diagnostics and their specifications.

3.10.2 Beam Intensity

Bunch-to-bunch and pulse-to-pulse intensity stability will be measured after the drive beam electron gun, after bunching, and just upstream of the target. see Chapter 2 for more details.

The positron beam intensities will be measured immediately after the bending magnet which separates the electrons from the positrons, after the first L-band section in the booster linac, and at the end of the booster linac.

Bunch intensities will be sampled for beam-loading compensation feedback. High-bandwidth gap monitors appear feasible and development is in progress.

3.10.3 Beam Position

Position monitors, which measure the average bunch offset in a pulse train, will be placed in between every accelerator section near to or captured in any quadrupole magnet located there to monitor and, along with X/Y steering coils, correct and stabilize orbits via feedback loops. The centroid position of the bunches will also be important. High-bandwidth beam position monitors are feasible and development is being pursued but is expected to be expensive (see Chapter 15 for details). Placement of these monitors is lattice-dependent and in general will be used to diagnose and correct for wakefield and beam-loading problems. For the drive beam electrons, a set of these monitors to measure X, X', Y, Y' will be placed after the gun, after bunching, at the end of the S-band linac, and at two high-dispersion regions at 80 MeV and at the end of the linac (3.11 GeV for NLC-I and 6.22 GeV for NLC-II) for energy measurement.

For the positron beam, two sets will be used, one after the first bending magnet (after the electrons are separated), and the other between the L-band linac and the pre-damping ring.

3.10.4 Beam Size

The size of a small group of bunches, or perhaps of each bunch, will be determined with standard SLC-type wire scanners using high-bandwidth detectors (such as those under development in Japan). This will facilitate optical matching and emittance control tuning needed to minimize intensity jitter due to tails in the transverse distributions. For the drive beam electrons, a set to measure transverse phase space distribution will be located at the 80-MeV point, and another set at the end of the S-band linac before the double-bend achromats leading to the two positron production/capture systems. See Chapter 2 for more details. For the positron beam, two full sets will be installed, one after the first bending magnet, and the other at the end of the L-band linac before the pre-damping ring.

3.10.5 Beam Bunch Length

For the drive beam electrons, a bunch length monitor will be installed at the 80 MeV point after the beam is bunched (see Chapter 2). For positrons, three monitors will be used, one in each of the two double-bend achromats following the capture accelerators, and a third at the end of the L-band linac.

3.10.6 Energy

The energy of the drive electron beam will be monitored before the positron target, not only to facilitate its stabilization through feedback, but also to measure the effects of tuning methods and schemes for machine protection recovery. This can be done through high-bandwidth beam position monitors as described earlier.

The positron beam energy will be monitored at each available point to ensure the beam is centered in the aperture and to measure the effects of tuning methods and schemes for machine protection recovery. Monitors will be placed after the first few sections, and between the end of the linac and the pre-damping ring. The energy of the individual bunches will also be monitored at these locations. This should be done with high-bandwidth beam position monitors as previously mentioned. Wires could be used and they have the advantage of measuring energy spread as well, but are too slow for other uses.

3.10.7 Energy Spread

The energy spread of the drive electron beam will be monitored and maintained at a minimum just before the positron target . Wires with high-bandwidth detectors can be used for reliable quantitative measurements as in SLC.

The energy spread of the positron beam will be measured using multi-strip beam size monitors and wires with high-bandwidth detectors and maintained at a minimum. Measurements will be done at two dispersion locations, *i.e.*, in the double-bend achromat using foils as in the SLC, and in the linac to pre-damping ring transfer line.

3.10.8 RF Phase and Amplitude Monitoring

The shape of rf pulses will be sampled at frequent intervals to ensure phase and amplitude stability. Slow drifts due to diurnal effects, etc., will be compensated with feedback loops. Pulse-to-pulse sampling and correlation with beam changes can be used to determine severity. The SLC prototype has these features.

3.11 Feedback and Stability

Software-driven feedbacks with high-speed data acquisition, calculation, and device control were integral to the success of the SLC prototype (see Appendix D), and will be used in the NLC positron source to control orbit, energy, and beam-loading compensation.

A significant fraction of the drive beam power will be deposited into the first L-band section in the positron capture accelerator. The amount of power deposition in the first section depends on the machine repetition rate and on the bunch structure (*i.e.*, number of bunches per pulse) as well. In order to maintain rf phase and amplitude stability in the first section during periods of machine rate and bunch structure changes, the temperature of the structure must be stabilized by appropriately adjusting the cooling water temperature with special hardware and control software. Of course, thermal stability is equally important for the remaining three sections in the capture accelerator. Depending on the amount of power deposition into these sections, similar temperature control hardware and software may also be necessary for them.

An intensity feedback using the gun pulsar with the capability of changing individual pulses is desirable for long-term stability. A beam-loading compensation feedback will be needed to keep bunch intensity variations minimized as pulse train intensity may change.

3.11.1 Intensity Uniformity Specifications

The specifications on the pulse-to-pulse, or train-to-train, and bunch-to-bunch intensity jitters for the positron beam are 0.5% and 2%, respectively. Assuming there is no intensity jitter amplification (or growth) from the drive electron beam to the positron beam, these jitter specifications should also apply to the drive electron beam. A stable DC electron gun coupled with a jitter limiting aperture downstream is expected to provide an electron beam meeting these jitter specifications. Beam intensity diagnostics will have a .2% resolution for pulse intensity measurements and a 1% resolution for bunch or bunchlet intensity measurements.

3.11.2 Transverse Orbit Stability

Feedbacks for maintaining transverse orbit will be used throughout the entire system, but will be particularly important at critical points such as at the positron target or at the entrance to the pre-damping ring. The SLC prototype feedback is adequate and beam position monitors that measure the average bunch offset in a pulse train will be used.

3.11.3 Energy Control

Energy feedback loops will be needed in three places: before the target to control the energy of drive beam electrons; after the capture accelerator to control the energy of positrons to be launched into the linac; and in the linac-to-ring line before the entrance to the pre-damping ring.

3.12 Operations and Tuning Procedures

Operations and tuning procedures will be designed to grow toward higher levels of automation as in the SLC prototype. Initial procedures for commissioning and optimizing the positron source will lead to the development of algorithms and feedbacks for maintaining the positron source.

The systems in the NLC positron source are very similar to the SLC prototype and many SLC procedures will be transferable for use on the NLC. However, modifications and additions will be made to address issues involving the significantly higher beam power, multibunch operation, and machine protection schemes.

3.13 Control System Needs

3.13.1 Specification

The control system needs to monitor and control some of the devices on the time scale of the repetition rate of the machine, *i.e.,* 180 Hz for NLC-I or 120 Hz for NLC-II. This will require effective software-based feedbacks such as are used in the SLC prototype. On a slower time scale, the system must monitor and control a variety of devices to include rf amplitude and phase, magnet fields, and beam monitors of all types. These are just a few of the generic tools necessary for successful NLC operation.

3.13.2 Special Requirements

The positron source has two special requirements on the control system. (i) For the L-band capture accelerator, especially the first section, rf phase control for maintaining a constant positron yield will require thermal stability of the accelerating structures. During rate recovery from a machine protection rate change, thermal stabilization in these sections will require a pre-established feedforward control that promptly adjusts the temperature of the structure cooling water. (ii) The backup power system needs to be controlled in such a way that uninterrupted power is delivered to the various vacuum pumps of the positron target chamber in the event of a power outage.

3.14 Other Considerations

While the SLC positron source has met its design goal in terms of beam intensity and demonstrated good reliability over many years of operation, its success has been marred by the extensive and constant tuning effort required for operating the source and by excessive beam intensity jitter often experienced. The difficulty in operating the SLC source is known to be caused mainly by the lack of dedicated accelerators for both the drive beam and the positron beam which greatly limits the tuning flexibility. The excessive intensity jitter appears attributable to successively smaller apertures in the positron system starting from the capture accelerator to the damping ring. Accompanying a loss of beam intensity at each aperture, there is generally a concurrent growth in the intensity jitter. The accumulative effect leads to an increase in the positron intensity jitter, after being extracted out of the damping ring, by $\geq$75% over that of the drive beam.

Since the NLC positron source will have dedicated accelerators for the drive beam and the positron beam, its operational characteristics are expected to be improved significantly over its SLC counterpart. Also, since the smallest aperture in the NLC positron system by design is given by the capture accelerator immediately following the production target, minimal intensity losses are expected for the positron beam in the rest of the machine, including the L-band linac, the pre-damping, and the main damping ring. Judging from the SLC experience, one may expect that intensity jitter growth from the drive beam to the positron beam in NLC may be insignificant. Thus, producing a sufficiently stable drive beam may be the only critical task. This is, however, not to say that beam intensity jitter in the NLC positron source is a trivial issue. On the contrary, since intensity stability is such a critical issue to the success of the NLC, intensity jitter in the NLC positron source warrants critical design considerations.

3.15 Summary

The NLC positron source has a conventional design based on the SLC positron source and will be built with existing technology only. The significantly higher beam intensity for NLC is realized by the combined use of a larger drive electron beam on the target and an L-band positron accelerator. The former permits a quadratic increase in the pulse energy of the drive beam, while the latter permits the accelerating structure to have a minimum iris radius slightly more than twice as large as that of the SLC S-band structure, which translates into a $>$16-fold increase in the four-dimensional transverse phase space admittance. The source upgrade from NLC-I to NLC-II is straightforward, simply requiring an 80% increase in the drive beam size (area) and a doubling of the drive beam energy (from 3.11 GeV to 6.22 GeV). Both the NLC-I and NLC-II sources are conservatively designed to have large intensity safety margins—about a factor of 2.5 higher than the highest conceivable operating intensity at the interaction point—to insure against possible unforeseen beam losses.

The design employs a rotating positron target, intended for eliminating beam intensity modulations induced by target motion, with multi-stage differential vacuum pumping along the target drive shaft. Two identical, inter-switchable, positron production and capture systems will be built side by side to improve the source efficiency.

References

[Ecklund 1981] S. Ecklund, "Positron Target Material Tests", SLAC–CN–128 (1981).

[Helm 1962] R.H. Helm, 1962, "Adiabatic Approximation for Dynamics of a Particle in the Field of a Tapered Solenoid", SLAC–4 (1962).

[Kulikov 1991] A.V. Kulikov, S.D. Ecklund, and E.M. Reuter, 1991, "SLC Positron Source Pulsed Flux Concentrator", SLAC–PUB–5473 (1991), in Conference Record of the 1991 IEEE Particle Accelerator Conference, San Francisco, CA, 2005 (1991).

[Lynch 1989] H.L. Lynch, 1989, "ETRANS", SLAC Memorandum (1989).

[Nelson 1985] W. Nelson, H. Hirayama and D. Rogers, 1984, "The EGS4 Code System", SLAC–Report–265 (1985).

[SLC 1984] SLC, *SLC Design Handbook*, SLAC (1984).

[Thompson 1990] K.A. Thompson, and R.D. Ruth, 1990, "Controlling transverse multibunch instabilities in linacs of high energy linear colliders", SLAC–PUB–4801, in *Phys. Rev. D* **41**, 964 (1990).

Contributors

- S. Ecklund
- P. Emma
- G. Gross
- A. Kulikov
- R. Miller
- W. Nelson
- L. Rinolfi
- H. Tang
- K. Thompson
- J. Turner
- T. Umemoto
- V. Vylet
- D. Yeremian
- Z. Li

4

NLC Damping Rings

Contents

4.1 Introduction

The NLC damping rings are designed to damp the incoming electron and positron beams to the very small emittances needed for collisions. The rings have three purposes: (1) damp the incoming emittances in all three planes, (2) damp incoming transients and provide a stable platform for the downstream portion of the accelerator, and (3) delay the bunches so that feedforward systems can be used to compensate for charge fluctuations.

To meet these goals, we have designed three damping rings: two identical main damping rings, one for the electrons (EDR) and one for the positrons (PDR), and a pre-damping ring for the positrons (PPDR). The pre-damping ring is needed because the incoming positron emittance is much larger than that of the electron beam. Each damping ring will store multiple trains of bunches at once. At every machine cycle, a single fully damped bunch train is extracted from the ring while a new bunch train is injected. In this manner, each bunch train can be damped for many machine cycles.

The parameters of the two main damping rings (MDR) are similar to the present generation of synchrotron light sources and the B-factory colliders that are being constructed in that they must store high-current beams ($\sim$1 A) while attaining small normalized emittances. Table 4-1 compares the NLC ring parameters with those of the SLAC B-factory Low Energy Ring (PEP-II LER), the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory (LBNL), and the Accelerator Test Facility (ATF) damping ring being constructed at KEK in Japan to verify many of the damping ring design concepts. In particular, the stored beam currents are less than half of what the PEP-II LER has been designed to store while the emittance, energy, and size of the rings are similar to those of the ALS and the ATF.

Thus, these other rings will be able to test and verify many of the accelerator physics issues that will arise in the NLC damping rings. In particular, strong coupled-bunch instabilities will be studied in the high current B-factories while issues associated with the very small beam emittances, such as intra-beam scattering and ion trapping, will be studied in the ALS and the KEK ATF damping ring. In addition, the PEP-II LER and the Advanced Photon Source (APS) at Argonne National Laboratory will be able to study the photoelectron-positron instability that is thought to arise in positron storage rings.

These similarities with other rings have also simplified the design process and we have been able to and will continue to benefit from the experience of these other accelerators. For example, the damping ring rf system is based on those developed for the SLAC B-factory and the ATF damping ring, the multibunch feedback systems are based upon the feedback systems which were designed for the SLAC B-factory and successfully verified on the ALS, and the vacuum system is similar to that used by the ALS. Furthermore, the design uses "C" quadrupole and sextupole magnets similar to those designed at the ALS and the APS at Argonne, a high-field permanent magnet wiggler very similar to a design developed and installed at Stanford Synchrotron Radiation Laboratory (SSRL), and a double kicker system for extraction similar to one to be tested in the KEK ATF.

Because the damping rings must form a stable and reliable platform for the rest of the linear collider, we have designed the rings conservatively. Whenever possible we have used concepts and designs that have been experimentally verified or will be verified in the near future and have not adopted some of the more exotic, although attractive, concepts that have been suggested. The one possible exception from this rule is the use of the TME lattice for the main damping rings instead of a standard TBA or Chasman-Green lattice. We felt that the advantages of the TME lattice—simpler and fewer bending magnets—justify this choice.

The damping ring complex consists of the injector lines for the damping rings, the positron pre-damping ring, the two main damping rings, and the transfer line between the pre-damping ring and the main positron damping ring. Section 4.2 gives an overview of the damping ring complex and defines the system requirements. Next, the determination

	NLC MDR	PEP-II LER	LBNL ALS	KEK ATF
Energy (GeV)	2.0	3.1	1.5	1.5
Circ. (m)	220	2200	200	140
Current (A)	1.2	3	0.6	0.6
$\gamma\epsilon_x$ (10^{-6}m-rad)	3	400	10	4
$\gamma\epsilon_y$ (10^{-6}m-rad)	0.02	16	0.2	0.04

Table 4-1. *Comparison of NLC main damping rings with parameters of other rings.*

of the parameters, various scaling relations, and some of the design tradeoffs that were made are discussed. The goal of this sections is to document our parameter choices so others may then explore alternate methods of optimization.

Section 4.3 presents the parameters and the optical designs of the main damping rings, pre-damping ring, injection and extraction lines, and PPDR-to-PDR transfer line. It also contains a description of the trajectory correction, emittance control and matching requirements, dynamic apertures, and the transport of the polarization through the damping ring complex.

The collective limitations are discussed in Section 4.4 and then, the rf and vacuum systems are described in Sections 4.5 and 4.6. This is followed by specifications of the feedback systems, supports and alignment techniques, magnet design and instrumentation. Finally, we discuss the ground vibration effects and some stability issues.

4.2 System Overview and Parameter Determination

In this section, we will first provide a brief overview of the damping ring complex and then we will enter a more detailed discussion of the requirements and parameter determination.

4.2.1 Overview

The damping ring complex consists of one damping ring for the electrons and both a pre-damping ring and a main damping ring for the positrons; as will be discussed, two positron rings are needed because the incoming positron emittance is very large. It also includes the transport line between the positron pre-damping ring and the main damping ring as well as elements in the extraction lines that complete the extraction system for the main damping rings. Finally, the damping ring complex includes the injection matching sections upstream of the rings which provide matching for the transverse phase space as well as the energy compressors to match the longitudinal phase space and spin rotators to orient the particle spin in the vertical direction for injection into the damping ring complex.

At present, most of the studies have concentrated on the main damping rings. Although relatively little of the detailed engineering has been performed, we have a good outline of the design and the problems we may encounter. In particular, we have identified and described solutions for the most difficult issues: these are the dynamic aperture, the vertical emittance, the impedance and instabilities, and the stability and jitter in both longitudinal and transverse phase space. As stated, much of the design rests on work being performed for the B-factories presently being constructed and the KEK ATF damping ring.

One area which has not been studied sufficiently is the design of the bending magnets. Although we have confidence that the magnets could be constructed, they are high-field combined-function magnets that need additional research to

fully verify their feasibility. We feel that it is justified to push their design since, if a design can be found, they are static devices that will not cause further trouble. If, instead, a reasonable design cannot be found, it is straightforward to scale the energy of the ring complex and reduce the requirements, although one also loses some of the tuning range in the present design and likely increases the cost.

In most aspects, the pre-damping ring has relatively loose requirements. The emittance and damping time requirements are not severe. The required beam stability is not very significant. The two difficult items are attaining the dynamic aperture and the injection/extraction kicker systems. At this time, the design is still in a preliminary stage but we have addressed the two issues just mentioned.

Main-Damping Rings

The Electron and Positron Damping Rings (EDR and PDR) are designed to damp beams with an injected emittance of $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad to an emittance of $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad. The rings will operate at 180 Hz and must provide sufficient damping to decrease an injected emittance by over four orders of magnitude.

The lattice is based on detuned Theoretical Minimum Emittance (TME) cells which were chosen because of eased requirements on the combined-function bending magnets. The damping is performed using both high-field bending magnets and ten 2.5-m sections of damping wiggler.

The rings operate with four trains of 90 bunches. The bunch trains are injected onto and extracted from the closed orbit using pulsed kickers and DC septa. To avoid coupled-bunch instabilities the rf cavities are based on those of the PEP-II B-factory and a transverse bunch-by-bunch feedback system is used. The rings are designed to operate with maximum bunch charges of 1.57×10^{10} particles; this is roughly 20% more than the maximum needed at the IP. In addition, the electron source has been designed to provide additional charge to allow for at least 10% losses during injection into the electron damping ring. Similarly, the positron source has been designed to produce at least 20% additional charge to provide for losses during injection into the pre-damping ring.

Finally, because the rings must generate extremely small beam emittances, there are tight jitter and alignment tolerances. Extensive effort has been made to include cancelations and tuning procedures in the design that will ease the tolerances to reasonable levels. In particular, all of the quadrupoles and all of the sextupoles will have independent power supplies. This will facilitate beam-based alignment as well as matching of the lattice functions.

Pre-Damping Ring

The Positron Pre-Damping Ring (PPDR) is designed to damp the large emittance positron beam from the positron source to an emittance of roughly $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad. At this point, the positrons are injected into the positron main damping ring (PDR) where they are damped to the desired final emittances. The pre-damping ring allows us to decouple the large aperture requirements for the incoming positron beams from the final emittance requirements of the linear collider.

The pre-damping- ring does not need to produce flat beams. Thus, to maximize the damping of the transverse phase space, the ring has a damping partition number $J_x \approx 1.4$ and operates on the coupling difference resonance. This increases the damping in both the horizontal and vertical planes. Furthermore, like the main damping ring, the ring damps multiple trains of bunches at once, the number of which is determined by the ring circumference. The initial design was only 114 m in circumference and stored only two trains. The present design is roughly 50% larger and stores three bunch trains. Although this increase in circumference will certainly increase the cost of the ring, the previous design was very tightly packed making the component design and maintenance substantially more difficult; experience with the SLC damping rings strongly suggests that the increase in circumference is cost-effective.

The magnets and vacuum system are being designed to provide sufficient aperture to accept a 2-GeV beam with an edge emittance of $\gamma\epsilon_{x,y} = 0.09$ m-rad and $|\delta p/p| \leq 1.5\%$ plus 2-mm clearance for misalignments and mis-steering. Given the nominal injected edge emittance of $\gamma\epsilon_{x,y} = 0.06$ m-rad, this provides a substantial margin for injection and internal mismatches. In addition, the injectors specifications assume that roughly 20% of the delivered charge is lost at injection into the pre-damping rings while the pre-damping rings are designed to operate with a maximum bunch charge that is roughly 20% greater than the maximum required at the IP.

Like the main damping rings, all of the quadrupoles and all of the sextupoles will have independent power supplies. This will facilitate beam-based alignment as well as matching of the lattice functions which is especially important in the pre-damping ring because of the limited aperture.

Injection Lines

The injection lines to the damping rings from the booster linacs are relatively straightforward transport lines. They have five primary functions:

- Provide diagnostics to monitor the beam energy, energy spread, transverse positions, and emittances.
- Provide transverse betatron matching for injection into the rings.
- Compress the incoming energy spread to match the longitudinal phase space to the damping rings.
- Rotate the longitudinal polarization of the electron beam into the vertical plane for injection into the electron damping ring.
- Include a beam dump after the diagnostics but before injection into the rings.

The spin rotation requires at least a 20° arc to rotate the longitudinal spin into the horizontal plane. We have designed a 60° arc (270° spin rotation) which is more suitable for the energy compressors. Finally, although the positron beam is not polarized, we are designing the positron system so that it could be converted to operate with polarized electrons for future $\gamma - \gamma$ or $e^- - e^-$ collisions. Thus, the geometrical constraints imposed by the spin rotation are applied to both the electron and positron injection lines.

Extraction Lines

The extraction lines from the damping rings have two purposes: they match the beams into the beginning of the spin rotator/bunch compressor beam lines that are described in Chapter 5, and they contain final elements for the extraction systems.

Bypass and Transfer Lines

There are four additional bypass and transfer lines that allow each of the rings to be bypassed and transfer the beams from the positron pre-damping ring to the main damping ring. These contain emittance diagnostics and beam dumps to allow tuning of the injector systems without operating portions of the ring complex.

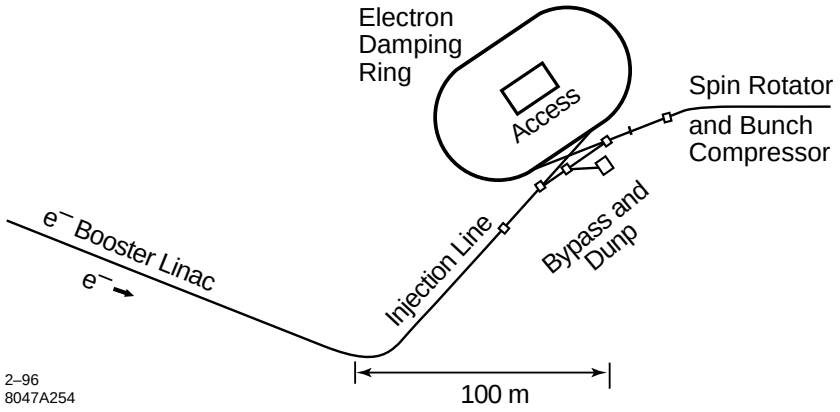


Figure 4-1. *Layout of electron damping ring system along with the two bunch compressors and the electron source.*

4.2.2 Damping Ring Complex Layout

The electron and positron damping ring systems are located at the low-energy ends of the respective linacs. The two systems are separated by roughly 20 km. On the extraction side, the systems are constrained by the prelinacs, the spin rotator in the low-energy bunch compressor which performs a 20° bend, and the bunch compressor turn-around (180° arc). At the injection end, the systems are constrained by the source linacs; the electron source is roughly 300 m in length while the positron source is in excess of 1 km.

The electron damping ring system includes an injection line with a 60° bend, and the ring injection and extraction sections which yield a net bending of roughly 20°. Figure 4-1 is a schematic of a possible layout.

The positron damping ring system also includes an injection line with a 60° bend, and the injection and extraction sections for both the pre-damping ring and the main damping rings. Figure 4-2 is a schematic of a possible layout which allows for the long positron source system.

4.2.3 Requirements

At this point, we will enter a more detail description of the damping ring requirements and the parameter determination. The primary requirements of the NLC damping rings are summarized in Table 4-2 where the limiting parameters are listed for each of the NLC energy upgrades. The rings must produce electron and positron beams with emittances of $\gamma\epsilon_x \leq 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y \leq 3 \times 10^{-8}$ m-rad at a repetition rate as high as 180 Hz. The beams in the damping rings consist of multiple trains of 90 bunches with a maximum single bunch charge of 1.31×10^{10} (this is 5% higher than the required charge at the IP to allow for beam loss due to collimation).

To satisfy these requirements, the damping ring complex is designed to operate with the parameters listed in Table 4-3 which exceed the requirements of all presently considered NLC upgrades. To provide operational flexibility, we are designing the rings to operate with a peak current roughly 20% higher than the nominal peak current. Many of the limiting requirements (damping times and bunch train lengths) are set by the NLC–Ic (500-GeV-c.m.) design and not the higher energy upgrades. The only parameter that increases in difficulty as the center-of-mass energy is upgraded is the beam charge. This is relevant for the design of the rf system and the vacuum chamber, and the study of collective effects (many of which depend upon the rf and chamber designs).

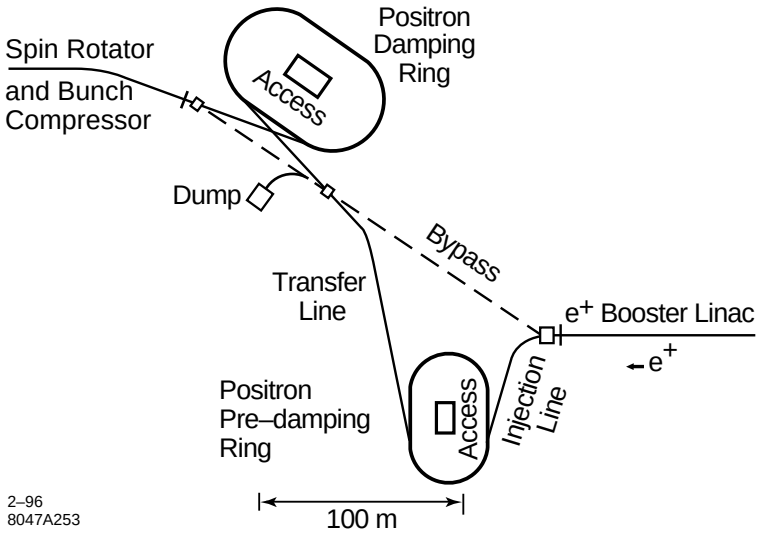


Figure 4-2. *Layout of positron damping ring complex along with the two bunch compressors and the positron source.*

	NLC–Ic (500 GeV)	NLC–IIc (1 TeV)	NLC–III (1.5 TeV)
$\gamma\epsilon_{x\,ext}$ $(10^{-6}$ m-rad$)$	3	3	3
$\gamma\epsilon_{y\,ext}$ $(10^{-8}$ m-rad$)$	3	3	3
Charge/Bunch (10^{10})	0.89	1.31	1.31
Bunches/Train	90	90	90
Bunch Spacing (ns)	1.4	1.4	1.4
Repetition Rate (Hz)	180	120	120

Table 4-2. *Damping ring beam requirements for NLC designs.*

$\gamma\epsilon_{x\,eq}$ / $\gamma\epsilon_{x\,ext}$ $(10^{-6}$ m-rad$)$	3.0 / 3.0
$\gamma\epsilon_{y\,eq}$ / $\gamma\epsilon_{y\,ext}$ $(10^{-8}$ m-rad$)$	2.0 / 3.0
Charge/Bunch	1.57×10^{10}
Bunches/Train	90 bunches/train
Bunch Spacing (ns)	1.4
Repetition Rate (Hz)	180

Table 4-3. *Requirements for NLC damping ring complex.*

	Injected Electrons	Injected Positrons
Charge/Bunch	1.9×10^{10}	2.5×10^{10}
$\gamma\epsilon_{x,y}$ (m-rad)	1×10^{-4} (rms)	0.06 (edge)
$\sigma_{\Delta E/E}$ (%)	1 (FWHM)	2 (FWHM)
$\Delta\gamma J$ (m-rad)	0.5×10^{-4} (rms)	0.03 (edge)
$\Delta E/E$ (%)	0.2 (rms)	0.2 (rms)

Table 4-4. *Electron and positron beams at entrance to damping rings.*

Table 4-4 lists the parameters of the beams assumed from the electron and positron injectors. The incoming electron emittance is specified as an rms value while the incoming positron emittance is specified as an edge emittance. The transverse injection errors are specified in units of the normalized actions and the permissible energy errors are given in %; the injection errors correspond to oscillation amplitudes that are equal to the rms and edge beam sizes for the electrons and positrons, respectively. Both of these values are set by the damping time requirements and not by the dynamic aperture or the multibunch sensitivity to transients both of which have looser tolerances.

4.2.4 Parameters: Damping Times and Equilibrium Emittances

At this point, we can discuss the parameters of the damping rings, in particular, the beam emittances and the required number of damping times. In a damping ring, the required number of damping times and the extracted emittances determine the primary parameters of the design. The emittances from the injectors are summarized in Table 4-4 for both the electrons and positrons and the requirements for the extracted beams are listed in Table 4-3. In examining these parameters, two points are immediately obvious.

First, the requirements for the positron and electron beams are extremely different. Because of the large incoming emittance, the injected positron beam needs substantially more damping and a much larger vacuum aperture. Therefore, we will use a pre-damping ring to damp the injected positron emittance to the level of the incoming electron beam. This allows us to partially separate the problems of large dynamic aperture, small equilibrium emittance, and fast damping. In addition, it allows us to design identical main damping rings for the electrons and positrons while the pre-damping ring is optimized for just the incoming positrons.

Second, the main damping rings must produce beams with a 100:1 emittance ratio, the vertical emittance being much smaller than the horizontal. Thus, the main rings will have to operate far from the coupling resonance. In a ring designed to produce flat beams, the vertical damping time is more critical than the horizontal and there is only one method of decreasing it: increased synchrotron radiation due to high-bending fields or damping wigglers. We will briefly discuss and discard other options subsequently.

In contrast, the pre-damping ring can operate on the difference coupling resonance since the extracted emittances of the pre-damping ring do not need to be asymmetric; the SLC damping rings operated in this mode until 1992 when it was decided to collide flat beams at the IP. Operating on the coupling resonance reduces the equilibrium emittance by a factor of two and it couples the horizontal damping partition into both planes, effectively increasing the vertical damping rate at the expense of the horizontal.

Now, we can specify the damping times and equilibrium emittances. Ideally, these determine the emittance of the extracted beam as expressed by:

$$\epsilon = \epsilon_i e^{-2N_\tau} + \epsilon_{eq}(1 - e^{-2N_\tau}) \quad , \tag{4.1}$$

	Main Damping Rings	Pre-damping Ring
Max. Injected $\gamma\epsilon_x$, $\gamma\epsilon_y$ (m-rad)	1.5×10^{-4}, 1.5×10^{-4}	0.09, 0.09 (edge)
Extracted $\gamma\epsilon_x$, $\gamma\epsilon_y$ (m-rad)	3×10^{-6}, 3×10^{-8}	1.5×10^{-4}, 1.5×10^{-4}
Equil. $\gamma\epsilon_{eq\,x}$, $\gamma\epsilon_{eq\,y}$ (m-rad)	3×10^{-6}, 2.0×10^{-8}	5×10^{-5}, 5×10^{-5}
$N_{\tau\,y}$	4.8	3.2
$\tau_{\rm eff}$ (ms)	1.16	1.74

Table 4-5. *Damping time requirements for NLC damping rings.*

where ϵ_i is the injected emittance, ϵ_{eq} is the equilibrium emittance, and N_τ is the number of damping times the beam has been stored. Because the damping needed in the vertical plane is greater than or equal to that in the horizontal while the damping rate in the vertical is less than or equal to that in the horizontal $(J_x \geq J_y)$, the number of damping times is determined from the vertical emittances. Table 4-5 lists the injected emittances from Table 4-4 assuming that the injection-offsets fully filament and add to the beam emittance. Then, we calculate the required number of vertical damping times for the pre- and main damping rings. We have assumed that the pre-damping ring is operating on the coupling resonance so that the extracted emittances are equal and we have assumed equilibrium vertical emittance of $\gamma\epsilon_{y\,eq} = 2.0 \times 10^{-8}$ m-rad for the main damping ring. Obviously, assuming a larger equilibrium vertical emittance in the main ring increases the required damping and therefore the difficulty of the design. Although requiring smaller a equilibrium emittances imposes stricter magnet and alignment tolerances, it is felt that the necessary alignment tolerances can be attained using beam-based alignment methods.

The NLC main damping rings will damp an injected beam vertical emittance of 1×10^{-4} to 7×10^{-9}. At the same time, we will design the rings to attain an equilibrium vertical emittance less than 2×10^{-8} m-rad; the tolerances are being specified to attain an expected vertical emittance of 1.5×10^{-8} m-rad. Thus, we are providing more damping that necessary. We believe that it is worthwhile investing in the additional damping capability since it provides stability against fluctuations in the incoming beam. Specifically, it allows for injection mismatches or other effects [1] that can increase the effective injected emittance by 50%. This additional damping will be reflected in the injection tolerance of $\gamma J_y \leq 0.5 \times 10^{-4}$. Finally, if necessary, the damping rates can be increased by operating the rings at a higher energy; the rings are being designed to operate at an energy between 1.8 GeV and 2.2 GeV with the nominal energy being 1.98 GeV.

Similar arguments apply to the pre-damping ring. Again, we want to keep the equilibrium emittance small to reduce the damping requirements. Because the pre-damping ring is coupled, the horizontal and vertical emittances and the horizontal and vertical damping rates are equal. In this case, we have chosen to design to coupled emittances of

$$\gamma\epsilon_{x,y\,\rm eq} \leq 5 \times 10^{-5}\ \text{m-rad} \qquad \text{(pre-DR)} \quad . \tag{4.2}$$

Although we have not included any margin in the damping time, we should note that in calculating the required number of damping times for the positron beam, we have used an effective rms injected emittance equal to 70% of the edge emittance.[2] At the time of this writing, we did not have details of the incoming positron beam distribution but

[1] In the SLC electron damping ring, the injected beam does not appear to damp at the expected damping rate during the first millisecond. The cause of this problem is unknown.

[2] When a beam is damped in a ring, the injected distribution is not modified until the emittance is damped close to the equilibrium emittance of the ring. Since an aperture limited beam, such as the incoming positrons, will not have a gaussian distribution, it is not reasonable to parameterize the emittance with the rms value. Instead, because we are interested in the core density of the beam, we will define the effective emittance using the FWHM: $\sigma_{\rm eff} = 0.4$ FWHM. Using this parameterization and assuming that the initial distribution is uniform, $\epsilon_{\rm eff} \approx 0.7\epsilon_{\rm edge}$. This is the value assumed for the pre-damping ring design.

this is undoubtedly an over-estimate of the effective rms emittance and therefore provides a margin on the damping requirements.

Finally, we can specify the required equilibrium horizontal emittance in the main damping rings. Since the rings are uncoupled and $J_x \geq J_y$, the horizontal damping time is less than or equal to the vertical. Therefore, the residual injected horizontal emittance is negligible and the horizontal emittance of the extracted beam is very nearly equal to the horizontal emittance of the ring:

$$\gamma\epsilon_{x\,\mathrm{eq}} \leq 3 \times 10^{-6}\,\text{m-rad} \qquad \text{(MDR)} \quad . \tag{4.3}$$

Now we can calculate the required damping times of the rings. The damping times are determined from the desired repetition rate (f_{rep}), the required number of damping times per bunch train (N_τ), and the number of trains stored simultaneously in the rings (N_{train}):

$$\tau_{x,y} \leq \frac{1}{f_{\mathrm{rep}}}\frac{N_{\mathrm{train}}}{N_\tau} \quad . \tag{4.4}$$

The maximum number of trains stored in the rings depends upon the length of the bunch trains and the rise and fall times of the injection and extraction kickers. The kicker rise and fall times depend upon the required kicker amplitude and stability [Mattison 1995]. The SLC damping ring kickers have rise and fall times of 60 ns. Although the stability requirements are tighter in the NLC main damping rings than in the SLC rings, the amplitude and physical design constraints are looser. Thus, it is thought that 60 ns will be sufficient for the rise and fall time of the NLC main damping ring kickers. The situation is different in the pre-damping ring. Here the amplitude requirements are more severe than those in the SLC rings but the stability requirements are comparable and again the physical design constraints are looser. Thus, again we feel that 60 ns will be sufficient for the rise and fall time of the pre-damping ring kickers.

Note that when damping multiple trains of bunches at the same time, the damping time *per se* is not the relevant parameter. Instead, we are concerned with the damping time scaled by the number of bunches in the ring which is roughly proportional to the ring circumference; two rings, one with twice the damping time but storing twice as many trains, will have the same damping performance. Thus, we define an effective damping time:

$$\tau_{\mathrm{eff}} \equiv \frac{\tau}{N_{\mathrm{train}}} = \tau\frac{T_{\mathrm{train}}}{T_0} \quad , \tag{4.5}$$

where T_{train} is the length required per bunch train including the gap between the trains, and T_0 is the ring circumference. The effective damping time then depends upon the repetition rate and the number of damping times per bunch. In the NLC main damping rings, the effective damping time is:

$$\tau_{\mathrm{eff}} \leq \frac{1}{f_{\mathrm{rep}}N_\tau} = 1.16\,\text{ms} \qquad \text{(EDR, PDR)} \quad , \tag{4.6}$$

while in the pre-damping ring

$$\tau_{\mathrm{eff}} \leq 1.74\,\text{ms} \qquad \text{(PPDR)} \quad . \tag{4.7}$$

4.2.5 Scaling Relations

Equations 4.2, 4.3, 4.6, and 4.7 determine the basic parameters. Initially to study these parameters we ignored the option of damping wigglers, assuming an isomagnetic ring, and we neglect the effect of intrabeam scattering. In this case, we can write simple expressions for $\gamma\epsilon_{x\,eq}$ and $\tau_{y\,\mathrm{eff}}$, the two quantities we wish to calculate:

$$\tau_{y\,\mathrm{eff}} = 1.69 \times 10^{14}\frac{T_{\mathrm{train}}}{\gamma^3 G_B} = \frac{2.88 \times 10^{12} T_{\mathrm{train}}}{B_0[kG]\gamma^2} \tag{4.8}$$

$$\gamma\epsilon_{x\,eq} = 3.84\times10^{-13}\frac{\gamma^3}{J_x}\langle\mathcal{H}\rangle_B G_B = \frac{65T_{\rm train}}{\tau_{y\,\rm eff}}\frac{\langle\mathcal{H}\rangle_B}{J_x} \quad . \tag{4.9}$$

Here, G_B is the inverse bending radius of the bending magnets, $\langle\mathcal{H}\rangle_B$ is the average of $\mathcal{H}_x$ over the bending magnets, B_0 is the magnetic field in kGauss, and we have assumed that $J_y = 1$. Finally, $T_{\rm train}$ is the train length which is determined by the length of the bunch train and the injection and extraction kickers.

Equations 4.8 and 4.9 show that the emittance of a ring can be decreased by reducing the dispersion in the bending magnets, reducing the strength of the bends, or decreasing the energy of the ring. Unfortunately, the damping times are increased by reducing the bending magnet strength or decreasing the energy of the ring. This implies that the dispersion in the bends is the only free parameter. Unfortunately, it is also constrained.

In the NLC design, we would like to limit the ring energy to roughly 2 GeV; reasons for this choice are discussed in Section 4.2.6. If we then apply these formula to the main damping rings, we find that we need a 30-kGauss bending field to attain the desired damping. This is not reasonable and therefore we consider additional methods of increasing the damping.

The most obvious method of increasing the damping is to use damping wigglers. Another method of increasing the damping is to change the horizontal damping partition. Unfortunately, this has no effect on the vertical damping if the ring is not coupled. Regardless, changing the horizontal damping partition will decrease the equilibrium horizontal emittance which is another quantity we need to reduce. The damping partition can be changed with a defocusing gradient in the bending magnets—Robinson wigglers are not effective because of the high dispersion that is needed [Raubenheimer 1988]. Unfortunately, damping wigglers reduce the change of J_x. Thus, to change J_x significantly, one is forced to use very high gradients in the bending magnets which makes the option less desirable.

The scaling formulas Eqs. 4.8 and 4.9 can be modified to include the effects of damping wigglers:

$$\tau_{y\,\rm eff} = \frac{2.88\times10^{12}T_{\rm train}}{B_0(kG)\gamma^2(1+F_w)} \tag{4.10}$$

$$\gamma\epsilon_{x0} = \frac{65T_{\rm train}}{\tau_{y\,\rm eff}(J_{x\,0}+F_w)}\frac{\left(\langle\mathcal{H}\rangle_B + 6\overline{\beta_x}F_w\rho_0/5\pi k_w^2\rho_w^3\right)}{\left(1+F_w\right)} \tag{4.11}$$

where $J_{x\,0}$ is the damping partition without wigglers and the parameter F_w is a measure of the effectiveness of the wigglers,

$$F_w \equiv \frac{L_w}{\pi}\frac{B_w^2}{B_0}\frac{1}{B\rho} = \frac{L_w}{\pi}\frac{\rho_0}{\rho_w^2} \quad . \tag{4.12}$$

In addition, L_w and B_w are the length and peak field of the wiggler; $\overline{\beta_x}$ is the average beta function in the wiggler, and k_w is the wiggler wave number, $k_w = 2\pi/\lambda_w$ where λ_w is the wiggler period. Also, ρ_0 and ρ_w are the bending radii of the main bends and the wiggler. Both bending radii are proportional to the energy over the respective magnetic fields. We have assumed sinusoidal wigglers, and thus the average of $B_w^2(s)$ over the wiggler is equal to $\frac{1}{2}\hat{B}_w^2$. In addition, we assume that $\overline{\beta_x} \gg \lambda_w/2\pi$ so that the $\eta_x'^2\beta_x$ term dominates in the excitation term from the wiggler.

As described in Section 4.2.7, in the NLC main damping ring design we have decided to equalize the contributions from the wigglers and the arcs to the radiation damping. This specifies that the term F_w is equal to 1. At this point, we can discuss the Halbach scaling formula for hybrid wigglers [Halbach 1985]:

$$B_w \le 3.44\exp\left[-\frac{g}{\lambda_w}\left(5.08 - 1.54\frac{g}{\lambda_w}\right)\right] \quad , \tag{4.13}$$

which is roughly valid for $0.08 < g/\lambda_w < 0.7$. This suggests that one can achieve a peak field of $B_w = 22$ kGauss (close to the saturation point of Vanadium Permendur) in a wiggler with a gap of 2 cm and a period of 20 cm. In

practice, this scaling result is optimistic when close to saturation but it provides a starting point from which one can determine the wiggler parameters.

4.2.6 Ring Energy

There are a number of considerations when determining the ring energy. First, the desire to maintain high-spin polarization while damping the beams suggests that the nominal ring energy should be chosen so that the spin tune is a half integer. This limits the energy to

$$E = (n + 0.5)440\,\text{MeV} \;= 1.10, 1.54, 1.98, 2.42, \cdots \text{GeV} \quad . \tag{4.14}$$

Second, we would prefer to keep the ring energy low for two primary reasons: it keeps the ring cost lower and it keeps the normalized longitudinal emittance small, making the bunch compression easier. The NLC requires that the damping ring bunch be compressed longitudinally by roughly a factor of 40. We do not want an uncorrelated energy spread much greater than 1% in the linac and thus we need to perform at least a portion of the bunch compression at an energy five times that of the damping ring. Unfortunately, at higher energies it becomes more difficult to perform the compression without degrading the beam emittances.

The problem with a low-energy ring is that it is harder to attain the required damping using reasonable magnet designs. We can understand the choices using the scaling formula. First, using Eqs. 4.10 and 4.12 and assuming that the fraction of damping due to the wigglers F_w and the wiggler peak field B_w are fixed, we find that, to meet the damping time requirements, the main bending field and the required length of wiggler scale as

$$B_0 \sim \frac{1}{\gamma^2} \qquad \text{and} \qquad L_w \sim \frac{1}{\gamma} \quad . \tag{4.15}$$

Thus, as the ring energy is increased, the bending magnets and wigglers become easier to design and construct.

Next, using Eq. 4.11 and noting that the dispersion invariant scales as $\langle \mathcal{H} \rangle_B \sim \Theta^3 \rho_0 \sim \Theta^3 \gamma / B_0$, we find that, to maintain the equilibrium emittance,

$$N_{\text{cell}} \sim \gamma \quad \text{and} \quad L_{\text{bend}} \sim \gamma^2 \quad . \tag{4.16}$$

Finally, given these relations, the length of the ring, the momentum compaction, and the synchrotron radiation power must scale as

$$L_{\text{ring}} \sim \gamma^3 \quad \text{and} \quad \alpha_p \sim 1/\gamma^2 \quad \text{and} \quad P_{SR} \sim U_0 \sim \gamma \quad , \tag{4.17}$$

where we assumed that the cell length scales in the same manner as the length of the bends. The cost of the rings will tend to increase with the length, while the cost of the rf systems will increase with the power required; both of these costs will increase with higher energy. In addition, the momentum compaction decreases with the square of the ring energy while rough scaling for the longitudinal microwave threshold scales as $\gamma \alpha_p$. This suggests that longitudinal stability may be more difficult at higher energy.

Thus, the determination of the nominal ring energy becomes an optimization between the difficulty of the bending and wiggler magnets versus the decrease in momentum compaction, and the increase in size and cost of the ring. We should attempt a cost optimization, but for this iteration of the NLC damping ring design, we have picked an energy of 1.98 GeV and a spin tune of roughly 4.5. This is the lowest energy that appears to yield reasonable designs for the bending magnets and wigglers. Of course, during operation, we may want to vary the ring energy to optimize for a lower repetition rate or a larger injected emittance. For this reason, we are designing the rings to operate over the 1.8–2.2-GeV range; this range extends between the two nearby integer spin resonances.

Finally, it is worth noting that intrabeam scattering does not present justification for increasing the ring energy. *When the effective damping times are held constant*, the emittance growth due to intrabeam scattering does not decrease with the ring energy [Raubenheimer 1991]. Using very simple scaling formula, the emittance growth depends upon the intrabeam scattering growth rate compared to the synchrotron radiation damping. The intrabeam scattering growth rate scales roughly as

$$\tau_{x\,\mathrm{IBS}}^{-1} \sim \frac{N\langle\mathcal{H}\rangle}{\gamma\epsilon_x\gamma\epsilon_y\gamma\epsilon_z} \tag{4.18}$$

which, for given normalized emittances, is independent of energy, while the damping rates scale inversely with the ring length which increases as γ^3. Thus, emittance growth due to intrabeam scattering actually becomes worse as the design energy is increased.

4.2.7 Wigglers vs. Arcs

At this point, we can discuss the difference between designing a wiggler with arcs or a ring with wigglers. In the former, the radiation damping in the wiggler dominates the radiation damping in the arcs, *i.e.*, the parameter F_w is much greater than 1. In this case, the contribution of the quantum excitation in the arcs is suppressed by roughly a factor of $(1 + F_w)^2$. This allows one to design the arcs with a much larger dispersion which leads to a larger contribution to the momentum compaction and makes the chromatic correction easier. Unfortunately, the gains in the momentum compaction and chromatic correction are partially offset by the additional length of wiggler that is required.

The Japan Linear Collider (JLC) damping rings [JLC] are an example of this design style. There, roughly 80% of the damping is performed in the wigglers and the requirements on the arcs are substantially reduced. But, the JLC design requires 140 m of high-field wiggler and the momentum compaction and dynamic aperture, if larger at all, do not appear to be significantly larger.

For the NLC main damping ring design, we have decided to pursue another approach: we want to divide the radiation damping equally between the arcs and the damping wiggler. We feel that this is a more efficient use of magnets and space. It exchanges some of the difficulties of dealing with a very long wiggler system with the arcs. Preliminary comparisons between the JLC and NLC rings show that they both have similar performance expectations.

4.2.8 Lattice Specification

At this point, we can apply the scaling formula to estimate the parameters for the NLC damping rings. First, we can estimate the bending field required in the pre-damping ring to meet the damping time requirements. The pre-damping ring needs to attain an effective damping time of $\tau_{y\,\mathrm{eff}} = 1.74\,\mathrm{ms}$ (Eq. 4.7). Assuming that damping wigglers are not used, limiting the change to the horizontal damping partition to $J_x = 1.3$, and allowing 60 ns for the rise and fall times of the injection and extraction kickers, Eq. 4.10 shows that the bending field must be $B_0 = 17.5\,\mathrm{kGauss}$; this is comfortably below the saturation point of standard magnet iron. Two points to note: first, the required field is independent of the number of trains stored in the ring at once, and second, without changing J_x, the required field would be 20 kGauss which is quite high.

A similar procedure can be used to estimate the bending fields and wigglers fields required in the main damping rings. Assuming that wigglers are used to perform half the required damping and that the bunch trains are separated by 60 ns, Eq. 4.10 shows that the bending field must be $B_0 = 15.1\,\mathrm{kGauss}$. If we limit the peak wiggler field to 22 kGauss, Eq. 4.12 shows that we need 26 m of wiggler.

Now, we have to specify a lattice that will meet the emittance requirements. For the main damping ring, we have considered four different lattice cell structures: a FOBO cell, which is a FODO cell where the defocusing quadrupole is replaced by a bending magnet with a defocusing gradient, a Chasman-Green lattice, a Triple Bend Achromat (TBA) lattice, and a Theoretical Minimum Emittance (TME) lattice. The advantages and disadvantages of each are summarized below:

- FOBO
  - \+ Simple FODO-like structure with well-constrained beta-functions
  - – Uses three quadrupoles per bending magnet (assuming symmetric form)
  - – Less efficient at generating emittance
  - – Difficult combined function bending magnet
- Chasman-Green
  - \+ Fairly efficient at generating emittance
  - \+ Two quadrupoles per bending magnet
  - \+ Excellent dynamic aperture behavior in R. Brinkmann's comparison [Brinkmann 1990].
  - \+ Presently used at ESRF and ELETTRA
- TBA
  - \+ Fairly efficient at generating emittance
  - \+ Two quadrupoles per bending magnet
  - \+ Presently utilized at ALS and PLS
- TME
  - \+ Very efficient at generating emittance (fewer bends)
  - \+ Easy to optimize ratio of α_p and $\mathcal{H}$
  - \+ Excellent dynamic aperture in studies
  - – Three quadrupoles per bending magnet

To study the different structures, we made preliminary designs for the NLC main damping rings using each of these four lattices [Moshammer 1993]; the rings were designed in a race-track form with two arcs, including dispersion suppressors, and roughly 10 m of straight section which could be expanded into injection and extraction and wiggler sections. Parameters and tracking results for the four designs are listed in Table 4-6 and the cell lattice functions are plotted in Figures 4-3, 4-4, 4-5, and 4-6. Notice that the dispersion is not matched to zero at the ends of the Chasman-Green and TBA cells; zero dispersion is not necessary after most cells in the ring and one can obtain smaller emittances by matching to nonzero values. The only advantage these preliminary results indicated was that for equal ring energy, the TME design used fewer and longer bending magnets with lower field gradients. This was thought to be sufficiently important that it outweighed the advantage of using a more standard lattice such as the Chasman-Green or TBA. Thus, we decided to complete a design based upon the TME structure and study the other lattices further in the future.

The optimization is different in the positron pre-damping ring. Here, the equilibrium emittance is easily attained and the primary issues become attaining the desired dynamic aperture and constraining the beta-functions to keep the

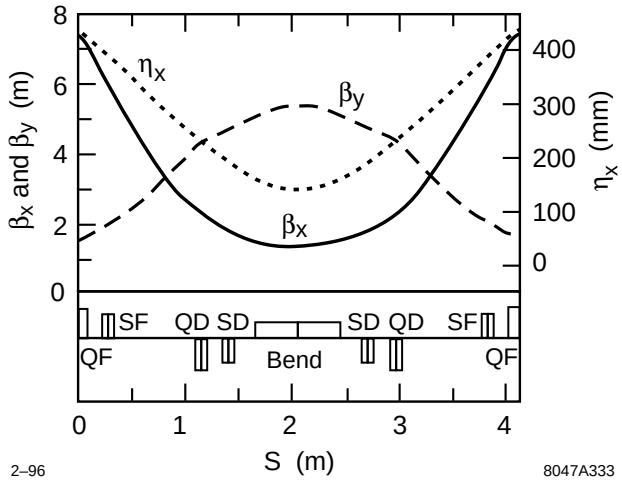


Figure 4-3. *Lattice functions in a FOBO cell.*

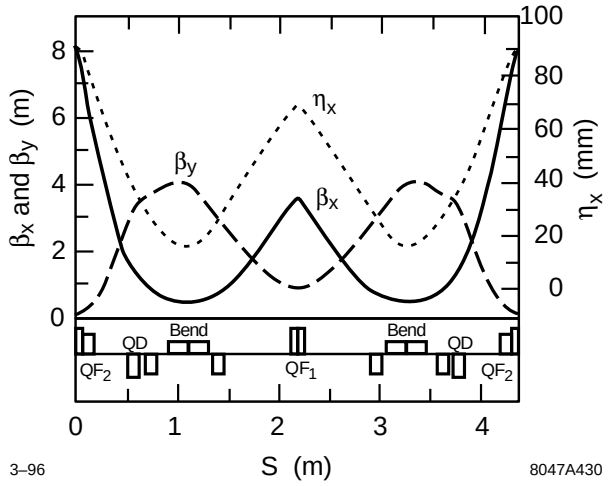


Figure 4-4. *Lattice functions in a Chasman-Green cell.*

Lattice (GeV)	Energy $(10^{-6}$ m-rad)	$\gamma\epsilon_{x\,\mathrm{eq}}$ (10^{-3})	α_p (m)	Circ.	N_{bends} $(10^{-3}$ m-rad)	Norm. Accept. X/Y
FOBO	2.2	4.4	1.4	133	80	36 / 17
Chasman-Green	2.0	3.0	0.6	154	72	4 / 14
TBA	2.0	3.4	0.7	177	72	14 / 10
TME	2.0	3.2	0.7	158	40	36 / 26

Table 4-6. *Parameters for different lattice structures [Moshammer, 1993].*

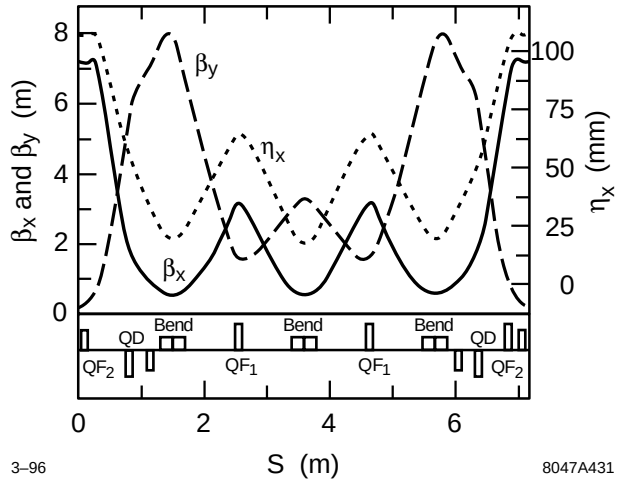


Figure 4-5. *Lattice functions in a TBA cell.*

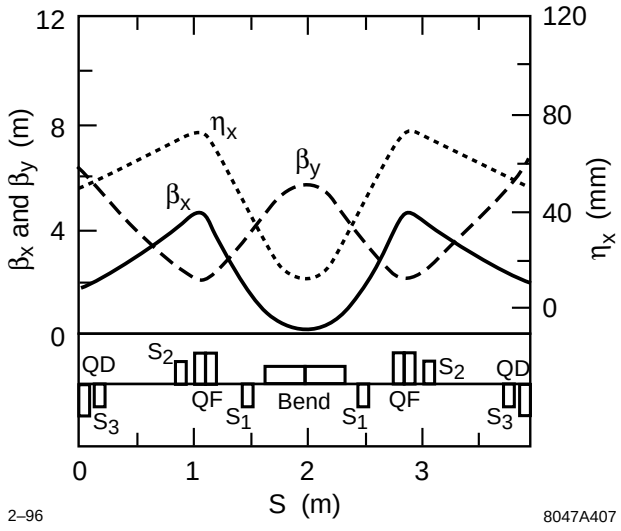


Figure 4-6. *Lattice functions in a TME cell.*

required vacuum aperture relatively small. The FOBO lattice which is basically a simple FODO structure seems to be the best choice in this case; typical lattice functions are plotted in Figure 4-3. The FOBO structure was felt to be superior to the more standard FODO structure because it allows one to increase the horizontal damping partition and thus increase the transverse damping rates when the ring operates on the betatron coupling resonance.

4.2.9 RF Parameters

In both the pre-damping ring and the main damping rings, the rf is chosen to be 714 MHz. This is as low as consistent with the bunch spacing of 1.4 ns. There are several advantages of the low-frequency choice. First, for a given rf voltage, the lower frequency has a higher stored energy in the cavities thereby reducing the beam loading and the bunch-to-bunch synchronous phase shift. Second, the coupled-bunch instability driven by the fundamental mode of the cavities has a smaller growth rate. Finally, two other advantages of the lower frequency is the increased bunch length in the main damping rings and the larger energy acceptance in the pre-damping ring.

The minimum rf voltage is determined by the synchrotron radiation, HOM losses, and the height of the energy bucket. In the pre-damping ring, we would need an rf voltage in excess of 3.5 MV to capture the full $\pm 2\%$ incoming energy spread. Instead, we use an energy compressor to reduce the injected energy spread at the expense of increasing the bunch length. We expect the compressor will be able to reduce the incoming energy spread by a factor of two. With the compressor, an rf voltage of 2.0 MV should be sufficient; this provides an energy bucket of $\pm 1.5\%$.

In the main damping rings, we are designing for an rf voltage of 1.0 MV. This is the lowest voltage that is consistent with the energy loss per turn of roughly 700 kV. The main rings have a very small momentum compaction and we want to keep the voltage low to keep the bunch length relatively long. Space is being left in the design to allow for future upgrades of the rf system to 1.5 or 2.0 MV.

Finally, the harmonic number of the rings is chosen to allow operation with fewer, but possibly longer, bunch trains at a reduced repetition rate. Since one would like to place the trains symmetrically in the ring, this requires that the harmonic number be a multiple of all the different numbers of trains being considered. Thus, while the PPDR has been designed to damp three trains of bunches at the same time, we have chosen an even harmonic number so that it could also operate with two evenly spaced bunch trains. Similarly, the EDR and PDR have been designed to operate with four trains of bunches at a 180-Hz repetition rate. To allow for operation with three bunch trains at 120 Hz, the harmonic number should be a multiple of 4 and 3. Unfortunately, the present design has a harmonic number of 532 which is not ideal. In the next design iteration, the rings will be modified to have an harmonic number of 540.

Finally, it should be noted that the harmonic numbers of the rings do not need to be related from timing considerations because only single bunch trains are injected on any machine cycle. In particular, the PPDR has a slightly larger gap between bunch trains allow more time for the injection and extraction kickers to rise and fall.

4.3 Lattice Design

In this section, we will describe the layouts and optics of the main damping rings, the pre-damping ring, the injection and extraction transport lines, and the PPDR to PDR transport line. After describing the optical designs, we will discuss the lattice related accelerator physics issues. This includes trajectory correction, emittance control and matching, dynamic and energy apertures, and polarization. These are all low-current effects. Collective effects are discussed in Section 4.4 and the rf and vacuum systems are described in Sections 4.5 and 4.6. In addition, the feedback systems

and support and alignment systems are described in Sections 4.7 and 4.9 and the magnet and diagnostics designs are describe in Sections 4.10 and 4.11.

At the time of this writing, the design of the main damping ring is fairly complete. We do not feel that there are any significant flaws in the design although a lot of detailed engineering is still required; this is particularly true of the rf cavities and the bending magnets. The only significant modification currently planned is increasing the ring circumference to increase the harmonic number from 532 to 540; this will facilitate operation with either 4 or 3 bunch trains. The circumference change will be made by increasing the length of the two straight sections by roughly 1.7 meters which will have the side benefit of making the vacuum chamber design simpler in the wiggler region.

The design of the pre-damping ring has not yet been investigated in the same depth as the main damping ring design. At this time, we have a design that appears to meet all of the requirements but many of the component designs have not yet been considered. As stated in the introduction, our belief is that the pre-damping ring is significantly more straightforward than the main damping rings except for two issues: the injection and extraction systems and the dynamic aperture. Both of these arise from the very large injected beam emittances and both of these are addressed in our design.

4.3.1 Main Damping Rings

As described, the Electron and Positron Damping Rings (EDR and PDR) are designed to damp beams with an injected emittance of $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad to an emittance of $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad. The damping and emittance requirements on the rings are described in Section 4.2.3. The determination of the design parameters is discussed in Sections 4.2.6–4.2.9; these parameters are summarized in Table 4-7.

The lattice is based on detuned Theoretical Minimum Emittance (TME) cells which were chosen because of eased requirements on the bending magnets. The chromaticity is corrected with three families of sextupoles and the dynamic aperture is more than sufficient. The damping is performed using both high-field bending magnets and a damping wiggler. The bunch trains are injected onto and extracted from the closed orbit using pulsed kickers and DC septa. Finally, all of the quadrupoles and all of the sextupoles will have independent power supplies. This facilitates beam-based alignment as well as matching of the lattice functions; both of these are important for attaining the desired vertical emittance.

In the next sections, we will describe the ring layout and then details of the ring lattice: the arcs and the dispersion suppressors, the wiggler and injection and extraction straight sections, and the placement of the rf cavities. In addition, the rf and vacuum systems are described in Sections 4.5 and 4.6 while the feedback systems and support and alignment systems are described in Sections 4.7 and 4.9. Finally, the magnet and diagnostics designs are describe in Sections 4.10 and 4.11.

Layout

The NLC main damping rings are roughly 223 m in circumference and measure 50 m by 80 m. They have a nominal energy of 1.98 GeV and required damping is achieved using both high field bending magnets and roughly 26 m of damping wiggler. The rings are designed in a race-track form with two arcs separated by straight sections; the layout is illustrated in Figure 4-7. Each arc consists of 19 TME cells plus dispersion matching sections. The straight sections are roughly 30 m in length. On one side of the ring, the straight section is devoted to the damping wigglers while the other straight section contains the injection and extraction components and the rf cavities. The arc cells, which are roughly 4 m in length, have one bending magnet, six sextupole magnets, three quadrupoles, and their own integral vacuum chamber. The wiggler straight section consists of 10 separate wiggler sections. The injection and extraction

Energy	$1.8 \sim 2.2$ GeV; 1.98 GeV nominal
Circ.	223 m
Max. Current	1.2 A
Max. N_{bunch}	1.57×10^{10}
Trains	4 trains of 90 bunches
Train Separation	60 ns
Bunch Separation	1.4 ns
Max. Rep. Rate	180 Hz
ν_x, ν_y, ν_s	23.81, 8.62, 0.004
$\gamma\epsilon_x$	2.56×10^{-6} m-rad
$\gamma\epsilon_x$, $\gamma\epsilon_y$ with IBS	3.1×10^{-6} m-rad, 1.5×10^{-8} m-rad
σ_ϵ	0.09%
σ_z	4.0 mm
$\xi_{x\ \mathrm{uncorr}}, \xi_{y\ \mathrm{uncorr}}$	-41.8, -22.4
τ_x, τ_y, τ_z	4.06 ms, 4.62 ms, 2.50 ms
U_{SR}	644 KV/turn
J_x	1.14
α_p	0.00047
V_{RF}	1.0 MV
f_{RF}	714 MHz
Lattice	40 TME Cells

Table 4-7. *Parameters for main damping ring (vers. 6.11)*

straight section contains: two kicker magnets and two septa for injection and extraction, two rf cavities, feedback striplines, and thirteen quadrupoles. Other components such as polarization monitors, optical monitors, and wire scanners will be added to this section.

At this time, we have not investigated the civil construction of the damping ring enclosures. Ideally, to minimize the disturbance to the ground, the ring enclosure will be constructed as a tunnel and not a vault. Although only schematically illustrated in Figure 4-7, the design goal will be to place all electronics, power supplies, kicker thyratrons, and rf klystrons in an accessible shielded region at the center of the ring. This will minimize cable and waveguide distances while allowing for maintenance and repair during operation.

Arcs

Each arc consists of 19 TME cells plus dispersion matching sections. Optimized TME cells are designed to attain the minimum emittance that is possible in a bending magnet. To do this, three quadrupoles are placed between the bending magnets and are used to minimize the average dispersion invariant $\mathcal{H}$ through the bending magnets. In general, this minimum is found when the dispersion and horizontal beta functions have minima at the center of the dipoles. Explicitly, the minimum emittance should be found when [Teng 1985]

$$\beta_{x\,0}^{TME} \approx \frac{L_B}{\sqrt{60}} \qquad \eta_{x\,0}^{TME} \approx \frac{L_B^2}{24\rho} \tag{4.19}$$

where β_0 and η_0 are the values at the center of the dipole and L_B is the length of the dipole.

Unfortunately, minimizing $\langle\mathcal{H}\rangle$ also implies a small value of the momentum compaction. We would like to keep the momentum compaction large since this eases the requirements on the longitudinal impedance due to the longitudinal

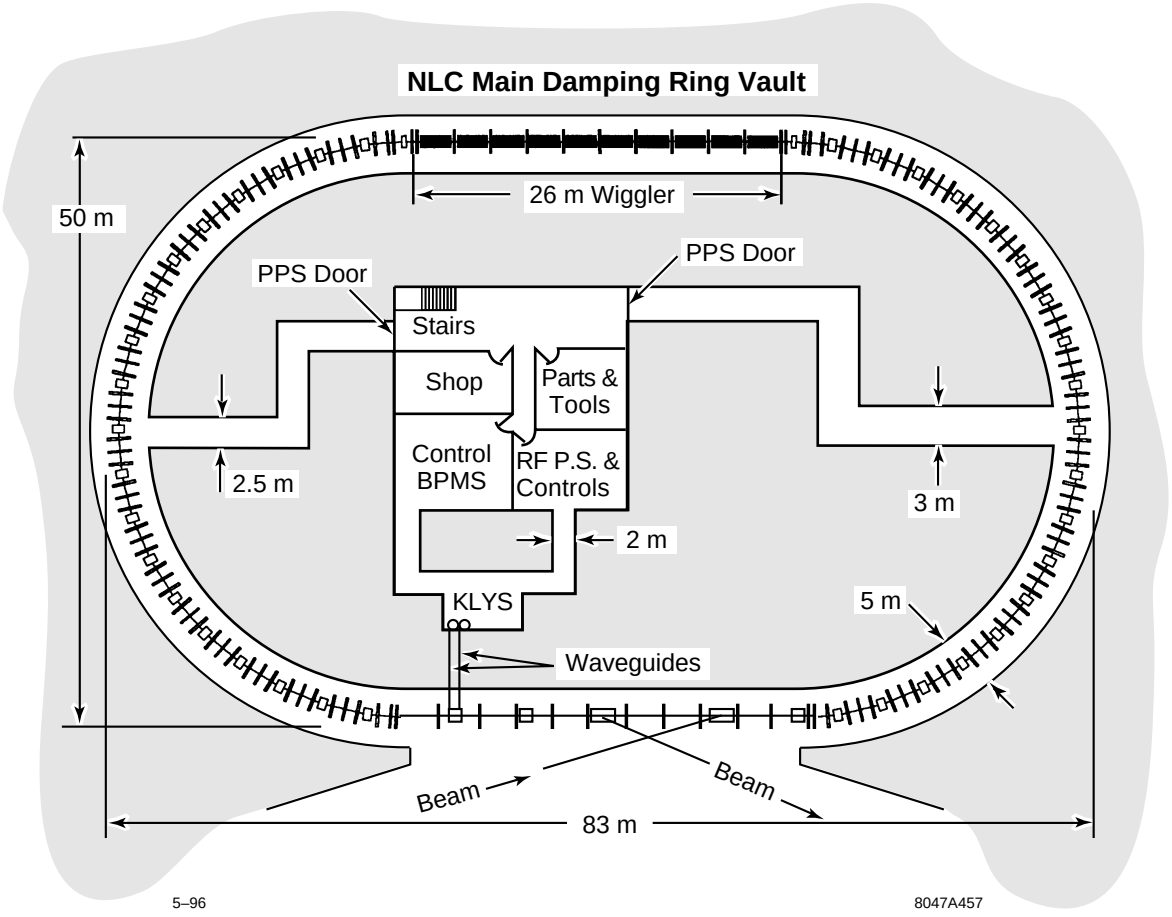


Figure 4-7. *Layout of main damping ring.*

microwave instability and it leads to a longer bunch length which is also desirable. Thus, we chose parameters in the cells to optimize the ratio of I_1/I_5 for a given bending angle rather than $\mathcal{H}$ (or I_5). In this case, the optimal values of the dispersion at the center of the bending magnet is equal to

$$\eta_0^\star \approx -\frac{L_B^2}{24\rho} + \frac{L_B}{\sqrt{15}}\sqrt{2L_B^2 + 5\beta_0^2} \quad . \tag{4.20}$$

This value is actually close the the optimal TME value.

In Figure 4-8, we plot the variation of I_5 (solid), which is proportional to $\mathcal{H}$, and I_5/I_1 (dashes), which is proportional to the ratio of $\mathcal{H}$ to α_p, versus β_0 for the case when $\eta_0 = \eta_0^\star$. Notice that both curves have minima near the optimum beta function but do not increase very rapidly as β_0 is increased. In practice, we detune the cell so that $\beta_0 \approx 2\beta_0^{TME}$; this reduces the chromaticity. In Figure 4-9, we plot the variation of I_5 (solid) and I_5/I_1 (dashes) against η_0 for $\beta_0 = 2\beta_0^{TME}$. In this case, the optimum $\mathcal{H}$ occurs at roughly η_0^{TME} while the optimum $\mathcal{H}/\alpha_p$ occurs around $\eta_0 \approx 2\eta_0^{TME}$. We have chosen to optimize the cells to this latter value.

Each TME cell consists of two QF quadrupoles, a single QD quadrupole, and a single bending magnet with a defocusing gradient. The defocusing gradient in the bending magnets increases both the quantum excitation and J_x. In a ring without damping wigglers, the increase in J_x is larger than the increase in the excitation and the equilibrium

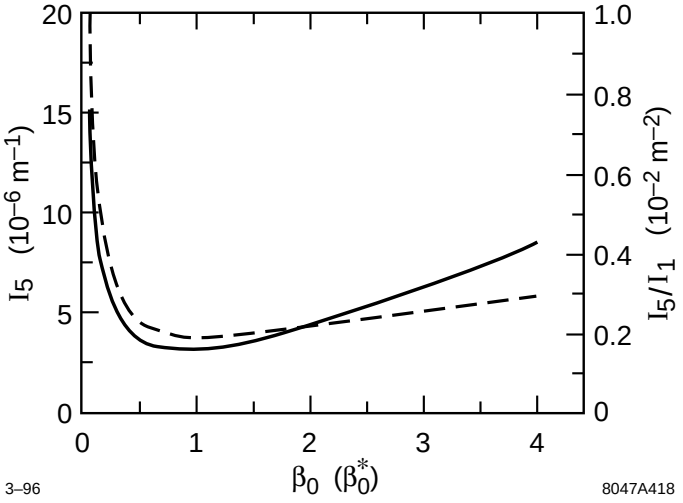


Figure 4-8. *I_5 (solid) and I_5/I_1 (dashes) versus β_0 for TME cells.*

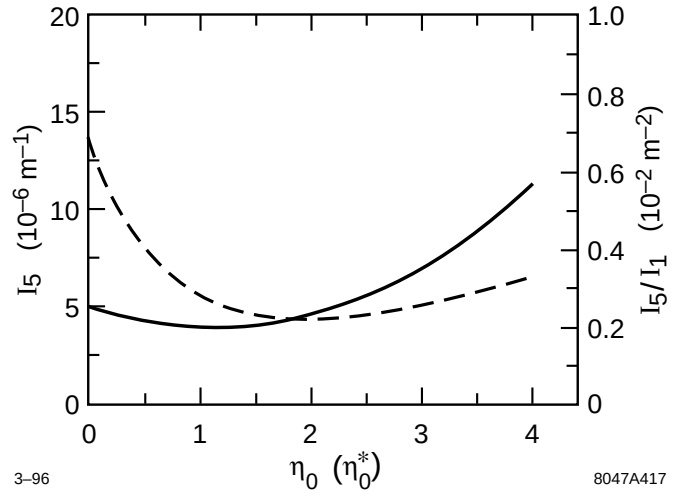


Figure 4-9. *I_5 (solid) and I_5/I_1 (dashes) versus η_0 for TME cells.*

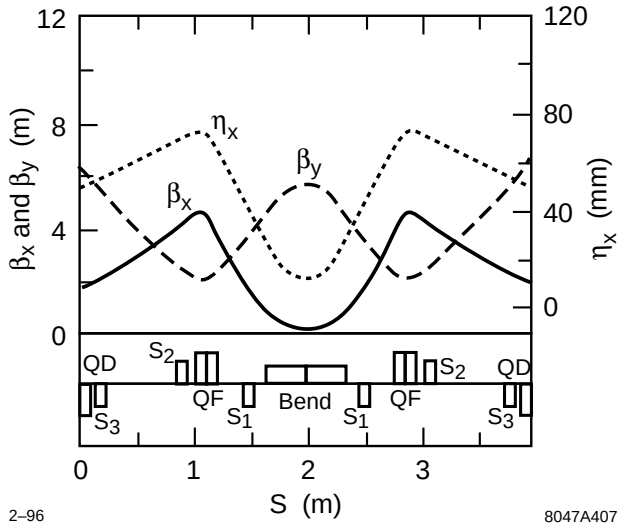


Figure 4-10. *β_x (solid), β_y (dashes), and η_x (dots) in main damping ring cell.*

emittance can be decreased by as much as a factor of two. But, the emittance and damping time reduction is less when damping wigglers are included since they reduce the change to J_x for a given gradient. Because the main damping rings do not need the additional horizontal damping and because it takes large gradients to achieve a significant reduction in the emittance, we have chosen a gradient that optimizes the matching through the cells while producing $J_x = 1.14$.

The length of each cell is roughly 3.9 m. The vacuum chamber is circular with a 1.25-cm inner radius and an antechamber is needed to handle the intense synchrotron radiation. The lattice functions for a cell are plotted in Figure 4-10 and the component layout is illustrated in Figure 4-11. The magnets and vacuum system are described in Sections 4.10 and 4.6.1, respectively.

The dipole steering correction is performed using trim windings on the sextupoles. There are two horizontal correctors and one vertical corrector in each cell. In addition, there are three button BPMs per cell. These are located at the sextupole magnets on either side of the bending magnet and one of the sextupoles adjacent to the defocusing quadrupole.

The chromatic correction is performed using three sextupole families in the arcs as illustrated in Figure 4-10. There is one defocusing family placed on either side of the bending magnets and a focusing and defocusing sextupole placed adjacent to the focusing and defocusing quadrupoles. Thus, there are six sextupoles per cell. At this time, it is not believed that additional sextupoles located in nondispersive regions or additional families of sextupoles are needed. The three sextupole families allow us to correct the chromaticities while also minimizing the tune shift with amplitude terms. This yields a dynamic aperture that is more than sufficient (Section 4.3.9).

Finally, the dispersion is matched to zero at the end of the arcs for the two straight sections. The matching is performed using the three upstream quadrupoles and a bending magnet with an integrated strength that is half of the arc bending magnets. We have assumed that the magnetic field of the matching bends is decreased to reduce excessive fringing fields. The gradient is scaled at the same time so that the matching bends can be constructed using the same pole shape as the main ring bends.

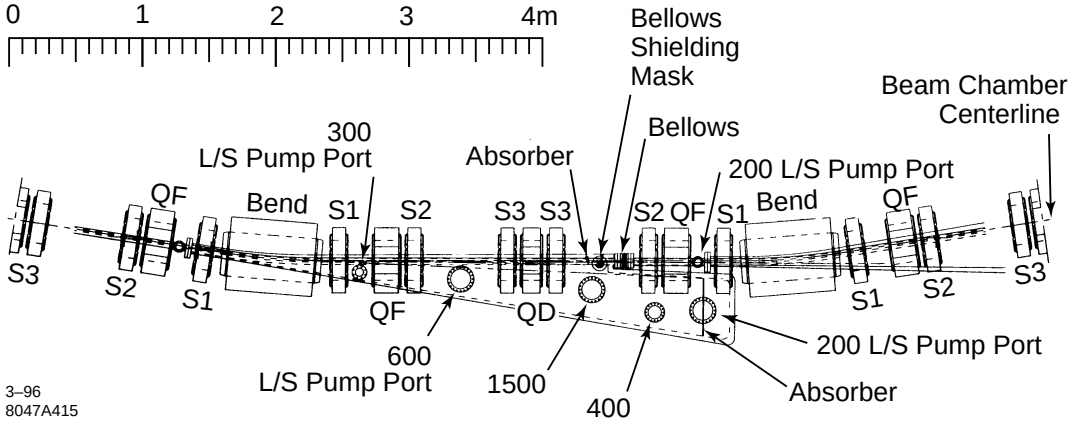


Figure 4-11. *Cell layout in a main damping ring cell showing magnet and vacuum chamber components; note the bending magnets are drawn as "H" magnets although the present design utilizes "C" bending magnets.*

The horizontal and vertical dispersion matching will be tuned using the matching quadrupoles, which have individual power supplies, and trim windings on the adjacent sextupoles that will be wired as skew quadrupoles. Four additional skew quadrupole windings are located in each arc for local correction of the vertical dispersion which is the most important contribution to the vertical emittance; this is discussed further in Section 4.3.8.

Wiggler Region

The wiggler region consists of five FODO cells with 3.0-m drifts between the quadrupoles. There are ten 2.6-m wiggler sections to be placed within the region with roughly 21 cm of space between the wiggler ends and the quadrupoles. The wigglers have a 2-cm full gap. To handle the intense synchrotron radiation, the vacuum chamber will have integrated pumping in ante-chambers on both sides of beam chamber. The adjacent quadrupoles have a 3-cm pole-tip radius so that they fit around the chamber. The wiggler magnet sections are described in Section 4.10.1: "Wiggler Magnets".

To minimize the emittance growth in the wigglers, we need to minimize the average horizontal beta function through the region. The average beta function is minimized at a phase advance of roughly 101.5° per cell, but this is a flat minimum and the chromaticity per cell increases rapidly as the phase advance increases. Thus, we have chosen horizontal and vertical phase advances of roughly 90° and 60° per cell.

The lattice functions are matched from the dispersion suppressors with two quadrupoles immediately following the suppressor and the first two quadrupoles at either end of the wiggler region. This allows the region to remain matched if the wiggler strength is varied. The lattice functions from the last TME cell through the wiggler region are plotted in Figure 4-12.

The dipole steering correction will be performed using trim windings on the quadrupoles wired as dipoles and button BPMs located next to the quadrupoles. The BPMs will be shielded to protect them from the large angle (low energy) synchrotron radiation.

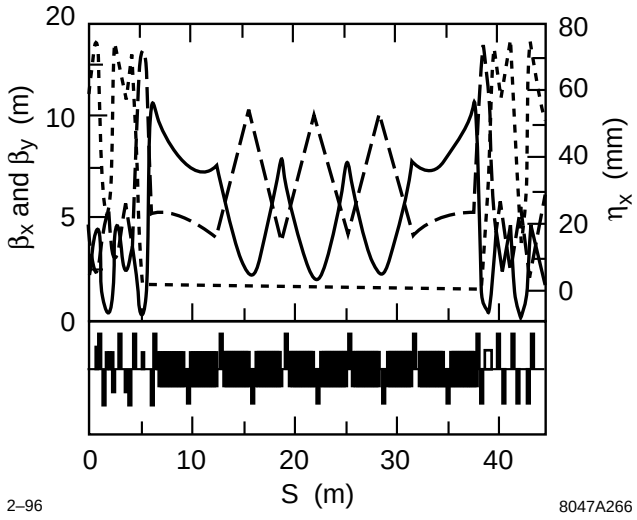


Figure 4-12. *β_x (solid), β_y (dashes), and η_x (dots) in main damping ring wiggler region.*

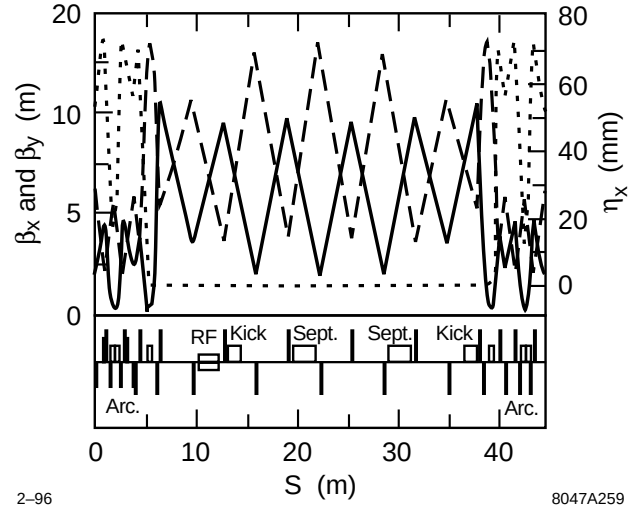


Figure 4-13. *β_x (solid), β_y (dashes), and η_x (dots) in main damping ring injection and extraction region.*

Injection-Extraction Region

The beams are injected and extracted in the injection-extraction straight section. To reduce the rf transients, beams are injected and extracted simultaneously. In this way, the beam current through the rf cavities is not interrupted.

The injection-extraction region is simply a mirror image of the wiggler region. It consists of five FODO cells with 3.0-m drifts between quadrupoles. The matching of the lattice functions is performed with four matching quadrupoles at either end of the region. The placement of the kickers and septa is illustrated in Figures 4-13 and 4-14 and the extracted beam trajectory, along with the position of the septum blade, is illustrated in Figure 4-15.

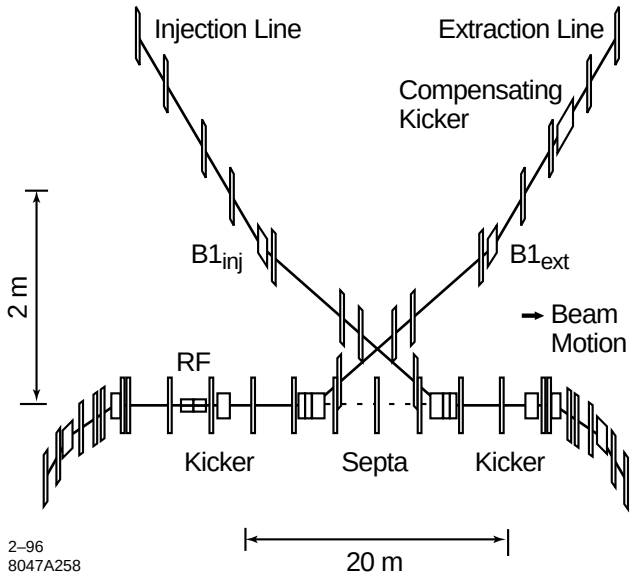


Figure 4-14. *Component layout in the main damping ring injection and extraction region.*

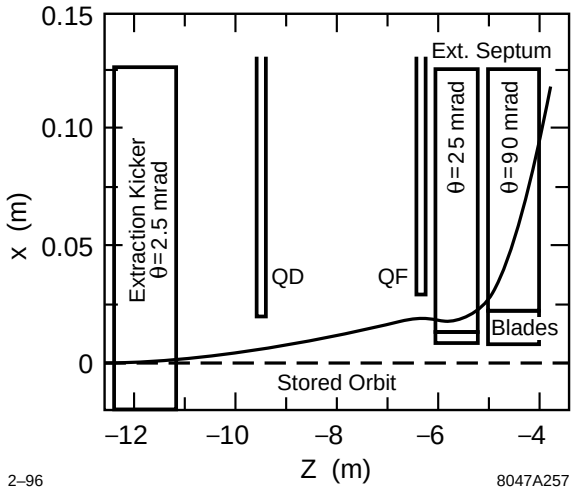


Figure 4-15. *Extracted beam trajectory from kicker through septum.*

The steering correction in this region will be performed using air-core dipole magnets. Experience with the SLC damping rings has illustrated the need for orthogonal horizontal and vertical angle and position closed orbit bumps through each of the septa as well as control of the injected and extracted beam trajectories.

The kickers are 1.2 m in length and provide a 2.5-mr horizontal deflection; these are described further in Section 4.10.3.

The septa are DC current sheet septa constructed in two pieces. The first septum is 83-cm long and provides a 25-mr deflection. The blade has a thickness of 5 mm and lies 9 mm from the stored beam where the residual field is negligible. The injected and extracted beam trajectories have 4.5-mm clearance from the blade; assuming an emittance of 10^{-4} m-rad, the 3-sigma horizontal beam size is 1.5 mm. The second septum is 1-m long and it provides a 90-mr deflection. The blade has a thickness of 15 mm and lies 9 mm from the stored beam where again the residual field is negligible. Both sections of the septum magnet are described further in Section 4.10.1.

The adjacent focusing quadrupoles need to have a 3-cm pole-tip radius to provide room for the injected or extracted beams, while the adjacent defocusing quadrupoles may need to include an entrance or exit beam line; at the adjacent defocusing quadrupoles, the injection and extraction trajectories are roughly 25 cm from the stored beam trajectory.

The injection and extraction lines cross about 60 cm from the stored beam trajectory. Initial calculations of the effect of the crossing suggest that the effect of the beams passing through each other is not insignificant. Thus, a vertical chicane will probably be used to deflect the incoming beam trajectory away from the extracted beam. The closest magnets are 1.2-m away from the crossing and thus there is plenty of longitudinal space for the chicane but attention will need to be given to the effect of the fringing fields on the extracted beam.

To reduce the tolerances on the septum field stability, a compensating bend B_1, which is powered in series with the septum, is located after the beam crossing on both the injection and extraction lines. Four quadrupoles are used to match the R_{22} transport element so that field fluctuations will cancel. This will ease the septum field tolerances by roughly an order of magnitude to $\Delta I/I \sim 5 \times 10^{-5}$.

In the same manner, the effects of kicker jitter on the extracted beam will be reduced using an achromatic (double) kicker scheme. Here, a compensation kicker is located in the extraction line after the septum compensation bend. Two additional quadrupoles are used to match the compensation kicker so that it is separated from the extraction kicker by a $+I$ transform in the horizontal plane. The compensation kicker is powered from the same thyratrons as the extraction kicker and is expected to reduce the sensitivity to kicker jitter by at least a factor of four. This will ease the tolerance $\Delta\theta/\theta$ on the extraction kicker from 0.0005 to 0.002. A similar achromatic kicker system will be tested at the ATF damping ring at KEK.

In addition, the sensitivity of the stored beam to the trailing edge of the kicker pulses is reduced by separating the injection and extraction kickers by $+I$ and adjusting the falling edge of the injection kicker to compensate for that of the extraction kicker; this can be done using a separate pulser as discussed in Section 4.10.3. Furthermore, the bunch trains will be ordered in the ring so that the trailing edge of the kicker pulse only effects the most recently injected bunch train allowing any residual deflection to be damped.

Finally, we should note that this injection and extraction region requires fairly difficult septa and kicker magnets. Although the magnets are thought to be possible, we also considered an alternate design that would require weaker kicker magnets. The alternate design consisted of two FODO cells with 5-m drift sections between quadrupoles. In this case, there is only one septum, a septum quadrupole, which both the injected and extracted beams use. Two of the disadvantages of this design were that the two beams must pass through each other and the resulting deflects of the extracted bunches can be as large as $0.2\sigma_y$ and the field tolerances in the septum quadrupole were very tight. For these reasons, we are not considering this design further.

RF Cavities

The rf cavities are located in a 3-m drift section before the extraction kicker; the cavity placement is illustrated schematically in Figures 4-13. The present rf design specifies two rf cavities, but the rf cavities and associated tapers are roughly 50 cm in length and thus additional cavities could be installed if necessary. In these sections, the dispersion is zero and the horizontal and vertical beta functions are roughly 6 m. It is expected that the cavity aperture will be roughly twice the nominal vacuum chamber radius and thus special masks will not be needed to protect the cavities from beam loss during injection or synchrotron radiation.

Detuned Lattice

We have studied a detuned version of the lattice for the main damping rings. The optics has an equilibrium emittance and a momentum compaction that are both roughly twice as large as the nominal lattice. In addition, the quadrupole and sextupole strengths are reduced by 30% while the dynamic aperture is increased by roughly 50% and the ring tunes are changed to $\nu_x = 20.185$ and $\nu_y = 6.385$. Such a lattice may prove useful during the initial commissioning stages.

4.3.2 Pre-Damping Ring

As described, the Positron Pre-Damping Ring (PPDR) is designed to damp the large emittance positron beam from the positron source to an emittance of roughly $\gamma\epsilon_{x,y} = 1 \times 10^{-4}$ m-rad. At this point, the positrons are injected into the positron main damping ring (PDR), described in Section 4.3.1 where they are damped to the desired final emittances. The pre-damping ring allows us to decouple the large aperture requirements for the incoming positron beams from the final emittance requirements of the linear collider.

The damping and emittance requirements on the PPDR are described in Section 4.2.3. The determination of the design parameters is discussed in Sections 4.2.6–4.2.9; these parameters are summarized in Table 4-8.

The pre-damping ring does not need to produce flat beams. Thus, to maximize the damping of the transverse phase space, the ring has $J_x \approx 1.4$ and operates on the coupling difference resonance. This increases the damping in both the horizontal and vertical planes. Furthermore, like the main damping ring, the ring damps multiple trains of bunches at once, the number of which is determined by the ring circumference. The initial design was only 114 m in circumference and stored only two trains. The present design is roughly 50% larger and stores three bunch trains. Although this increase in circumference will certainly increase the cost of the ring, the previous design was very tightly packed making the component design and maintenance substantially more difficult; experience with the SLC damping rings strongly suggests that the increase in circumference is cost-effective.

The magnets and vacuum system are being designed to provide sufficient aperture to accept a 2-GeV beam with an edge emittance of $\gamma\epsilon_{x,y} = 0.09$ m-rad and $|\delta p/p| \leq 1.5\%$ plus 2-mm clearance for misalignments and mis-steering. Given the nominal injected edge emittance of $\gamma\epsilon_{x,y} = 0.06$ m-rad, this provides a substantial margin for injection and internal mismatches.

Like the main damping rings, all of the quadrupoles and all of the sextupoles will have independent power supplies. This will facilitate beam-based alignment as well as matching of the lattice functions which is especially important in the pre-damping ring because of the limited aperture.

In the next sections, we will describe the ring layout and then details of the ring lattice: the arcs and the dispersion suppressors, the wiggler and injection and extraction straight sections, and the placement of the rf cavities. Issues

Energy	$1.8 \sim 2.2$ GeV; 1.98 GeV nominal
Circ.	171 m
Max. Current	1.2 A
Max. N_{bunch}	1.57×10^{10}
Trains	3 trains of 90 bunches
Train Separation	64 ns
Bunch Separation	1.4 ns
Max. Rep. Rate	180 Hz
ν_x, ν_y, ν_s	10.18, 8.18, 0.018
$\gamma\epsilon_x$	7.7×10^{-5} m-rad
$\gamma\epsilon_{x,y}$ (coupled)	4.5×10^{-5} m-rad
σ_ϵ	0.1%
σ_z	7.5 mm
$\xi_{x\,\mathrm{uncorr}}$, $\xi_{y\,\mathrm{uncorr}}$	-13.9, -10.0
τ_x, τ_y, τ_z	4.44 ms, 6.15 ms, 2.73 ms
U_{SR}	371 KV/turn
J_x	1.39
α_p	0.0051
V_{RF}	2.0 MV
f_{RF}	714 MHz
Lattice	30 FOBO Cells

Table 4-8. *Parameters for pre-damping ring (vers. 3.4)*

such as trajectory correction, dynamic aperture, and matching are described later in this section (Section 4.3), while collective effects are discussed in Section 4.4. In addition, the rf and vacuum systems are described in Sections 4.5 and 4.6 while the feedback systems and support and alignment systems are described in Sections 4.7 and 4.9. Finally, the magnet and diagnostics designs are describe in Sections 4.10 and 4.11.

Layout

The design of the pre-damping ring looks very similar to the SLC damping rings, except the lengths are scaled by a factor of five. The ring has a race-track form with dispersion-free straight sections for injection and extraction. The ring is roughly 60 meters by 40 meters and is illustrated schematically in Figure 4-16.

In the design, the injection and extraction regions are in the straight sections on opposite sides of the ring. To reduce the requirements on these components, the systems are designed to occupy most of these regions. To minimize rf transients during injection and extraction, a new bunch train will be injected one half turn after a train is extracted. Furthermore, the rf cavities are placed downstream of the injection kicker and upstream of the extraction kicker so that the injection/extraction process will not interrupt the beam current seen be the rf cavities.

Arcs

Each arc is constructed from 14 FOBO cells plus dispersion suppressors. The FOBO cells are similar to FODO cells except a bend with a defocusing gradient is used instead of the defocusing quadrupole. This lattice structure was chosen because it keeps the lattice functions well constrained and it is simple. It has significantly better emittance

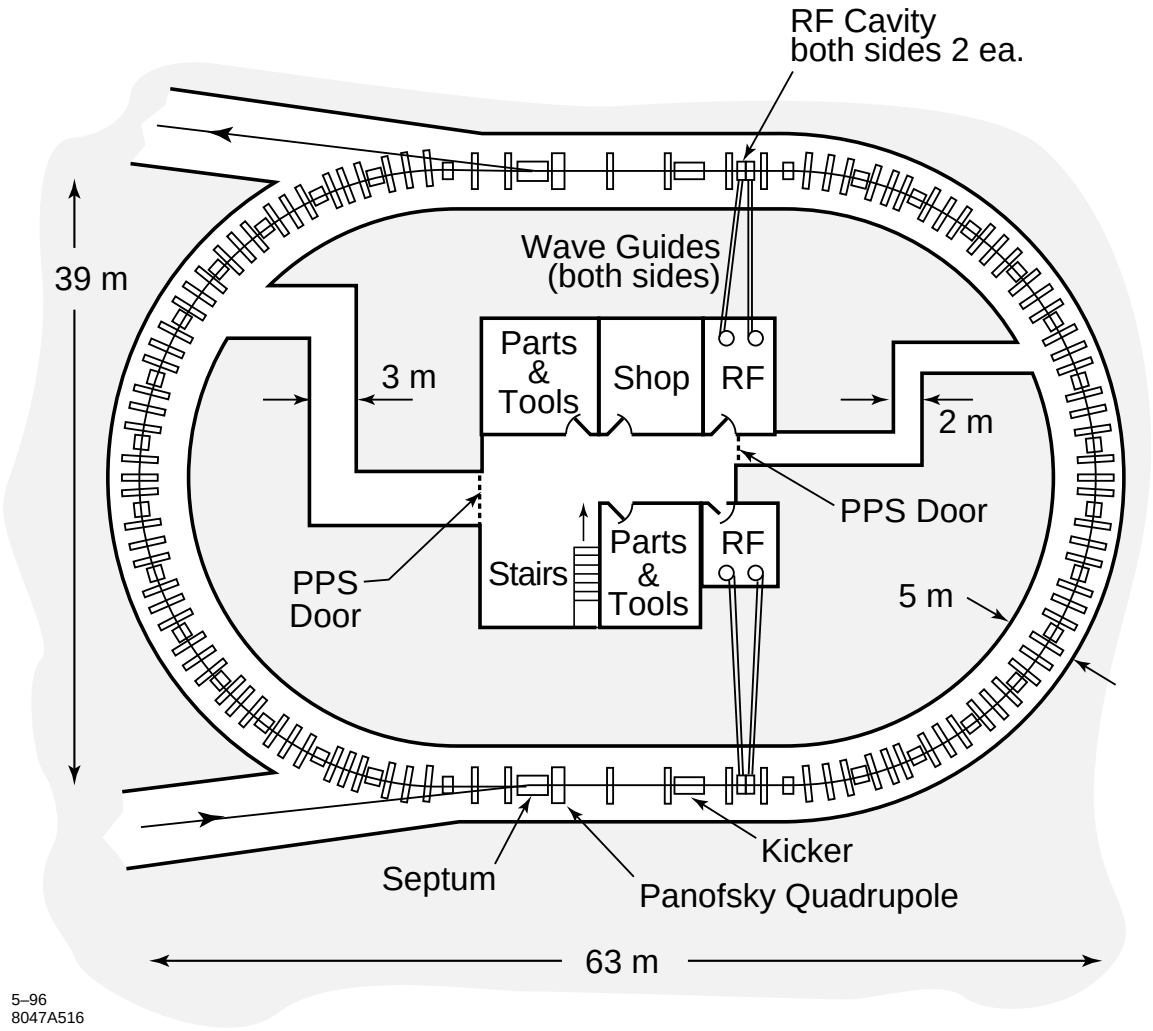


Figure 4-16. *Layout of pre-damping ring.*

performance than the normal FODO lattice and the single central bending magnet has smaller fringing field losses than if it were split as in a normal FODO cell.

Although the bending magnet has a defocusing gradient and could, in theory, completely replace the defocusing quadrupole, small defocusing quadrupoles are included at either end of the bending magnet. These provide additional vertical focusing to reduce the aperture and gradient requirements in the bend and they provide additional tuning control to optimize the lattice. Thus, each cell contains a single bending magnet, a single QF magnet, and two QD magnets. These magnets are described in Section 4.10. The cells are 4.1-m in length and the horizontal and vertical phase advances per cell are roughly 0.29 and 0.20. These tunes were chosen to balance the emittance contribution against the chromaticity generated by the cells while constraining the vertical beta function in the bending magnet. Finally, the lattice functions and magnet locations are plotted in Figure 4-17.

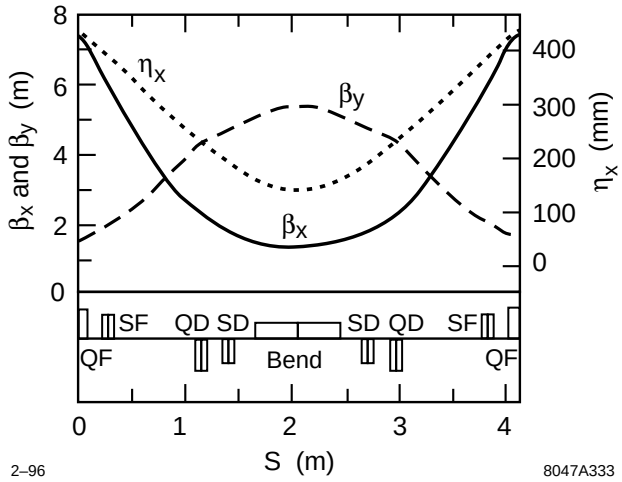


Figure 4-17. *β_x (solid), β_y (dashes), and η_x (dots) in the pre-damping ring FOBO cell.*

The dipole steering correction is performed using trim windings on the sextupoles. There is one horizontal corrector and one vertical dipole corrector in each cell. In addition, there are two BPMs per cell: one located adjacent to the bending magnet and one located next to the focusing quadrupole.

The chromatic correction is performed using two families of sextupoles: a set of defocusing sextupoles located between the bending magnet and the defocusing quadrupoles and a set of focusing sextupoles located between the defocusing and focusing quadrupoles. The symmetric placement of the sextupoles optimizes the dynamic aperture although it doubles the required number of sextupole magnets; this should be further investigated in the future.

The required vacuum aperture is determined from $X, Y = \sqrt{2\beta\epsilon_{x,y}} + \eta_{x,y}\delta p/p$. To provide sufficient aperture for an edge emittance of $\gamma\epsilon_{x,y} = 0.09$ m-rad and $|\delta p/p| = \pm 1.5\%$ plus 2-mm clearance around the beam for misalignments and mis-steering, the inner chamber size must be 5.8 by 1.8 cm at the focusing quadrupoles and 2.0 by 3.6 cm at the bending magnet. To minimize the impedance of the chamber, we will use a uniform elliptical cross section of an inner aperture of 6 cm by 3.6 cm; this will be discussed further in Section 4.6.2.

Finally, the dispersion is matched to zero at the end of the arcs for the two straight sections. The matching is performed using the two upstream quadrupoles and a bending magnet with an integrated dipole strength that is half that of the arc bending magnets. These matching bends have the same magnetic field as the main arc bends and half the effective length, but they have no field gradient to ease the field requirements and the fringing fields. As in the main damping ring, the horizontal and vertical dispersion matches will be tuned using the dispersion matching quadrupoles and trim windings on the sextupoles wired as skew quadrupoles.

Injection/Extraction

The injection and extraction regions are constructed from FODO cells in the dispersion-free straight sections. The two regions are located on opposite sides of the ring and utilize most of the respective straight sections to reduce the requirements on the kickers and the septa. Both systems are identical except for the number of kicker modules that are used; the extraction line needs fewer because of the smaller out-going beam. The lattice functions and component placement in the injection linac are illustrated in Figure 4-18.

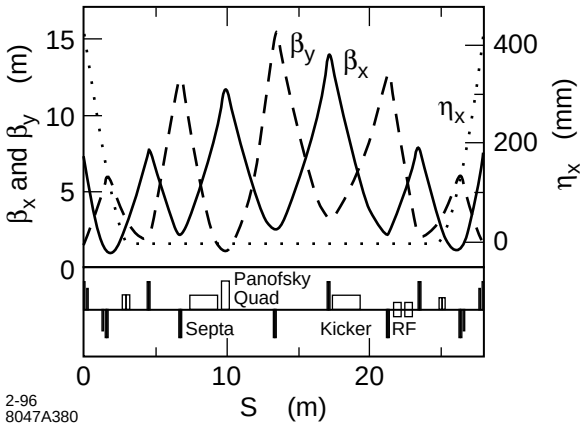


Figure 4-18. *β_x (solid), β_y (dashes), and η_x (dots) in the pre-damping ring injection region.*

The injection kickers are constructed in six segments with a total length of 2 m and provide an 8-mr horizontal deflection. The large deflection is needed because of the very large injected beam emittances. The extraction kicker is similar except it would be constructed from five of the kicker segments and would provide a 6.6-mr horizontal deflection. Both of these are described further in Section 4.10.3.

The septa are DC Lambertson septa that deflect the beam vertically. They are 2-m in length and have 5-kGauss fields providing a 150-mr deflection. The return blades are 2.3-cm thick and are located 3.3-cm from the stored and injected beam trajectories and 1.8-cm from the extracted beam trajectory; the extracted beams sizes are smaller and thus less space is required. In all cases, this leaves >1-cm clearance from the edge of the beam to the blade, assuming an injected edge emittance of 0.09 m-rad and an extracted beam emittance of 1×10^{-4} m-rad; the septa are described further in Section 4.10.2.

The focusing quadrupoles, adjacent to the septa, are constructed as dual Panofsky quadrupoles. This is necessary to make optimal use of the kickers. Both halves of the dual quad have the same strength and the same polarity. They are 50-cm in length and provide an integrated gradient of 28 kGauss. The horizontal aperture of each quadrupole is 6.2 cm. The center current sheet is 2-cm thick and, assuming an injected edge emittance of $\gamma\epsilon_{x,y} = 0.09$ m-rad, the blades are 8 mm from the stored and injected beams; these quadrupoles are described further in Section 4.10.2.

Finally, the adjacent defocusing quadrupoles may need to include an entrance or exit beam line; at the adjacent defocusing quadrupoles, the injection and extraction trajectories are roughly 30 cm from the stored beam trajectory.

After the beam has cleared the ring, it is deflected horizontally and then vertically to return to the plane of the damping ring. The optics of this section has not yet been developed but no significant problems are expected. Beyond this point the injection and extraction lines will be similar to those of the main damping rings.

Using the same technique as in the main damping rings, the effects of kicker jitter on the extracted beam will be reduced using an achromatic (double) kicker scheme. Here, a compensation kicker is located in the extraction line after the septum compensation bend. Six quadrupoles are used to match the compensation kicker so that it is separated from the extraction kicker by a $+I$ transform in the horizontal plane. The compensation kicker is powered from the same thyratrons as the extraction kicker and is expected to reduce the sensitivity to kicker jitter by at least a factor of four. This will ease the tolerance $\Delta\theta/\theta$ on the extraction kicker from 0.0016 to 0.0064. A similar achromatic kicker system will be tested at the ATF damping ring at KEK.

RF Cavities

The PPDR uses four 714-MHz rf cavities to provide the 2-MV rf field; the rf system is described in Section 4.5. To eliminate the transient beam loading when bunch trains are injected and extracted, one pair of rf cavities is placed in a 2-m drift section after the injection kicker and the other pair is located in a similar drift before the extraction kicker; these are illustrated in Figure 4-18. Although the cavity aperture is large compared to the vacuum chamber radius, additional masking will probably be needed to protect the cavities from beam loss at injection.

4.3.3 Injection Line

The injection lines transport the beams from the source linacs to the damping rings. The damping ring injection components from the septum compensation bend through the injection kicker, are considered part of the damping rings and are described in Sections 4.3.1 and 4.3.2. The injection lines have five primary functions:

- Provide diagnostics to monitor incoming the beam energy, energy spread, transverse position, and emittance.
- Provide transverse betatron matching for injection into the rings.
- Compress the incoming energy spread and match the longitudinal phase space to the damping rings.
- Rotate the longitudinal polarization of the electron beam into the vertical plane for injection into the electron damping ring.
- Include a beam dump just before injection into the rings.

Although the positron beam is not presently polarized, we are designing the positron injection line so that it could be easily upgraded to operate with polarized beams, *i.e.,* either polarized positrons or polarized electrons which would be needed for γ-γ or e^--e^- collisions. Thus, the geometrical constraints imposed by the spin rotation are applied to both the electron and positron injection lines and the two lines are very similar.

Layout

The layout of the electron injection line is plotted in Figure 4-19; the positron line is similar, although there is a longer rf section after the arc instead of the solenoid. Both injection lines are roughly 70 m in length. They include an arc of 60° which:

- Generates the R_{56} for the energy compressor.
- Provides for the energy diagnostics.
- Rotates the polarization vector by 270° in the x–z plane so that it is oriented in the x direction.

Following the arc, a 2.5-m S-band (2.856 GHz) rf (electrons) or 4-m L-band (1.428 GHz) rf (positrons) accelerator section is used to perform the energy compression. The L-band rf is needed in the positron line because the incoming positron beam emittance is very large and the aperture of the S-band structure is not sufficient; unfortunately, one needs twice the energy gain from the L-band structure as would be needed from an S-band structure.

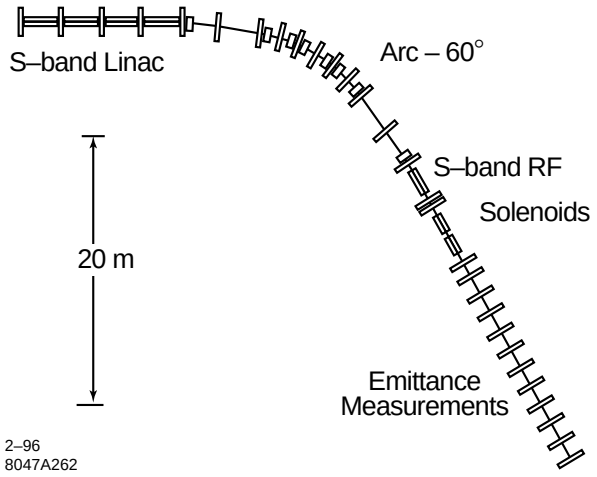


Figure 4-19. *Component layout of the electron injection line.*

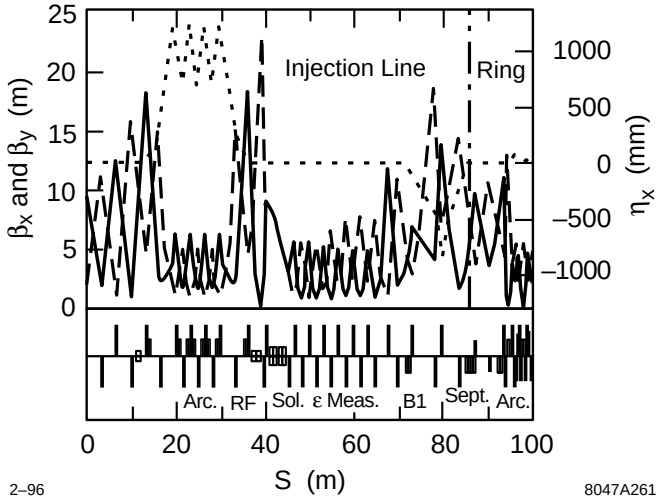


Figure 4-20. *β_x (solid), β_y (dashes), and η_x (dots) in the electron injection line with the solenoid on; the damping ring boundary is denoted by the vertical dashed line.*

Next, the electron side includes a solenoid to rotate the polarization from the horizontal plane to the vertical. This is followed by a series of FODO cells with five conventional wire scanners for emittance measurement; these wire scanners would be similar to those used throughout the SLC.

Finally, the beam is matched from the injection line into the ring. This section includes the septum compensation bend which generates a high dispersion that is ideal for a final energy and energy spread measurement. This portion of the line is described in Sections 4.3.1 and 4.3.2 and the component placement is illustrated in Figures 4-14 and 4-18. The bypass line, which includes a pulsed beam dump, intersects the injection line just before the septum compensation bend. A bending magnet is used to direct the beam to either the bypass line and beam dump or to the damping ring.

Optical Design

The electron injection line optical functions with the solenoids on are plotted in Figure 4-20; the positron line is similar. The lines are matched from the electron or positron source linacs into an arc that rotates the longitudinal spin polarization into the horizontal direction and acts as an energy diagnostic. The arc consists of three cells with a horizontal phase advance of 60° and a peak dispersion of 1.2 m.

This arc provides an $R_{56} = 0.8$ m for the energy compression and allows for a very accurate measurement of the beam energy by separating two BPMs by 180° and summing the horizontal measurements; the sum of the two measurements is independent of incoming betatron oscillations. The high dispersion also allows for accurate measurements of the electron and positron beam energy spread using wire scanners. For the electron beam, the dispersive beam size is roughly 1 cm while the betatron beam size is roughly 0.4 mm while in the positron system the expected energy spread extends between $\pm 2\%$ with a corresponding beam size of 2.4 cm at the high-dispersion points in the arc while the beam edge emittance of $\gamma\epsilon = 0.06$ m-rad corresponds to a size of 1 cm at the same location. Although this resolution is thought to be sufficient, if it is not, the dispersion can be increased in the design by increasing the length of the arc cells; this will be evaluated in the future.

In the electron line, the arc is followed by a 2.5-m S-band rf section which completes the energy compression and then two 1.5-m solenoids to complete the spin rotation. To prevent emittance dilution due to betatron coupling, the beta and alpha functions are set to be equal through the solenoid. The matching quadrupoles before and after the solenoid allow the betatron match to be maintained as the solenoid excitation is varied. The positron line differs in that it uses a 4-m L-band acceleration section and the solenoid is not needed.

Following the solenoids there is an emittance measurement section which consists of four FODO cells with five wire scanners interspersed. These wires will be used to measure the beam emittance, the betatron match, and the coupling mismatch from the solenoid; the matching quadrupoles can then be used to adjust the matching which when properly set will minimize the coupling mismatch as well as the betatron mismatch.

Finally, the beam line is matched past the septum compensation bend and into the ring; details of this are described in Sections 4.3.1 and 4.3.2. There is a dispersion maximum of 80 cm roughly halfway between the septum and the compensating bend where the final energy and energy spread will be measured just before injection into the rings.

Energy Compressors

The energy compressors reduce the incoming energy spread and the incoming energy jitter while increasing the bunch length. The compressor is required for the positron system because the pre-damping ring rf bucket height is only $\pm 1.5\%$ while the incoming beam has an energy deviation extending between $\pm 2\%$. The compressor also matches the beam to the rf bucket in the ring, reducing the amplitude of the longitudinal quadrupole oscillation at injection.

In the case of the electron beam, the compressor is not completely necessary since the rf bucket height is sufficient to accept the incoming beam. Regardless, the compressor will reduce the sensitivity to energy jitter and improve the longitudinal match at injection. In both cases, the compressors are designed to compress the energy spread and increase the bunch length by a factor of two. Parameters of the two compressors are given in Table 4-9 and the incoming and outgoing longitudinal phase spaces for the electrons and positrons are plotted in Figures 4-21 and 4-22, respectively.

Details of the energy compressor rf design can be found in Chapter 6. In both the S-band (electron) and L-band (positron) compressors, the structures will be Damped-Detuned Structures similar to those developed for the X-band main linacs. The L-band rf power will be provided by two 35-MW klystrons and a SLED rf pulse compressor. In the S-band case, the rf will be provided by a single 56-MW klystron which is similar to the SLAC 5045 klystrons. In both cases, the multibunch loading will be compensated using the ΔT (early injection) method. The positron compressor

	Electron	Positron
$(\Delta E/E)_{inj}$ (95%) (%)	±0.6	±1.8
$\sigma_{z\,inj}$ (mm)	1.5	3.7
R_{56} (m)	0.8	0.8
T_{566} (m)	2.0	2.0
f_{RF} (MHz)	2856	1428
V_{RF} (MV)	42	80
L_{RF} (m)	2.5	4
$(\Delta E/E)_{ext}(95\%)$ (%)	±0.4%	±1.0%
$\sigma_{z\,ext}$ (mm)	3.33	6.97

Table 4-9. *Parameters of electron and positron energy compressors.*

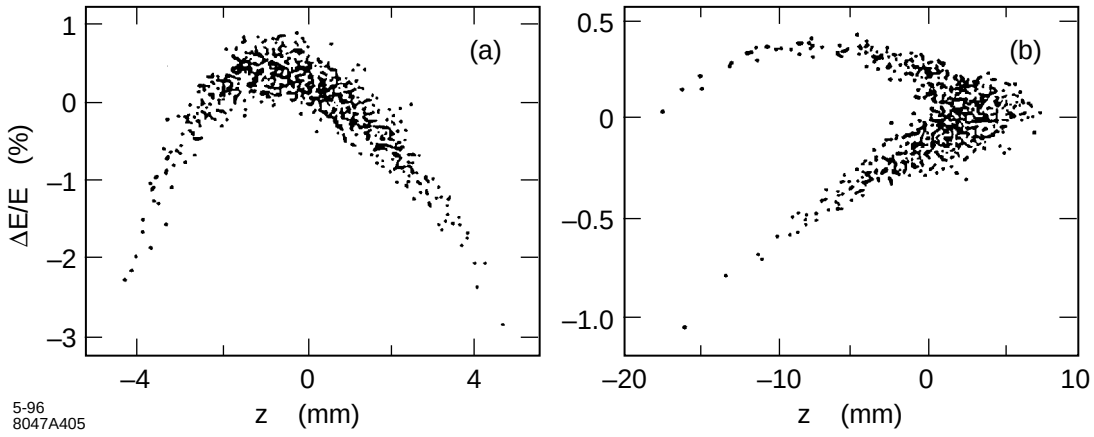


Figure 4-21. *Longitudinal phase space of electrons before (a) and after (b) energy compressor.*

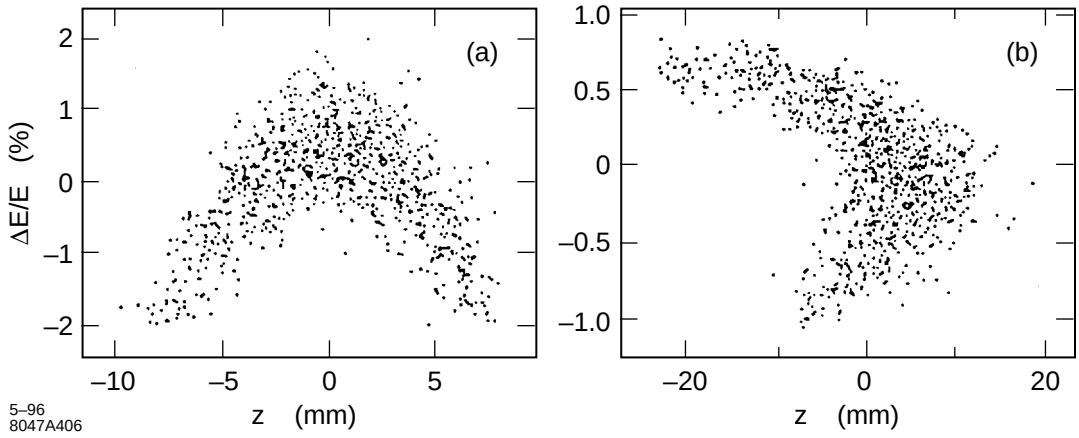


Figure 4-22. *Longitudinal phase space of positrons before (a) and after (b) energy compressor.*

has been specified to compensate for a maximum beam current of 2.75 A (2.4×10^{10} per bunch) which is roughly 85% greater than the charge needed at the IP to provide for beam loss through the injector system. The electron compressor has been designed to compensate for maximum beam current of 2.2 A (1.9×10^{10} per bunch) which is roughly 50% greater than the charge needed at the IP. In both cases, the beam loading should be compensated to an error less than a few 10^{-4} relative energy deviation; the tuning of the beam loading compensation will be done using either the energy spectrometer in the beam dumps or the dispersion in the arcs of the rings.

Polarization

The longitudinal polarization of the electron beam must be rotated into the vertical plane before injection into the main damping ring. This is performed using the 60° arc which rotates the polarization vector through 270° in the x–z plane at the nominal beam energy of 1.98 GeV. A solenoid is then used to rotate the polarization into a vertical orientation. The rotation in the solenoidal field is given by

$$\theta_s = \left[1 - \frac{g-2}{2}\right] \frac{B_z L_{sol}}{B\rho} \quad . \tag{4.21}$$

To perform the 90° rotation at 1.98 GeV requires an integrated field of 103.9 kGauss-m. We assumed two superconducting solenoids with lengths of 1.5 m and fields of 34.62 kGauss; these are similar to the superconducting solenoids used in the SLC damping ring injection and extraction lines.

4.3.4 Extraction Line

The main damping ring extraction lines only extend through the end of the achromatic kicker systems; the achromatic kicker system is discussed in Section 4.3.1. The extraction line optics is shown in Figure 4-23 and the component layout is illustrated in Figure 4-14 which is in Section 4.3.1.

The beam line downstream of the extraction line is part of the spin rotator and bunch compressor section and is described in Chapter 5. Regardless, it is worthwhile noting the requirements from the damping ring viewpoint.

- Energy and energy spread measurement: This will be done using the dispersion in the 20° arc which is part of the spin rotator system and is described in Chapter 5. In the same manner as in the injection line arc, BPMs located 180° apart in horizontal phase advance can provide a very accurate measure the the beam energy deviations.
- Bunch length measurement: It will be desirable to have a continuous parasitic measurement of the bunch length from the damping rings. This will be done by measuring the beam size in the bends of the bunch compressor. The calibration depends upon the compressor rf voltage and the dispersion in the bending magnet; this is discussed in Chapter 5.
- Transverse phase space measurement: The present design has a laser wire scanner located after the beta-matching quadrupoles at the end of the extraction line. This will provide a continuous parasitic diagnostic of the damping ring performance and can be used for quad-scan emittance measurements when necessary. In addition, there are four laser wire scanners in the prelinac at the end of the first bunch compressor. The emittance from the damping rings can be measured by turning off the bunch compressor spin rotation solenoids and the bunch compressor rf. This will be the main diagnostic for most tuning operations.
- Matching: The beta-matching will be performed using six quadrupoles at the end of the extraction line and the beginning of the bunch compressor system which is described in Chapter 5. The dispersion-matching will

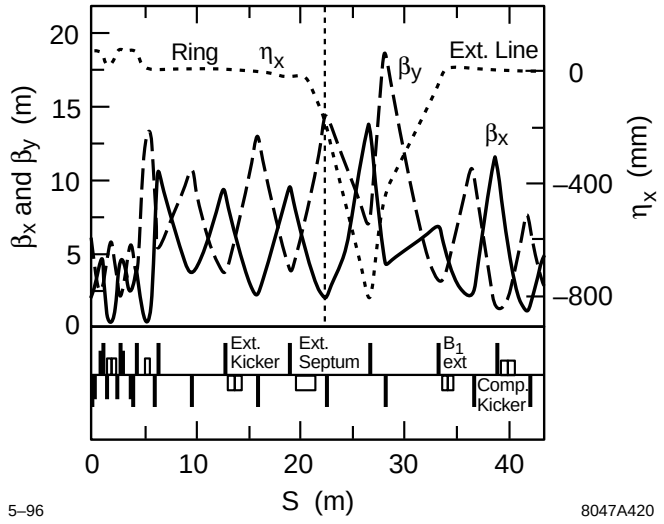


Figure 4-23. *β_x (solid), β_y (dashes), and η_x (dots) in the main damping ring extraction line; the damping ring boundary is denoted by the vertical dashed line.*

be performed using both trim quadrupoles and trim skew quadrupoles located in the extraction line and trim quadrupoles located in the 20° arc of the spin rotator system. Finally, the residual betatron coupling will be corrected with a skew correction system located between the spin rotators and the first bunch compressor; this is also described in Chapter 5.

- Pulsed dump line: A pulsed dump line will be located in the prelinac following the emittance diagnostics at the end of the first bunch compressor. This will allow full operation of the damping ring complex for commissioning and tuning.

4.3.5 PPDR-to-PDR Transfer Line

The design of the PPDR-to-PDR transfer line has not been completed yet. In concept it will combine the pre-damping ring extraction line, described in Section 4.3.2, with the main damping ring injection line, described in Section 4.3.1. In addition, the line should include a single wire scanner with a variable quadrupole to perform a quad-scan emittance measurement and a beam dump before the main damping ring. If the beams are polarized, the spin polarization will already be oriented vertically and no further manipulation is necessary.

4.3.6 Bypass Lines

The design of the damping ring bypass lines has not been completed yet. As illustrated in Figures 4-1 and 4-2, the primary component of the bypass lines will be the beam dumps. Since the bypass lines will branch off after the emittance diagnostics, these will allow tuning of the electron and positron sources while the damping rings are being maintained. These beam dumps will need to absorb the full beam power of roughly 130 kW.

Sextupoles ($X/Y/\Theta$) (μm/ μr)	100 / 50 / 1000
Quad. QFH ($X/Y/\Theta$) (μm/ μr)	100 / 100 / 600
Quad. QM2 ($X/Y/\Theta$) (μm/ μr)	100 / 100 / 600
Other Quad. ($X/Y/\Theta$) (μm/ μr)	100 / 100 / 1000
Bends ($X/Y/\Theta$) (μm/ μr)	200 / 200 / 1000
BPMs ($X/Y/\Theta$) (μm/ μr)	100 / 100 / 1000
All magnets have 0.1% rms strength errors.	

Table 4-10. *Rms errors for EDR and PDR orbit and emittance simulations.*

	Rms Θ_X (μrad)	Max Θ_X (μrad]	Rms Θ_Y (μrad)	Max Θ_Y (μrad)
Arcs	155	433	132	468
Straights	60	201	54	223

Table 4-11. *Corrector strengths from 100 simulations in the EDR and PDR with errors from Table 4-10.*

4.3.7 Trajectory Correction

The main damping rings and downstream systems all have very tight alignment and field tolerances to prevent dilution of the transverse emittances. In all cases, these tolerances will be attained using beam-based alignment techniques where the alignment accuracy depends on beam measurements and not mechanical systems; the techniques to be used in the damping rings are described in Section 4.9.

In the downstream systems, the alignment will be implemented using remote magnet movers. The tolerances are sufficiently loose in the damping rings that there is some question as to whether magnet movers are required; once the alignment is known, one could mechanically perform the differential moves to attain the desired alignment. This question will be determined in the future, for this design, we have assumed magnet movers.

In this section, we will describe the placements and required strengths of the trajectory correctors. These correctors must be able to provide sufficient steering to store beams before the beam-based alignment procedures are started. In addition, they must correct for residual alignment errors after beam-based alignment. In most cases, the corrector strengths required to simply store the beams are very weak compared to those required to steer the trajectory and minimize the BPM readings.

Main Damping Rings

In the main damping rings, each arc cell (Figure 4-10) contains two horizontal dipole correctors, a single vertical dipole corrector, and three BPMs. The horizontal correctors will be located at the S1 sextupole magnets (the sextupoles adjacent to the bending magnet). The vertical corrector will be located at either one of the S3 sextupoles (the sextupoles adjacent to the QD magnet). Both horizontal and vertical dipole correctors will consist of trim windings on the sextupole magnets and are discussed in Section 4.10.1.

Like the horizontal correctors, two BPMs will be located between the S1 sextupole magnets (the sextupoles adjacent to the bending magnet) and the QF quadrupole magnets. At this time, we have not determined how the BPMs are

mounted but, ideally, they would be constrained transversely at either the sextupole or the quadrupole magnet. The sextupoles are the preferred choice because they are separated by roughly 108° in horizontal phase advance while the QF magnets are separated by 144° and the orbit relative to the magnetic center of the sextupoles is more important than that of the quadrupoles. The third BPM would be located at the QD magnet; it is not strictly necessary for the trajectory correction or emittance control, but it provides redundancy if one of the other BPMs in the cell fails.

In both straight sections, BPMs will be located at every quadrupole, while horizontal dipole correctors will be located at the focusing quadrupoles and vertical dipole correctors will be located at the defocusing quadrupoles. Here, the dipole correctors will consist of trim windings on the quadrupoles.

Finally, additional horizontal and vertical dipole correction elements will be needed in the injection/extraction region to allow optimal control of the trajectory through this region. Small air core dipoles will be used for this function.

The orbit correction has been simulated using a standard set of errors listed in Table 4-10. All errors have Gaussian distributions truncated at ± 2.5 sigma. For the simulations, the arc BPMs and horizontal correctors were placed at the non-optimal QF magnets. After correction, the residual trajectory at the monitors was roughly 100-μm rms. The rms and maximum corrector strengths found from 100 simulations are listed in Table 4-11. The correctors are subdivided into arc correctors and straight section correctors. It is suggested to design the arc correctors for a maximum deflection of 0.50 mr (33 G-m field) and the straight section correctors for a maximum deflection of 0.25 mr (16 G-m field). These strengths will be more than sufficient to store beams after the preliminary conventional alignment where the expected errors are roughly four times those in Table 4-10. In addition, if during normal operation, the errors are significantly larger and many of the correctors are at maximum strength, the rings must be realigned.

Pre-Damping Ring

In the pre-damping ring, each arc cell (Figure 4-17) contains one horizontal dipole corrector, a single vertical dipole corrector, and two BPMs. The horizontal corrector will be located at one of the SF sextupole magnets (the sextupoles adjacent to the focusing quadrupole). The vertical corrector will be located at either one of the SD sextupoles (the sextupoles adjacent to the bending magnet). Both horizontal and vertical dipole correctors will consist of trim windings on the sextupole magnets and are discussed in Section 4.10.2.

Like the correctors, one BPM will be located between the SF sextupole magnets and the adjacent QF quadrupole magnets. The other BPM would be located between the bending magnet and one of the SD magnets. At this time, we have not determined how the BPMs are mounted but, ideally, they would be constrained transversely at either the sextupole or the quadrupole magnet.

In both straight sections, BPMs will be located at every quadrupole while horizontal dipole correctors will be located at the focusing quadrupoles and vertical dipole correctors will be located at the defocusing quadrupoles. Here, the dipole correctors will consist of trim windings on the quadrupoles.

The orbit correction has been simulated using a standard set of errors listed in Table 4-12. All errors have Gaussian distributions truncated at ± 2.5 sigma. For the simulations, the arc BPMs and horizontal correctors were placed at the non-optimal QF magnets while the defocusing correctors were placed at one of the QD magnets. After correction, the residual trajectory at the monitors was roughly 110-μm rms while the actual trajectory had an rms and maximum values of roughly 130 μm and 269 μm in the horizontal plane and 390 μm and 1014 μm in the vertical plane. These trajectory offsets are consistent with the 2-mm steering allowance provided in the aperture calculations.

The rms and maximum corrector strengths found from 100 simulations are listed in Table 4-13. It is planned to design the correctors for a maximum deflection of 1.0 mr (67 G-m field). These strengths will be more than sufficient to store beams after the preliminary conventional alignment where the expected errors are comparable to those in Table 4-12.

Sextupoles ($X/Y/\Theta$) (μm/ μr)	200 / 200 / 2000
Quadrupoles ($X/Y/\Theta$) (μm/ μr)	200 / 200 / 2000
Bends ($X/Y/\Theta$) (μm/ μr)	200 / 200 / 1200
BPMs ($X/Y/\Theta$) (μm/ μr)	200 / 200 / 2000
All magnets have 0.1% rms strength errors.	

Table 4-12. *Rms errors for PPDR orbit and matching simulations.*

Rms Θ_X (μrad)	Max Θ_X (μrad)	Rms Θ_Y (μrad)	Max Θ_Y (μrad)
166	532	347	1104

Table 4-13. *Corrector strengths from 100 simulations in the PPDR with errors from Table 4-12.*

Transfer Lines

No simulations have been performed at this time. The primary issues will be the aperture issues in the positron injection line, injection onto the closed orbit in both rings, trajectory control through the extraction septa. The aperture issues in the positron injection line will be resolved by including a $\pm$5-mm clearance for steering and including numerous BPMs and dipole correctors. The injection and extraction trajectories will be adjusted with additional steering correctors in the injection/extraction regions. The apertures of the septa have been chosen so that there should be plenty of room to optimize the trajectories.

4.3.8 Emittance Control and Matching

In addition to causing orbit deviations, alignment and field errors generate dispersion and betatron mismatches. In all of the damping rings, these increase the horizontal and vertical emittances but are only significant in the main damping rings where the equilibrium emittances are very small. In fact, in the main damping rings, the vertical emittance requirement determines alignment tolerances; other effects such as steering and dynamic aperture have looser tolerances.

In the pre-damping ring, which operates on the betatron coupling difference resonance, the tolerances for emittance control are relatively loose; other effects such as physical and dynamic aperture limitations are more important. Thus, the matching into- and in the pre-damping ring needs to be controlled. In addition, the pre-damping ring has a problem which is opposite that in the main damping rings, namely, the width of the coupling resonance is extremely small and, for stable operation, this will probably need to be increased with skew quadrupoles or vertical orbit bumps through the sextupoles.

In the transport lines, similar effects will increase the effective emittances injected into the rings, thereby increasing the time needed to damp to the desired emittances, and possibly leading to particle loss if the mismatched beam exceeds the aperture. We will first describe the issues for the main damping rings where the very small vertical equilibrium emittance determines the alignment tolerances; the tolerances needed to attain the vertical emittance are more severe than those due to the dynamic aperture limitations. Next, we will discuss the pre-damping ring and finally we will discuss the injection and extraction matching requirements.

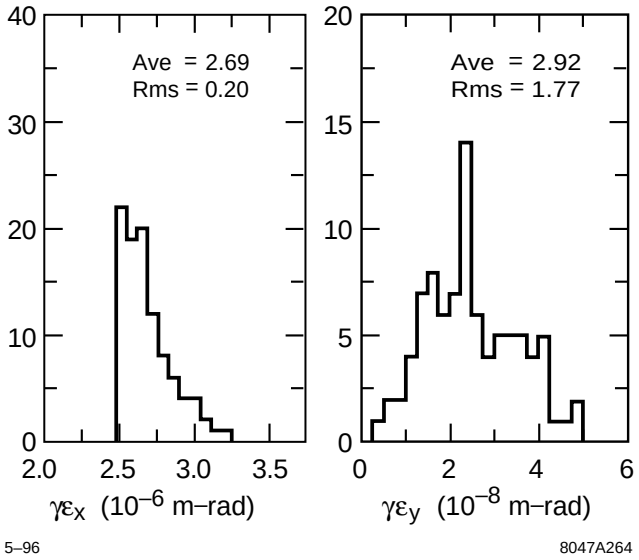


Figure 4-24. *Horizontal and vertical emittances from 100 simulations of errors and only using orbit correction; the average horizontal emittance is $\gamma\epsilon_x = 2.69 \times 10^{-6}$ m-rad and the average vertical emittance is $\gamma\epsilon_y = 2.92 \times 10^{-8}$.*

Although these effects will set tolerances, this section does not discuss how these tolerances will be attained and maintained stably. This will be discussed in Sections 4.9 and 4.8.

Main Damping Rings

In the NLC main damping rings, the primary sensitivity of the equilibrium emittances arises from dispersion mismatches; this is true in both the horizontal and vertical planes. Other effects such as betatron mismatches and betatron coupling are less important. In particular, the expected vertical emittance dilution from vertical dispersion due to random errors is roughly five times greater than the dilution due to coupling from the same errors.

The three elements that have the greatest contribution to the emittances are the horizontal and vertical misalignments of the arc sextupoles, rotational misalignments of the arc focusing quadrupoles, and rotational misalignments of the bending magnets. To reduce the emittance dilutions, it is most efficient to set lower tolerances on these sources and have more relaxed tolerances on the other elements.

We have performed 100 simulations using the errors listed in Table 4-10. As before, all errors have Gaussian distributions truncated at ± 2.5 sigma. Orbit correction was performed as discussed in the previous section. Without any other correction of the lattice, we found an average horizontal emittance of $\gamma\epsilon_x = 2.69 \times 10^{-6}$ m-rad and an average vertical emittance of $\gamma\epsilon_y = 2.92 \times 10^{-8}$ m-rad; for comparison, the design values for horizontal and vertical emittance are $\gamma\epsilon_x = 2.6 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 1.5 \times 10^{-8}$ m-rad. Histograms of the resulting emittance values are plotted in Figure 4-24.

In general, the dispersion generated by the alignment errors is easily corrected. The mismatch will oscillate much like an uncorrected close orbit and a few dispersion correctors can reduce the rms value significantly. The horizontal dispersion can be corrected using weak quadrupoles in regions of large horizontal dispersion and the vertical dispersion can be corrected using weak skew quadrupoles in regions of large horizontal dispersion.

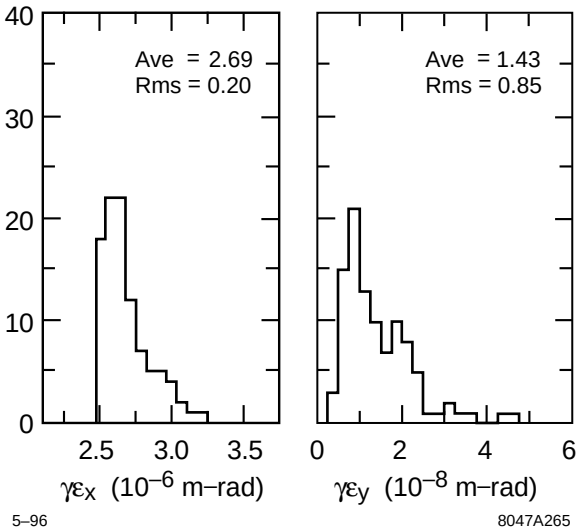


Figure 4-25. *Horizontal and vertical emittances from 100 simulations of errors where two weak skew quadrupoles were used to reduce the vertical dispersion; the average vertical emittance is* $\gamma\epsilon_y = 1.43 \times 10^{-8}$.

In Figure 4-25, we plot the residual emittance after using just two skew quadrupoles, located in the dispersion suppressor cells, to minimize the rms vertical dispersion. Notice that both the average of the vertical emittance and the width of the distribution are decreased. In this case, the average vertical emittance is $\gamma\epsilon_y = 1.43\times10^{-8}$ m-rad and 95% of the simulations have emittances below $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad; further correction would be possible with additional elements. This minimization can be performed either by directly measuring the dispersion or by optimizing the emittance of the extracted beam; in the simulation, we optimized the extracted beam emittance. Finally, as mentioned, a similar procedure can be used to correct the horizontal dispersion. We have not yet implemented such a solution but it will be straight-forward.

The current design for the main damping rings specifies individual power supplies for all the quadrupoles and sextupoles. Thus, there are numerous quadrupoles that can be used to correct betatron mismatches and horizontal dispersion mismatches. In addition, we are considering using magnet movers on most of the quadrupoles and sextupoles. If remote magnet movers are adopted then skew correctors to correct the vertical dispersion will exist all around the ring, *i.e.,* vertical motion of the sextupoles. If we do not choose to use the remote magnet movers, then eight sextupoles in each arc will have trim windings wired as skew quadrupoles.

As discussed previously, the alignment tolerances will be attained using beam-based alignment practices which are discussed in Section 4.9 while the tolerances required to maintain the required stability are discussed in Section 4.8.

Pre-Damping Ring

As stated, the issues are completely different in the pre-damping ring. Here, because the aperture is tight, the primary problem becomes maintaining the matched beta and dispersion functions. With the individual quadrupole power supplies it is straight-forward to both measure and correct any betatron mismatches that may exist. Similar arguments apply to correcting dispersion mismatches which are easily measured by storing the beam and varying the rf frequency.

Finally, because of the accurate alignment that is required to attain the desired dynamic aperture, the width of the betatron resonance is relatively small. This will make reliable operation on the coupling resonance difficult and so

skew quadrupoles will be used to increase the width of the resonance; this technique was also used on the SLC Positron Damping Ring.

Injection Line

The injection lines have relatively loose alignment and field tolerances. The primary issue is matching the beams to the damping rings. This is required to prevent emittance growth due to filamentation and possible beam loss if the mismatched beam exceeds the damping ring aperture.

Both injection lines have emittance diagnostics just upstream of the rings so that incoming mismatches can be corrected. The fine tuning of the beam match will be performed using gated synchrotron light cameras, located in the rings, to measure the beam size on different turns after injection. This technique has been used to successfully match the incoming phase space to the SLC damping rings [Minty 1995] and we expect to be able to match the incoming beam sizes to better than 10%.

Extraction Line

In the extraction lines from the main damping rings, the beam emittances are asymmetric and very small. Thus, betatron coupling and dispersive errors can lead to significant emittance growth. The betatron coupling will set tolerances on the allowable magnet roll and skew fields. This is particularly true for the extraction septa which tend to have nonlinear fields close to the septa blades. Similarly, the dispersive errors will set tolerances on the allowable orbit deviations, magnet alignment, and matching of the dispersion.

In the NLC damping rings, we are including multiple steering correctors to optimize in the injection and extraction trajectory through the septa. In addition, we are designing the septa with substantial clearance between the beam trajectory and the septa blades to minimize the nonlinearity seen by the beams.

The extraction lines will include six matching quadrupoles to match the beta functions into the spin rotators beam lines which follow; these are described in Chapter 5. In addition, residual betatron coupling from the damping rings or the extract lines will be corrected in the first bunch compressor skew correction section; this is also described in Chapter 5. In simulations, this correction section has had no difficulty removing large coupling from the extracted beams. We do not plan on installing any additional skew quadrupoles to control the betatron coupling in the extraction lines.

Trim quadrupole and skew quadrupoles will be located in the dispersive region of the extraction lines (Figure 4-23) to correct for residual dispersion leaking from the damping rings. In addition, correcting quadrupole elements will be placed in the 20° arc which is part of the spin rotator that follows the damping ring extraction lines.

Most tuning of the extracted beam will be performed using the emittance diagnostic station at the end of the first bunch compressor; this is described more fully in Chapter 5. To measure the extracted beams, the compressor rf and spin rotators will be turned off. In addition, a single laser wire will be located in the extraction line to monitor the ring performance. It will be possible to perform quad-scan emittance measurements on this wire for further tuning of the extracted beam.

Transfer Line

At this time, we have not yet designed the PPDR-PDR transfer line. It is expected that the beam line will be a combination of the injection and extraction lines just discussed and is not thought to present any significant problems.

Source	Error	Rms Magnitude
All Elements except Bends	$X,\ Y$ (μm)	100, 100
All Bends	$X,\ Y$ (μm)	200, 200
All Elements	Z [mm], Θ (mr)	1, 1
All Magnets	$\Delta B/B\ (10^{-3})$	1

Table 4-14. *Rms errors for main damping ring dynamic aperture simulations.*

The line will include a single conventional wire scanner to monitor the pre-damping ring performance. In addition, quad-scan emittance measurements could be made on this wire when desired.

4.3.9 Dynamic Aperture

The dynamic aperture is usually a limitation in low-emittance strong-focusing rings such as the NLC main damping rings. It is of particular concern in these cases, because the the rings must be designed to have minimal injected beam loss and the injected beam emittances are many orders of magnitude larger than the extracted emittances. The dynamic aperture is also a limitation for the pre-damping ring because of the very large injected beam emittances.

At this time, we have developed solutions for both the main damping rings and the pre-damping ring with very satisfactory performance. The tolerances in the main damping rings to attain a dynamic aperture that is in excess of $10\sigma_{x,y}$ of the injected beams are looser than the tolerances that are required to attain the desired extracted vertical emittance. We believe that these tolerances will be attained using beam-based alignment techniques and do not present a limitation.

In the pre-damping ring, the tolerances to attain a dynamic aperture in excess of the physical aperture are even looser. Here, the alignment tolerances could probably be attained using conventional alignment practices although the ability to use beam-based alignment will exist.

Finally, the multipole tolerances on the magnets in both rings are also straightforward. The tolerances on the main and pre-damping ring magnets are looser than the multipoles components that are being measured on the KEK ATF damping ring magnets and are comparable to those of the PEP-II LER ring magnets.

Main Damping Rings

The goals for the dynamic aperture are an aperture in excess of 10 sigma for an injected beam with the nominal injected emittance of 1×10^{-4} m-rad and a momentum aperture of $\pm 1\%$. This will allow the rings to operate during the commissioning phases where the injection beam properties will likely be poor and it will reduce the sensitivity to injection trajectory errors. Of course, this goal far exceeds the aperture requirements. The electron and positron injectors are being designed assuming a 10% loss of charge at injection into the main damping rings for which an aperture of 3 sigma of the injected beam size would more than suffice.

The dynamic aperture studies have been performed using MAD (Vers. 8.1) and TRACY; it should be noted that both of these codes use large ring approximations and the validity of this should be verified. The results, including energy errors of $|\delta p/p| \leq 1\%$, are illustrated in Figure 4-26. The outer curve shows the acceptance of the bare lattice while

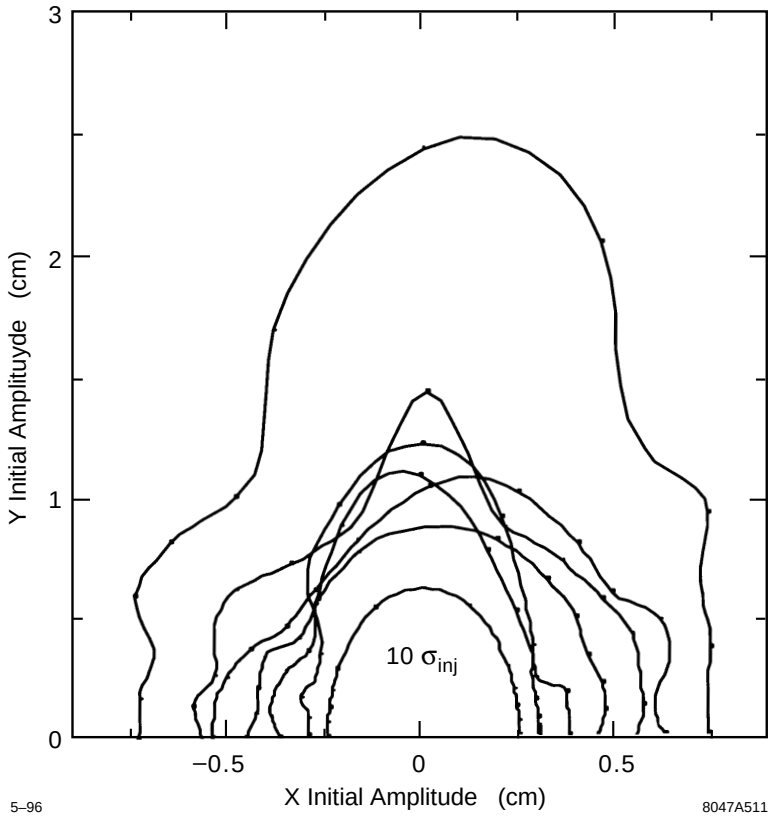


Figure 4-26. *Main damping ring dynamic aperture with multipole and alignment errors; the outer solid curve is the aperture of the bare lattice while the middle five curves are the aperture with different sets of errors and the inner solid curve is 10 sigma of the injected beam.*

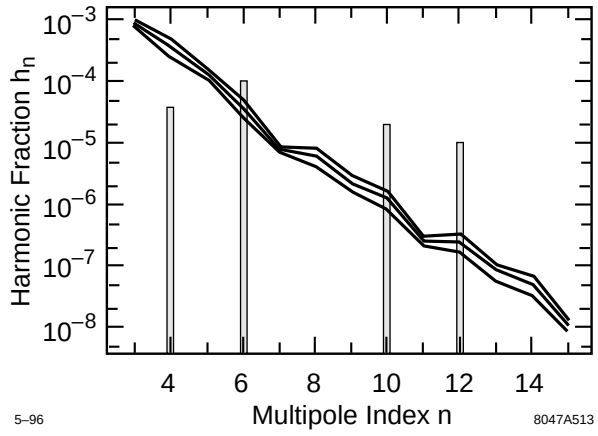


Figure 4-27. *Systematic (bars) and random (solid line) multipoles in the quadrupoles where the harmonic fraction is evaluated at 2/3 of the quadrupole bore.*

the inner curve is the goal of 10 times the injected beam size. One can see that the bare rings have an acceptance that is more than double the desired aperture.

The five middle curves in Figure 4-26 show the aperture when alignment and multipole errors are included, each curve corresponding to a different seed of random errors. In each case, the ring orbits were corrected with the dipole correctors and the tunes were corrected back to the nominal values using the arc quadrupoles.

The rms alignment errors that were used are listed in Table 4-14; these are looser than the tolerances required to attain the desired vertical emittance. In addition, all magnets were assumed to have strength errors of 10^{-3} and higher-order multipole errors.

The quadrupole magnet multipoles that were included are shown in Figure 4-27. Here, the bars represent the systematic multipoles while the solid line is the magnitude of the random multipoles. In both cases, the harmonic fraction shown is the multipole component relative to the quadrupole component at 2/3 of the the magnet bore; the bore varies from 15 mm to 30 mm for the different quadrupole families.

In order to evaluate the magnitude of the random multipoles in the quadrupole magnets, we assumed that the quadrupoles would be constructed in two pieces, as was done for both the ALS and the APS synchrotron light sources. The random contribution arising from fabrication errors then has two components: an uncorrelated pole-to-pole piece (significantly reduced relative to four-separate-pole construction), which we take to be 0.77 mils (20 μm) rms in the pole placement in the transverse plane; and the two-pole correlated contribution which can characterized in terms of a relative translation (in the transverse plane) and rotation of the two halves. The former we take to be 1.5 mils (40 μm) in each of the two dimensions, and the latter to be 1.1 mr, corresponding to a linear displacement of 8.9 mils (225 μm) at a 20 cm arm length.

These pole-to-pole errors contribute to all multipoles and yield a symmetric (but non-uniform) multipole orientation angle distribution, while half-magnet translation errors generate even index pole contributions (*i.e.*, 8, 12, ...), and rotation errors generate odd index poles (*i.e.*, 6, 10, ...). Neglecting fringe fields and assuming ideal pole shapes, the anomalous multipoles generated by the above pole misplacements are precisely determined (to first order) by two dimensional magnetostatics [Halbach 1969].

The estimated systematic components are based on preliminary designs and experience from similar magnet designs. The 8- and 24-pole contributions occur only in the magnets which must accommodate a wide vacuum ante-chamber and thus necessarily deviate from reflection symmetry around each pole. The detailed design alternatives that affect the systematic multipoles have been given minimal attention; thus these systematic multipoles should be viewed as representative rather than definitive.

Finally, similar random and systematic multipoles have been included in the combined function dipoles. The systematic multipoles tentatively adopted are slightly looser than those achieved in prototyping the KEK ATF combined function magnet. It will be possible to extend the analysis of random multipoles done above for quadrupoles once a sufficiently definite design of the combined function pole shape is achieved.

With the errors, the dynamic aperture still exceeds the goal aperture although it is significantly smaller than the bare aperture of the ring. Of course, the dynamic apertures are sensitive to the correction that has been performed. As a trivial example, we considered the seed that led to the worst aperture and corrected the tunes with the straight section quadrupoles rather than the arc quadrupoles. In this case, the aperture improves by roughly 60%; in other cases, correcting the tunes with the straight section quadrupoles rather than the arc quadrupoles has little effect. Finally, we have also studied the aperture for tune variations of 0.01 about the nominal tunes and saw no noticeable change.

Finally, we have not yet included the multipoles from the sextupole magnets or the effect of the nonlinearity due to the wigglers. Both of these are expected to have a small effect but still need to be added.

Source	Error	Rms Magnitude
All Elements	$X,\ Y$ (μm)	200, 200
All Elements	Z (mm), Θ (mr)	2, 2
All Magnets	$\Delta K/K\ (10^{-3})$	1

Table 4-15. *Rms errors for PPDR dynamic aperture simulations.*

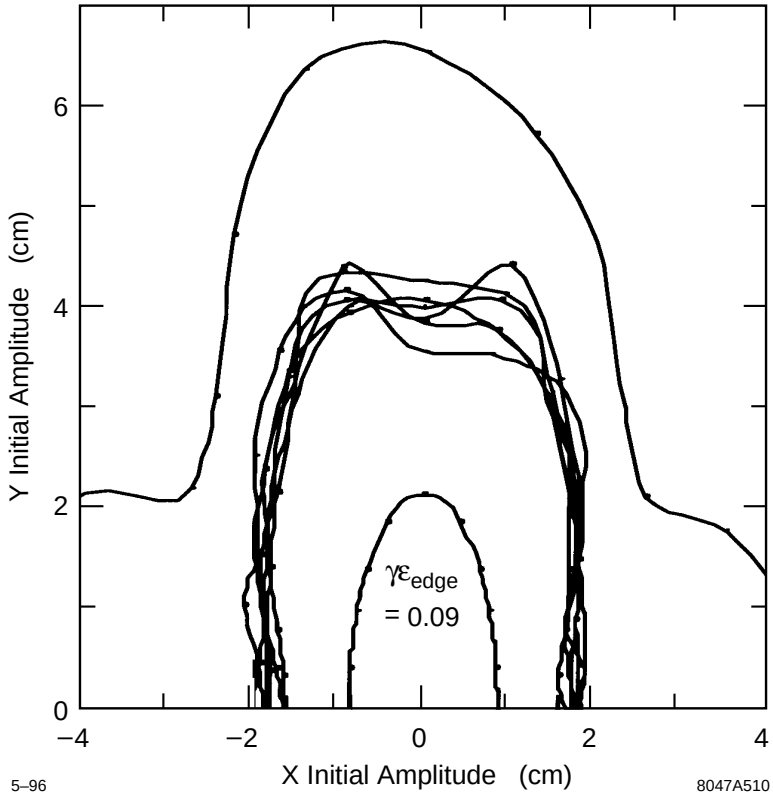


Figure 4-28. *Main damping ring dynamic aperture with multipole and alignment errors; the outer solid curve is the aperture of the bare lattice while the middle six curves are the aperture with different sets of errors and the inner solid curve is that required for an edge emittance of* $\gamma\epsilon_{x,y} = 0.09$ *m-rad.*

Pre-Damping Ring

In the pre-damping ring, the goal for the dynamic aperture is an acceptance in excess of the nominal injected edge emittance of 0.09 m-rad with a momentum aperture of $\pm 2\%$. This corresponds to a normalized acceptance of 0.09 m-rad where the acceptance is defined in terms of the particle action:

$$J_x = \frac{1}{2}\left(\gamma x^2 + 2\alpha x x' + \beta x'^2\right) \quad . \tag{4.22}$$

It should be noted that this is much larger than the aperture required. The positron source has been specified assuming 20% beam loss during injection into the pre-damping ring.

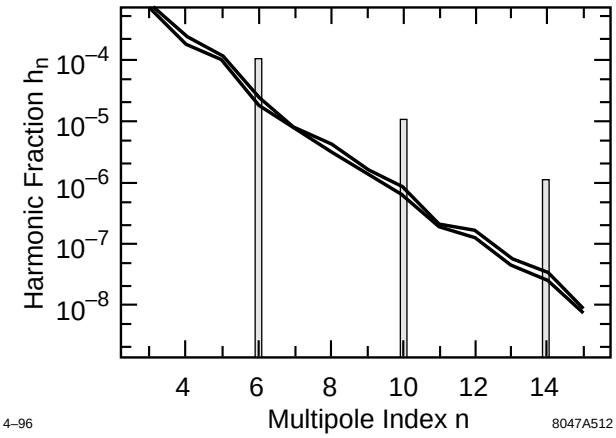


Figure 4-29. *Systematic (bars) and random (solid line) multipoles in the quadrupoles where the harmonic fraction is evaluated at 2/3 of the quadrupole bore.*

Again, the aperture studies have been performed using MAD (Vers. 8.1) and TRACY. We have studied the effect of alignment and magnet multipole errors. The results, including energy errors of $|\delta p/p| \leq 2\%$, are illustrated in Figure 4-28. The outer curve shows the acceptance of the bare lattice while the inner curve is the goal aperture of $\gamma J = 0.09$ m-rad. One can see that the bare rings have an acceptance that is more than three times the required aperture.

The six middle curves in Figure 4-28 show the aperture when alignment and multipole errors are included, each curve corresponding to a different seed of random errors. In each case, the ring orbits were corrected with the dipole correctors and the tunes were corrected back to the nominal values using the arc quadrupoles.

The rms alignment errors that were used are listed in Table 4-15; these are relatively loose tolerances and could be expected to be attained using conventional alignment practices. In addition, all magnets were assumed to have strength errors of 10^{-3} and higher-order multipole errors.

The quadrupole magnet multipoles that were included are shown in Figure 4-29. Again, the bars represent the systematic multipoles while the solid line is the magnitude of the random multipoles. In both cases, the harmonic fraction shown is the multipole component relative to the quadrupole component at 2/3 of the the magnet bore; the bore varies from 30 mm to 40 mm for the different quadrupole families.

As was done for the main rings, we assume that the quadrupoles would be constructed in two pieces with tolerances on 0.77 mils (20 μm) on the relative placement of the two poles on one piece and 1.5 mils (40 μm) on the placement of the two pieces. We also assumed random and systematic multipoles in the combined function dipoles that are based on those achieved in the prototype KEK ATF combined function magnet.

With these errors, we find that the aperture is almost double the required aperture. In the future, we will improve our estimates of the multipoles in the bending magnets and include multipoles in the sextupole magnets but we do not expect the aperture to change significantly.

4.3.10 Polarization

The polarization is assumed to be longitudinal at the exit of the electron source linac. It is rotated into the vertical plane in the injection line for injection into the electron main damping ring. The rotation is performed using the 60° arc followed by a solenoid. A similar arc is included in the positron injection line to allow the line to be easily upgraded for future polarized beams.

The polarization loss in the injection line due to the 1% energy spread in the beam is less than 1%. In contrast, if the injection line energy is changed, there will be large losses in polarization. This loss arises because the 60° arc no longer rotates the polarization by 270°. The fractional polarization loss is equal to:

$$\frac{\Delta P}{P} = (1 - \cos \Delta\Theta) \tag{4.23}$$

where $\Delta\Theta$ is the deviation from the 270° rotation: $\Delta\Theta = 270° \Delta E / 1.98\,\text{GeV}$. Thus, changing the ring energy by 100 MeV to 1.9 GeV or 2.1 GeV would result in a 2.7% loss in polarization through the injection line; the loss would increase to roughly 10% at 1.8 or 2.2 GeV.

At this time, we have not looked in detail at polarization issues in the damping rings. At the nominal operating energy of 1.98 GeV, the rings have a spin tune of 4.5; at this operating point, the depolarization is expected to be small. The operating energy range of roughly $\pm$ 220 MeV corresponds to shifts of ± 0.5 in spin tune. At these points, the spin tunes would be integral and we would expect full depolarization. Many other resonances exist that will limit the polarization within the ± 200-MeV operating range; these have not yet been studied although it would likely be possible to minimize their effect by modifying the ring tunes. We also have yet to study the effect of the beam emittance and the synchrotron radiation on the beam polarization. Both of these effects are expected to be negligible but have yet been considered.

Finally, after extraction from the damping rings, the polarization is rotated to the desired orientation in the bunch length compressors; these spin rotators are described in Chapter 5.

4.4 Collective Limitations

The NLC damping rings will operate with long trains of bunches at a high average beam current and thus coupled-bunch instabilities are an important limitation. In addition, to obtain the desired magnetic fields, the vacuum apertures are relatively small and the broad-band impedance sources can be large. Thus, both single and multibunch instabilities are potential limitations in the NLC damping rings.

Furthermore, because the performance requirements on the damping rings are tight, it is important to avoid instabilities of all sorts. This is also true of the so-called "benign" instabilities. For example, the "strong" longitudinal microwave instability, usually considered benign, sometimes exhibits a bursting mode that has been observed in a number of storage rings; such uncivilized behavior is not acceptable in a damping ring.

In this section, we review "collective" limitations for the main damping rings. We start by describing the impedance model for the ring. This is based upon MAFIA [Klatt 1986] calculations of the bellows, slots, masks, and BPMs, the resistive wall impedance, and the rf cavities; the rf cavities and the vacuum system are described in greater detail in Sections 4.5 and 4.6. Next, we discuss the single bunch instabilities, the synchronous phase variation along the bunch trains, and the coupled bunch instabilities. Finally, we address other effects including beam-gas scattering, intrabeam scattering, lifetime limitations, and ion trapping.

In addition to the usual series of instabilities, the damping rings may also be subject to some new effects that have been predicted theoretically but have not yet been seen experimentally. In particular, there are a photoelectron-positron instability, referred to as the "Ohmi effect," and a fast beam-ion instability that have been postulated. We address both of these issues although, at this time, the true implications are not clear.

In general, there are no known limitations in the main damping rings. The coupled-bunch instabilities require HOM damped rf cavities and a transverse coupled bunch feedback system but these technologies are being well developed for the new B-factories as well as for the KEK ATF damping ring. In addition, the broad-band impedance must be keep minimal to avoid the microwave instability but this should be possible with careful vacuum chamber design.

The two unknowns, the fast beam-ion instability and the electron-positron instability, will be verified in the near future. If they do present a limitation, there are a number of ways to ameliorate the effects; these are addressed in the respective sections.

Finally, this section deals primarily with the main damping ring. In most respects, the issues are less severe in the pre-damping ring since the bunch length is roughly twice as long, the vacuum aperture is much larger, and the beam emittances are much larger. The collective issues in the PPDR is are summarized in Section 4.4.11.

4.4.1 Main Damping Ring: Impedance Model

The ring impedance determines the beam stability and heating due to excitation of the higher-order modes (HOMs). The short damping ring bunches have a bunch spectrum that starts rolling off at 10 GHz, which is comparable to the 9-GHz cutoff frequency of the 1.25-cm-radius beam pipe. With such a broad bunch spectrum, even small irregularities of the beam pipe cross-section can give noticeable contribution to the beam impedance. Therefore, a careful design and analysis of the vacuum system and other impedance sources is required.

The total longitudinal impedance can be written as the sum of the narrow-band impedance (NB) and the broad-band impedance (BB):

$$Z_l^{tot} = Z_l^{NB} + Z_l^{BB} \quad . \tag{4.24}$$

The narrow-band impedance is primarily determined by the rf cavity HOMs and a few NB HOMs of the BPMs and bellows:

$$Z_l^{NB} = i\sum_m \frac{\omega_m}{2}\left(\frac{R}{Q}\right)_m \left[\frac{1}{\omega - \omega_m + i\omega_m/(2Q_m^L)} + \frac{1}{\omega + \omega_m + i\omega_m/(2Q_m^L)}\right] \quad . \tag{4.25}$$

The broad-band impedance is determined by the high frequency modes of the cavities and small discontinuities of the beam pipe such as the slot of the ante-chamber, bellows, BPMs, rf contacts, shallow tapers, etc. We can estimate the impedance by scaling from the inductance of the PEP-II. Scaling proportional to the machine circumference $2\pi R$ gives $L = 12$ nH. However, this scaling implies that the total number of inductive components in the ring scales as machine radius and that the inductive character of a component is not changed, *i.e.*, the dimensions of the elements scales in proportion to the rms beam size. A more accurate estimate can be obtained by modeling the vacuum components with the code MAFIA [Klatt 1986] as will be described below.

The broad-band impedance can be expressed:

$$Z_l^{BB} = -i\frac{Z_0}{(1 - i\omega a/c)^{3/2}}\frac{\omega L}{4\pi c} + (1-i)\,1.05\sqrt{\frac{\omega R}{c}} + (1+i)\frac{R_{cav}}{\sqrt{\omega}}\theta(\omega - \omega_c) \quad [\Omega] \quad . \tag{4.26}$$

The impedance of the inductive components also have a small resistive part in addition to the inductance. This is included in the first term in Eq. 4.26; the pure inductive impedance is modified to roll off at high frequencies as

Model	freq (MHz)	R_s [MΩ]	R/Q [Ω]	E_p (kV/m)	H_p (A/m)	k_l V/pC	$k_\perp$ V/pC/m
1	728.06	3.99	114.14	74.1	66.46	1.21	10.1
2	726.88	4.89	137.01	88.7	61.99	2.99	75.6
3	717.33	4.54	132.21	71.2	68.73	2.78	72.0

Table 4-16. *RF cavity models considered for MDR with $\sigma_z = 3.3$ mm.*

$1/\sqrt{\omega}$ which corresponds to the behavior of a cavity impedance at high frequencies. Such a model well describes the dependence of the loss factor on the bunch length. The parameter a defines the roll-off frequency and can be obtained from the loss factor of a bunch with the nominal rms length $\sigma_z = 4$ mm.

RF Cavities

The main contribution to the narrow-band impedance comes from the two rf cavities. As a first estimate, we modeled the cavities after the PEP-II rf cavity, scaled to the desired 714-MHz frequency. Since it is desirable to increase the beam pipe radius b at the cavity location to minimize the transverse kick factor, we considered three models with different beam pipe radii b and nose cone radii r:

1) b=3.1 cm, r=0.66 cm

2) b=1.4 cm, r=0.66 cm

3) b=1.4 cm, r=1.86 cm.

Table 4-16 gives frequency, shunt impedance R_s, R/Q, and the maximum E and H fields at the surface of the cavity, E_p and H_p, for the fundamental mode.

Cavity model 1, with 3.1-cm beam-pipe radius, has a total loss factor of $k_l = 1.21$ V/pC for an rms bunch length $\sigma_z = 3.3$ mm; this is shorter than the expected damping ring bunch length of $\sigma_z = 4.0$ mm. The maximum longitudinal impedance is $\Re Z_l = 0.21$ kΩ at 1 GHz while the maximum transverse impedance is $\Re Z_\perp = 170$ kΩ$/m^2$ at 2 GHz. The total loss factor of cavity model 2, with the 1.4-cm beam pipe, is $k_l = 3.0$ V/pC for the same 3.3-mm bunch length. The transverse kick factor is increased from $k_\perp = 10.2$ V/pC/m for cavity model 1 to $k_\perp = 75.6$ V/pC/m for cavity model 2 without a substantial gain in the shunt impedance of the fundamental mode.

The larger radius $r_c = 3.1$ cm of the beam pipe at the cavity location reduces the impedances which is preferable. The two rf cavities could be mounted on the same large radius beam pipe, which may be tapered to the regular 1.25-cm beam pipe at both ends. Separation of the cavities and the distance from a cavity to the nearest taper has to be optimized to avoid reduction of the Q factor of the fundamental mode and cross-talk between the cavities. The longitudinal loss and transverse kick factors for a model with a shallow 15-cm-long taper on both sides gives $k_l = 1.725\, V/pC$ and $k_\perp = 39.43$ V/pC/m; these numbers may be reduced by using longer tapers.

Such a design, however, leaves a large number of HOMs with frequencies up to the 9-GHz cutoff frequency trapped between the tapers. They can be damped by using absorbers placed on the tapers, adding wave guide dampers loaded with ferrite, or by using wave guides with ridges matched to coaxial cables. Similar designs have been used at KEK B-factory and DAFNE Φ-factory. The first method may be difficult to use because the absorber efficiency deteriorates at high frequencies and because the absorbers themselves would be sources of the resistive-wall impedance, outgassing, and dust.

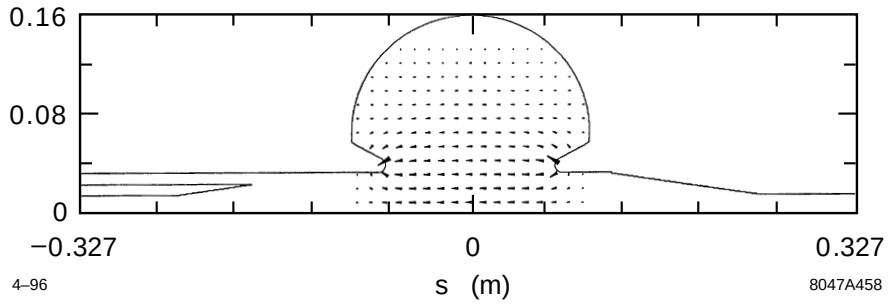


Figure 4-30. *Rf cavity model showing coax damping waveguide in the taper but not showing the damping waveguides on the cavity.*

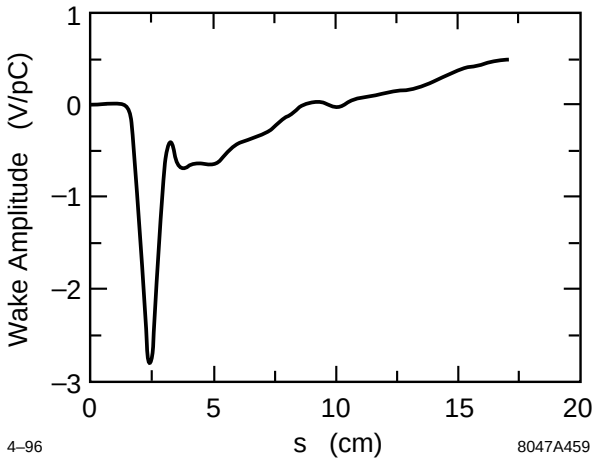


Figure 4-31. *Longitudinal wake for the rf cavity.*

A conceptual design using the wave-guide dampers is shown in Figure 4-30. A cylindrical coaxial wave guide, having a zero cutoff frequency, is connected to a taper. The outer radius of the wave guide is 3.1 cm and the inner radius of the wave guide was chosen to optimize the damping while minimizing the transverse kick to the beam. The wave guide is terminated with absorbers. The effect of the dampers on the Q-factor of the fundamental mode may be reduced by placing a choke on the wave guide.

Preliminary studies using MAFIA demonstrate that the fundamental mode is not affected by the wave guide damper as shown in Figure 4-30. The longitudinal wakeield for this design is shown in Figure 4-31; it is thought that the short-range wake can be further minimized with optimization of the parameters. At this time, we are performing 2- and 3-D MAFIA studies of the cavity to optimize the placement of the waveguide on the taper and those on the cavity body.

The high frequency contribution of the rf cavities to the broad-band impedance, Eq. 4.26, can be estimated comparing the loss factors of the modes with frequencies $\omega < \omega_c$ with the total loss factor of a cavity. The difference, Δk_l,

$$\Delta k_l = \frac{2R_{cav}}{Z_0}\sqrt{\frac{\omega_c}{c\sigma_z}}[\Gamma(1/4) - 4\sqrt{\frac{\omega_c\sigma_z}{c}}] \tag{4.27}$$

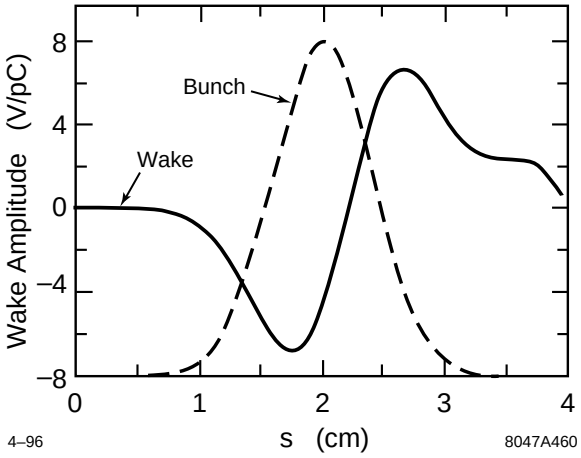


Figure 4-32. *Longitudinal wake of the slot termination.*

defines the coefficient which appears in Eq. 4.26

$$R_{cav} = \frac{Z_0 \Delta k_l}{2\Gamma(1/4)} \sqrt{c\sigma_z} \tag{4.28}$$

for $f_c = 5.34$ GHz. The sum of the NB contribution from the modes $f < f_c$ and the BB tail from $f > f_c$ has a weak dependence on the choice of f_c.

Resistive Wall

The resistive-wall impedance for a round 223-m Al beam pipe is

$$Z_l = (1-i)\, 1.05\sqrt{n}\,\Omega \tag{4.29}$$

$$Z_\perp = (1-i)\frac{0.48}{\sqrt{n}}\frac{M\Omega}{m} \tag{4.30}$$

where n is the harmonic number, $n_{max} = R/\sigma_z = 9000$ for $\sigma_z = 4.0$-mm bunch. The conductivity $\sigma_c = 3.8 \times 10^7$ $\Omega^{-1}m^{-1}$ and the beam pipe radius $b = 1.25$ cm were taken in this estimate. Note, that the skin depth is $\delta = 69.5/\sqrt{n}\,\mu$m.

The loss factor of the resistive wall is

$$k_l = \frac{\Re Z_l}{\sqrt{n}} \frac{2\Gamma(3/4)}{Z_0 \sigma_z} \sqrt{\frac{R}{\sigma_z}} \tag{4.31}$$

and, in this case, is equal to $k_l = 1.4$ V/pC. It should be noted that this estimate does not include the smaller vacuum aperture in the wiggler region where the full gap is only 16 mm. This is partially offset by the larger chamber in the injection/extraction region, and the average loss factor is not changed significantly.

Ante-Chamber Slots and Bellows

To provide adequate pumping with the minimum impedance and to reduce heating and adverse effects of the secondary photon and electron emission, the beam chamber is connected with the ante-chamber by a narrow slot 5-mm wide and

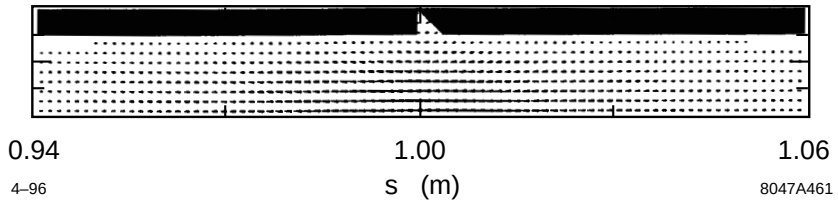


Figure 4-33. *MAFIA model of the bellows rf contact showing the trapped mode.*

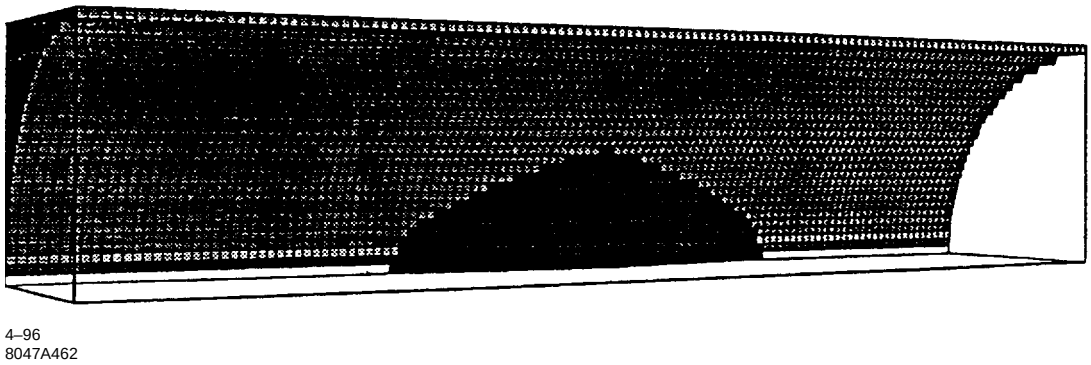


Figure 4-34. *MAFIA model for the bellows mask.*

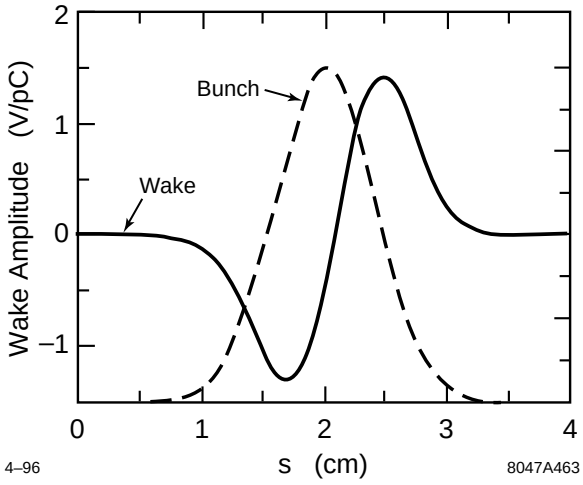


Figure 4-35. *Longitudinal wake for the bellows mask.*

20-mm deep. The slot is terminated with a 25° taper at the masks for the bellows. The impedance of the slot is inductive, see Figure 4-32. The inductance is 0.001 nH for the $\sigma_z = 4$-mm bunch length. The total inductance of 50 slots per ring is 0.05 nH, a very small contribution to the impedance budget.

The impedance of the shielded bellows is given by slots between shield fingers, and by a small cavity of 4-mm length near the ends of the fingers. The contribution of the slots can be estimated analytically and is negligibly small. The 2D MAFIA simulations of the small cavity showed that there is a trapped mode, see Figure 4-33, with frequency 9.099 GHz, which is below the beam pipe cutoff frequency of 9.186 GHz. The mode is localized in the region with length longer than length of the cavity. The trapped mode has a high $Q_0 = 15000$ and $R/Q = 0.23\,\Omega$. The power loss by 360 bunches with 1.5×10^{10} particles per bunch is 6.7 W but, only 25%, *i.e.*, 1.7 W, is deposited within the cavity region. This estimate is conservative because the beam pipe is interrupted by the chamber slot within the localization length of the trapped mode reducing the loaded Q factor of the mode.

Masks

A mask, that is 25-mm long and 2-mm thick, is located at the downstream end of the chamber slot to shield the bellows from the synchrotron radiation. The wakefield of the mask, modeled as shown in Figure 4-34, is depicted in Figure 4-35. The wakefield is mostly inductive but with a substantial resistive component of roughly 5×10^{-3} V/pC. The inductance of a mask is 0.0048 nH which agrees roughly with analytic estimates. The total inductance of the 50 masks and slots is 0.24 nH; the inductance will increase to 1.71 nH if the mask thickness is increased from 2 mm to 4 mm.

BPMs

The design of the BPMs in a high current ring is difficult because of the need to avoid trapped modes that can lead to excessive heating. The BPMs in the NLC damping ring consist of four buttons placed symmetrically around the beam pipe. In the initial design, the buttons have diameters of 8 mm and a ceramic ring of 2 mm thickness for vacuum seal is located at a distance of 9-mm away from the button. Calculations with 3D MAFIA were used to determine the impedance and signal sensitivity of the BPM. The MAFIA model of the BPM is outlined in Figure 4-36.

The total loss parameter of a four-button BPM is found to be $k_l = 2.03 \times 10^{-2}$ V/pC. The longitudinal wakefield of a BPM has a long tail of oscillations at large distances related to several trapped modes. The impedance spectrum is shown in Figure 4-37. Two prominent peaks in the relevant frequency range appear. The field pattern of the resonant mode of high Q at 9.1 GHz corresponds to a mode which is localized mostly in the beam pipe and extends to long distances around a BPM, (Figure 4-38(a)). For 150 BPMs, the total impedance of this mode is 2.1 kΩ, which is small compared with the stability limit of 6.5 kΩ which is calculated from Eq. 4.41 assuming that its frequency is resonant with the revolution frequency (f_b); this is the worst case. The actual frequency of the mode in units of the revolution frequency of 714 MHz, $f_r/f_b = 12.7$, is far from an integer value. Hence, the coupled-bunch effects of this mode will be small. Additionally, due to large spatial extent, the heating, produced by the mode at the BPM button, is small. Another resonance at 12.0 GHz is the TE_{11} is localized inside of a button, see Figure 4-38c. The total impedance for all the BPMs for this mode is 2.4 kΩ, which is small compared with the stability limit of 7.5 kΩ. The Q value of this resonance is about 50, and, hence, the heating effect arising from coupled bunches is negligible. Furthermore, there is a small peak at around 9.8 GHz between these two resonances. It is identified as the ceramic mode (Figure 4-38b). Since the ceramic is situated far away from the button, this mode is not strongly excited by the beam.

Thus, in the present design, there are a number of narrow-band impedances which appear to be below the coupled bunch stability limits. Detailed heating effects in the BPM need to be addressed further. Finally, the signal at the coax-

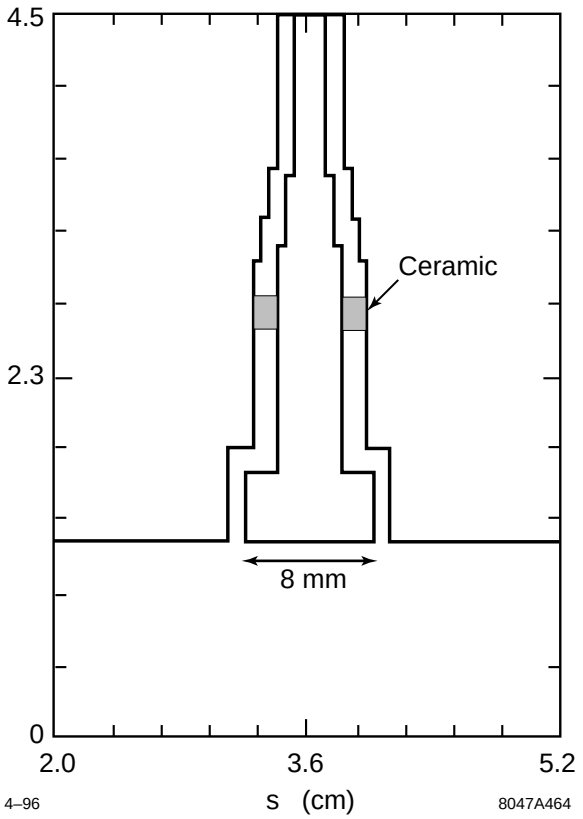


Figure 4-36. *2D projection of 3D MAFIA model for the BPMs.*

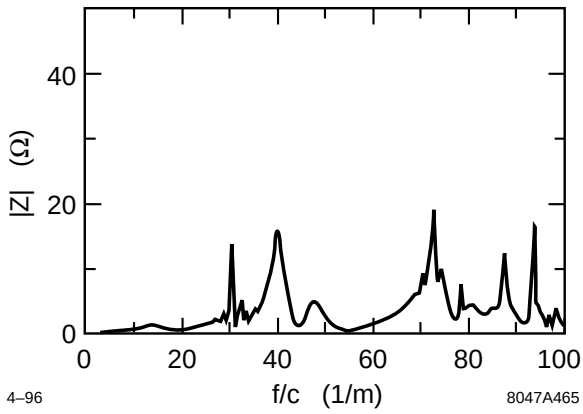


Figure 4-37. *Longitudinal impedance of a four button BPM.*

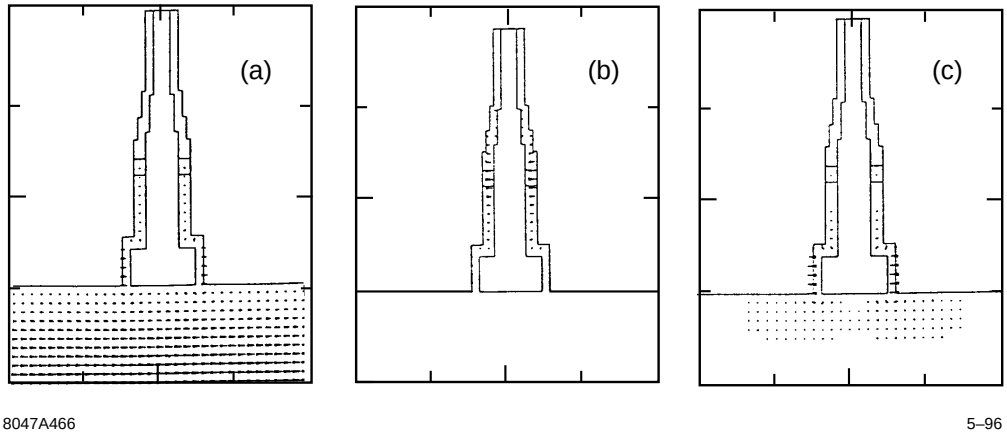


Figure 4-38. *Trapped modes in the BPMs.*

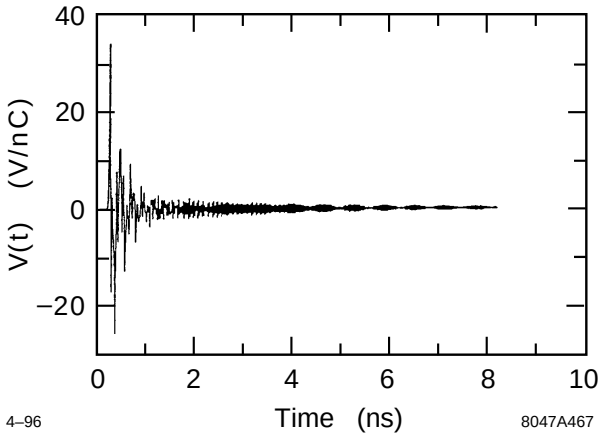


Figure 4-39. *BPM response to a passing bunch.*

ial cable as a function of time is shown in Figure 4-39 and the transfer impedance for the signal is shown in Figure 4-40. Further investigation is needed to determine whether the output signal satisfies the resolution requirements.

Summary

The contributions to longitudinal loss factor are listed in Table 4-17. The main contribution comes from two rf cavities, BPMs, and resistive walls. The other components of the ring are mostly inductive but, for short bunches, still have noticeable loss factors. The main source of the inductance is the synchrotron radiation masks, 50 of them give $L = 0.24$ nH. Some components, such as feedback, injection and extraction kickers, diagnostics, dampers, etc. are not yet included. We expect that they will not substantially change the total (high frequency) impedance.

The transverse impedance is dominated by the HOMs of the cavities and the resistive wall.

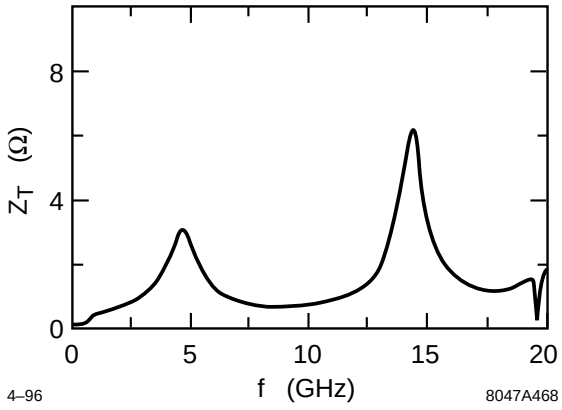


Figure 4-40. *Transfer impedance for the BPMs.*

	number	k_l V/pC
RF cavities	2	3.45
Res. Wall		1.44
Ante-chamber		0.0225
Bellows	50	0.365
4-button BPMs	150	3.04
Masks	50	0.25
Total:		8.5

Table 4-17. *Loss factor summary for the main damping rings.*

The total HOM power can be estimated from the total loss factor k_l as

$$P = I_{beam}^2 \frac{Z_0 k_l s_b}{4\pi} \tag{4.32}$$

which gives a rough estimate that does not take into account the real beam spectrum and an enhancement for the modes at frequencies equal to multiples of the bunch spacing frequency.

This yield $P = 15.4$ kW for $k_l = 8.5$ V/pC, maximum beam current $I_{beam} = 1.2$ A, and the $s_B/c = 1.4$-ns bunch spacing; the HOM power propagating in the beam pipe is smaller due to the narrow-band modes trapped in the rf cavities and BPMs. This is not expected to present a limitation but needs further investigation.

4.4.2 Single-Bunch Potential Well Distortion and Longitudinal Microwave Instability

The two important questions concerning the single bunch longitudinal phase space are (a) What is the average bunch shape and energy spread in the ring? and (b) What is the threshold current to the microwave instability? The answer to the first question is needed in order to know the longitudinal phase space properties of the bunch in subsequent parts of the collider. As for the second question, it is known that the microwave instability is accompanied by an increase in beam energy spread (which increases the longitudinal emittance) and, more importantly, that it can also be accompanied by unpredictable transient beam behavior (*e.g.*, the "saw-tooth" variation in bunch length that has been

observed in the SLC damping rings [Krejcik 1993]), a phenomenon that can greatly degrade the performance of the linear collider as a whole.

The microwave instability is also known to come in at least two forms: One, which we call the "strong" type of instability, has a high growth rate that depends only weakly on the synchrotron radiation damping rate, a threshold that can be roughly approximated by the Boussard criterion [Boussard 1975], and that can be thought of as generated by the coupling of two synchrotron modes with different azimuthal mode numbers. The other, a type that has only recently been identified [Oide 1994, Chao 1995, Bane 1995a], we call the "weak" type of instability; it has a growth rate that is sensitive to the radiation damping time, has a threshold that has no connection to the Boussard criterion, and can be described as the coupling of two modes with the same azimuthal but different radial mode numbers. The weak type of instability is found in storage rings with a predominantly resistive impedance.

In this section, we find that for the NLC damping rings at our highest design current ($N = 1.57 \times 10^{10}$) we may be running near the threshold of the weak instability but we are significantly below that of the strong instability. However, the weak instability is not a serious problem like the strong instability.

Consider, the SLC damping rings. In 1994 a completely redesigned vacuum chamber was installed. Before this time a strong type of instability was observed, with the current—due to the microwave instability—limited to the threshold value [Krejcik 1993]; since the upgrade only a weak type of instability has been seen, and the SLC runs routinely a factor of 3 above the threshold current [Bane 1995b]. We also believe the weak instability is less serious in that, unlike the strong instability, it is very sensitive to the tune spread within the bunch; by slightly increasing the tune spread (such as, for example, by adding a modestly powered, higher harmonic cavity to the ring [Chao 1995]), through Landau damping, the threshold to this kind of instability can be increased substantially.

In this section, we first obtain the wakefields of the main vacuum chamber objects, driven by a bunch that is significantly shorter than the nominal bunch length, using the computer program MAFIA [Klatt 1986]. These individual wake functions are then summed in the correct proportion to give a wakefield representing the entire ring. Finally, this wakefield is used as a Green function for calculating the effect of potential well distortion and for obtaining the threshold to the microwave instability. To obtain the bunch shape below threshold this wakefield is applied to the Ha¨ssinski formula [Ha¨ssinksi 1973]; for locating the threshold current it is applied to the Vlasov equation, which is solved in a perturbative manner using K. Oide's computer program [Oide 1990].

Calculations

One cannot predict the average bunch properties nor the threshold current of a ring without knowing the vacuum chamber geometry, from which they can calculate the wakefield or the impedance. As was the case for the SLC damping rings small changes in vacuum chamber cross-section—such as those connected to bellows, transitions, and masks—can dominate in their contribution to bunch lengthening and the microwave instability over larger objects such as the rf cavities [Bane 1988]. Therefore, it is important to know the details of the geometry of the vacuum chamber before beginning this work (as it is important to build the vacuum chamber to tight specifications during the construction phase). For the NLC damping rings, the main contributors to the longitudinal impedance have been designed: they are the 2 rf cavities, the 150 beam position monitors (BPMs), the longitudinal resistive-wall wakefield, the 50 synchrotron radiation masks, and the 50 tapers and pumping slots. It is assumed that the bellows that allow the ring to be assembled are shielded, leaving an impedance that is insignificant. Furthermore, it is assumed that the kicker chamber and septum region also do not contribute significantly to the impedance (Note that the important objects are all described in detail in Section 4.4.1).

The nominal rms bunch length in the NLC damping ring is 4 mm; for all our wakefield calculations we take a gaussian bunch with rms $\sigma_z = 1$ mm as the driving bunch and calculate the wakefield to 15 mm behind the bunch center (due to the large number of mesh points required for finding the wake of the rf cavities and the BPMs we could not easily

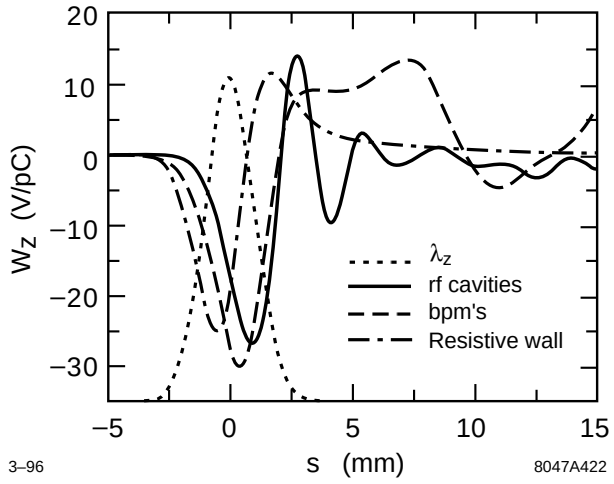


Figure 4-41. *The longitudinal wakefield left by a gaussian bunch, with $\sigma_z = 1$ mm, due to the rf cavities (the solid curve), the BPMs (the dashed curve), and the resistive-wall wakefield (the dot-dashes). The driving bunch shape is given by the dotted curve, with the head to the left.*

use a shorter driving bunch). We begin with the rf cavities. There are two single–cell cavities; each cavity is set back from the beam tube by a pair of gradual tapers. The total wakefield of the two rf cavities, as calculated by MAFIA, is shown in Figure 4-41 (the solid curve). We see contributions of both the setbacks and the cavity itself in the shape of the wake. The BPMs are of the button type; each BPM consists of four cylindrical, 8-mm-diameter electrodes that are set symmetrically (about the axis) into the beam tube wall; the impedance is due to the annular space that electrically isolates each button from the tube wall. The wakefield of the 150 BPMs is given by the dashed curve in Figure 4-41. We notice ringing at a frequency beyond 20 GHz.

The longitudinal resistive-wall wakefield can normally be ignored in impedance calculations of storage rings; for the NLC, however, its contribution is significant. For a gaussian bunch the resistive-wall wakefield is given by [Piwinski 1972]

$$W_z(s) = \frac{C}{4b\sigma_z^{3/2}}\sqrt{\frac{c}{2\pi\sigma}}f(s/\sigma_z) \quad , \tag{4.33}$$

with C the machine circumference, b the tube radius, σ_z the rms bunch length, σ the conductivity of the vacuum chamber walls; with

$$f(u) = |u|^{3/2}e^{-u^2/4}(I_{1/4} - I_{-3/4} \mp I_{-1/4} \pm I_{3/4})|_{u^2/4} \quad . \tag{4.34}$$

In Eq. 4.34 the upper signs apply for $u < 0$, the lower signs for $u > 0$; the modified Bessel functions I are evaluated at $u^2/4$. For the NLC damping ring we take $C = 223.5$ m, $b = 1.25$ cm, $\sigma_z = 1$ mm, and $\sigma = 3.2 \times 10^{17}$ s^{-1} (aluminum). The resulting resistive-wall wakefield is given by the dot-dashed curve in Figure 4-41. We note that, in amplitude, the cavities, the BPMs, and the resistive-wall contribute about equally to the total wakefield of a 1-mm bunch. As final ingredients we have also obtained the wakefields of the 50 synchrotron radiation masks and the 50 tapers and pumping slots; these contributions are small compared to the wakefields discussed above.

Summing all the contributions we obtain the wakefield that we will use as a Green function to represent the entire NLC damping ring (see Figure 4-42). One important subtlety is that, since it has a leading tail, our Green function is not causal. To obtain a causal Green function we will consider two ways of modifying the leading tail. The first method, which we call the capacitive approximation, is obtained by flipping the leading tail to the back, about $s = 0$, and adding it to the wake. The second method, which we call the resistive approximation, is obtained by replacing the leading tail by a delta function at the origin that has the same area as the leading tail. In the following we will

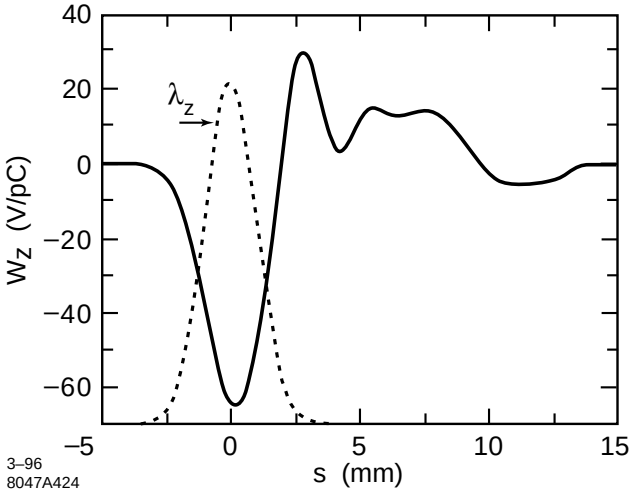


Figure 4-42. *The total longitudinal wakefield left by a gaussian bunch, with $\sigma_z = 1$ mm, on each turn in the NLC damping ring. The driving bunch shape is given by the dotted curve, with the head to the left.*

compare the results using these two versions of the Green function. If the bunch used to calculate the Green function is sufficiently short, then the results for the two cases should be nearly the same.

To obtain the longitudinal bunch distribution λ_z we apply the total wake function to the Ha¨ssinski equation [Ha¨ssinksi 1973]

$$\lambda_z(s) = K \exp\left(-\frac{s^2}{2\sigma_{z0}^2} + \frac{1}{V'_{rf}\sigma_{z0}^2}\int_0^s V_{ind}(s')\,ds'\right) \quad , \tag{4.35}$$

with K a constant that gives $\int_{-\infty}^{\infty}\lambda_z(s)\,ds = 1$, σ_{z0} the nominal (zero current) bunch length, V'_{rf} the slope of the rf voltage at the synchronous phase, and

$$V_{ind}(s) = eN\int_0^{\infty}\lambda_z(s-s')W_z(s')\,ds' \quad , \tag{4.36}$$

with eN the total bunch charge. For $N = 1.57\times10^{10}$, $\sigma_{z0} = 4$ mm, $V_{rf} = 1$ MV, rf frequency $f_{rf} = 714$ MHz, and synchronous phase (from the peak) $\phi = 50°$, we obtain the bunch shape and induced voltage shown in Figure 4-43 (the head of the bunch is to the left). [The result is essentially the same for both the capacitive and the resistive approximation to the Green function.] We note that the impedance is clearly resistive in character, *i.e.*, V_{ind} is roughly proportional to λ_z, and that the bunch shape is hardly perturbed from the nominal, gaussian one. The rms bunch length is increased by 6% to 4.23 mm and the centroid is shifted forward by 1.46 mm, to make up for the higher mode losses.

To estimate the threshold for the strong microwave instability we use the Boussard criterion [Boussard 1975]:

$$\frac{e\hat{I}|Z/n|}{2\pi\alpha_p E\sigma_\delta^2} \lesssim 1 \quad , \tag{4.37}$$

with $\hat{I}$ the peak current; $|Z/n| = |Z|\omega_0/\omega$ where Z is the impedance at frequency ω and ω_0 is the revolution frequency (for an inductive impedance $|Z/n|$ is a constant); α_p is the momentum compaction, E the beam energy and σ_δ the rms relative energy spread. The above criterion, if satisfied, indicates that the beam is stable. Note, however, that the Boussard criterion is only a very rough criterion, used to obtain an approximate idea of where the threshold to the strong instability might be. For the NLC, since the impedance is resistive it can be characterized by a resistance R, with

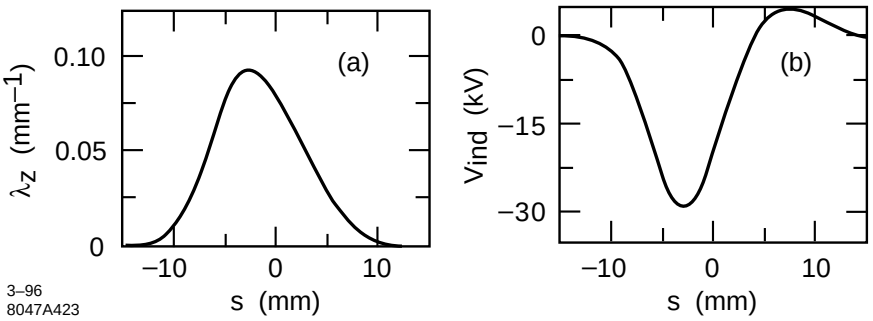


Figure 4-43. *The NLC damping ring bunch shape (a) and the induced voltage (b) when $N = 1.57 \times 10^{10}$, $V_{rf} = 1$ MV, $\phi = 50^{\circ}$. The head of the bunch is to the left.*

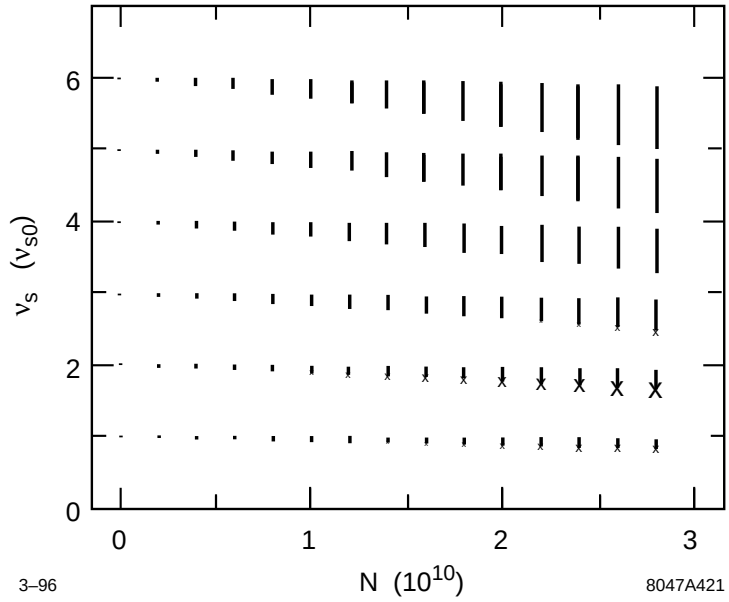


Figure 4-44. *The modes obtained by the Vlasov method when the capacitive model to Green function is used. Stable modes are represented by a dot, unstable modes by an `X', with the size proportional to the growth rate.*

the induced voltage given by $V_{ind} = -eNcR\lambda_z$. Fitting to this equation we find that the effective resistance of the damping ring $R \approx 250\,\Omega$. Let us take as typical bunch frequency $\omega \sim c/\sigma_z$, which gives $|Z/n| \approx R\sigma_z\omega_0/c = 0.03\,\Omega$. Then, using the same parameters as above, and taking $\alpha_p = 4.7 \times 10^{-4}$, $E = 2\,\text{GeV}$, and $\sigma_\delta = 9 \times 10^{-4}$, we find that the left hand side of Eq. 4.37 equals 0.5, suggesting that, at the highest envisioned current, we are still below the threshold for a strong instability.

To get a better estimate of the instability threshold, we have applied our wakefield to K. Oide's computer program that solves the linearized, time-independent Vlasov equation [Oide 1990]. This program uses a perturbation approach to find the threshold; it includes the potential well distortion as a zeroth order effect. This program solves for the modes of the system, and finds the mode that first becomes unstable. (For the SLC damping ring, the results of this program have been shown to agree very well with turn-by-turn tracking using a large number of macro-particles [Bane 1995a, Bane 1993]). For our simulation, we take as nominal synchrotron frequency $\nu_{s0} = 0.004$, and truncate the system matrix to six mode numbers in the azimuthal direction and 60 in the radial direction.

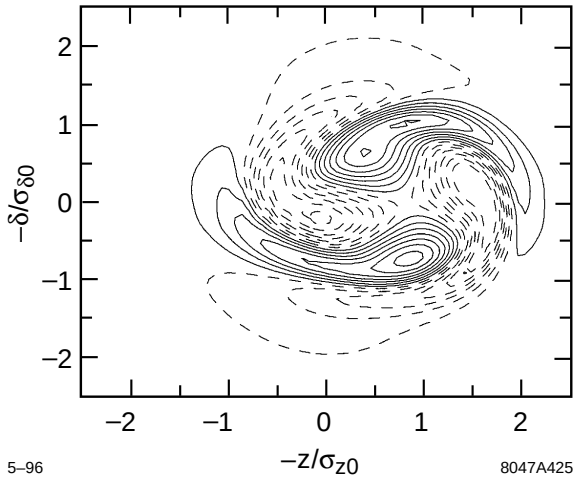


Figure 4-45. *The shape of the weak mode found at $N = 2 \times 10^{10}$ when the capacitive approximation to the Green function is used.*

The results are shown in Figure 4-44 where we plot the dependence of mode frequency on current for the capacitive approximation to the Green function. Stable modes are represented by a dot (note that at $N = 0$ for every azimuthal mode number there are 60 dots with the same frequency); unstable ones are indicated by an 'X', with the size of the symbol proportional to the growth rate (synchrotron radiation damping is not included in the calculation). We see that for each azimuthal mode number, at least up to the nominal maximum current ($N = 1.57 \times 10^{10}$), the frequencies are spread and depressed only slightly, which is typical behavior for a resistive impedance. The synchrotron radiation damping time of the NLC damping ring is $t_z = 2.5$ ms. The growth rate of an unstable mode first matches this value at $N = 1.65 \times 10^{10}$. The unstable mode is a quadrupole mode of the weak type, with frequency $1.8\nu_{s0}$ (its shape is plotted in Figure 4-45). In this example, the nearest strong instability is at $N = 7 \times 10^{10}$.

If, instead, we take the resistive approximation to the Green function we find that the growth rate of the weak instability is below the synchrotron radiation damping all the way up to the threshold of the strong instability, which, in this case, we find at $N = 8 \times 10^{10}$. The true result for our model damping ring is probably somewhere between the results given by the capacitive and the resistive model to the Green function, which suggests that at the highest design current no microwave instability will be encountered in the NLC damping rings. To obtain a more precise result of the location of the instability thresholds we need to get the wakefields for the NLC damping ring vacuum chamber due to a bunch with an rms that is significantly shorter than the current value of 1 mm.

4.4.3 Single-Bunch Transverse Mode Coupling

Transverse coupling will set a limit on the maximum single bunch beam current. In addition, the mode coupling threshold will be modified by the coupling between the bunches [Berg 1995]. In this section, we calculate the single bunch mode coupling threshold and then describe the modification of the results due to the bunch trains; the coupled-bunch instabilities are described in a later section.

In the NLC damping ring, the vertical plane results are universally worse than those for the horizontal plane, so only the vertical plane will be discussed here. The details of these calculations is given in Ref. [Berg 1995].

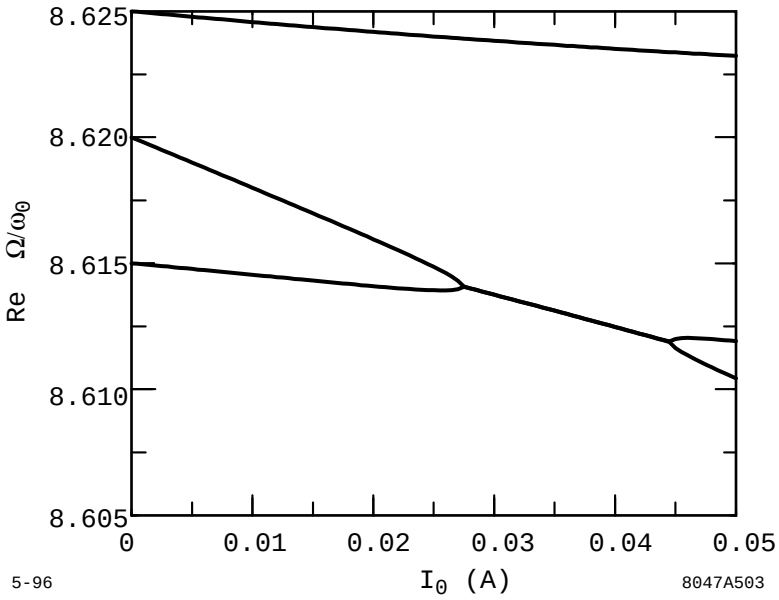


Figure 4-46. *Single-bunch vertical mode coupling.*

We use the broadband model of the transverse impedance and the transverse cavity higher-order modes given in Section 4.4.1. Note that the cavity higher-order modes as well as the part of the broadband impedance due to cavity tails must be enhanced by a factor of the β-function at the cavities (taken to be 6.5 m) over the average β-function (4-m vertical).

If only a single bunch is considered, transverse-mode coupling is seen at around 28 mA, as shown in Figure 4-46; the threshold is almost identical whether the rf cavities are de$Q$ed to 100 or 3000. For comparison, the damping ring single bunch current is much less, roughly 3.3 mA.

If we consider 532 symmetrically placed bunches, we find that the rigid ($m = 0$) coupled-bunch motion is only very slightly affected by mode coupling at the typical currents but the $m = 1$ coupled-bunch growth rates are strongly affected. If the rf cavities are de$Q$ed to around 100, the resulting growth rates are well below radiation damping at a beam current of 1.2 A. If the some cavity modes are left with Qs around 3000, the growth rates can become comparable to to the radiation damping rate. Additional details will be discussed in the section on transverse coupled-bunch instabilities.

4.4.4 Bunch-to-Bunch Synchronous Phase Variation

At any given time, there will be four trains of 90 bunches in the main damping rings. Within a train, the bunches populate adjacent buckets and between trains there is a gap that extends over 43 buckets. A consequence of an uneven filling scheme is that within each train the synchronous phase will vary from bunch to bunch. After extraction the beam enters the bunch compressors and then the X-band linac. The phase variation in the ring, if uncompensated, will lead to a phase variation in the X-band linac which, in turn, will result in an unacceptable spread in the final energy of the bunches in a train. The synchronous phase variation, however, can be compensated, either in the damping ring itself or in the bunch compressors that are described in Chapter 5. The present NLC design assumes that later method of

compensation and the bunch compressors have been designed to accommodate a ± 20 ps variation in the synchronous phase.

In this section, we first find the synchronous phase variation across the bunch trains of the NLC damping rings, and then we study a method of compensating this variation—one employing passive sub-harmonic cavities—in the damping ring itself. More details of this compensation technique can be found in Ref. [Bane 1995c].

The Variation in Synchronous Phase

Consider a damping ring filled with bunch trains, with N bunches per train, and let us suppose that we can ignore the effects of higher cavity modes. Let us further assume that the bunch gap is short, *i.e.*, $|\Delta\bar{\omega} T_g| \ll 1$, with T_g the gap time interval, and

$$\Delta\bar{\omega} = \Delta\omega + i\frac{\omega_{rf}}{2Q_L} \quad , \tag{4.38}$$

where $\Delta\omega$ is the frequency detuning, ω_{rf} the fundamental frequency, and Q_L the loaded Q of the cavity. Then the steady-state phase difference between the first and last (N^{th}) bunch in each train is given by [Lambertson 1995]

$$\Delta\phi_{1N} \approx -\frac{2kI_0T_g}{V_c \sin\phi} \quad , \tag{4.39}$$

with k the loss factor ($= \frac{1}{2}\omega_{rf} R/Q$), I_0 the average beam current, V_c the peak voltage, and ϕ the nominal synchronous phase ($\phi = \cos^{-1}[U_0/eV_c]$, where U_0 is the energy loss per turn). The average beam current I_0 is given by

$$I_0 = \frac{Nq}{NT_b + T_g} \quad , \tag{4.40}$$

with q the charge per bunch and T_b the bunch-to-bunch time interval within a train. The negative sign in Eq. 4.39 indicates that the synchronous phase of the last bunch is more forward (toward the crest) than the first bunch. Note also that in the NLC damping rings not only is the gap short but also the bunch train itself, *i.e.*, $|\Delta\bar{\omega} N T_b| \ll 2\pi$; it can easily be shown [Bane 1995c] that in such a case the synchronous phase varies linearly with bunch number within each bunch train.

For the NLC damping rings the rf frequency is 714 MHz and the R/Q for the two rf cavities combined is 240 Ω; therefore, $k = 0.54$ V/pC. We take $N = 90$, $q = 2.5$ nC, $T_b = 1.4$ ns, and $T_g = 60.2$ ns; therefore, $I_0 = 1.2$ A. Taking $V_c = 1$ MV and $\phi = 50°$ we obtain $\Delta\phi_{1N} = -5.8°$. We have also numerically simulated this problem using turn-by-turn tracking [Thompson 1991]. The unloaded Q of the cavities was taken to be $Q_0 = 25500$; for minimum reflected power $Q_L = 2200$ and $\psi = -44.5°$ (remember: $\tan\psi = 2Q_L\Delta\omega/\omega_{rf}$). The results are shown in Figure 4-47. Plotted are the shift in synchronous phase $\Delta\phi$ versus bunch number. We see that, over each train, the variation is linear with bunch number, and that the total phase shift is $-5.9°$, in good agreement with our analytical results. Note that both the length of the gap and the length of the bunch train are short according to our criteria: $|\Delta\bar{\omega} T_g| = 0.003 \ll 1$ and $|\Delta\bar{\omega} N T_b| = 0.2 \ll 2\pi$.

It may occasionally happen that one bunch train has a slightly different charge than the others. To study the sensitivity to this we have repeated the numerical tracking but with the bunches of the first train having only 95% of the nominal charge. The results are shown in Figure 4-48. In this case the maximum phase deviation in the first train has been reduced to 80%, in the other three trains to 95%, of the value found for equally populated trains. Note, however, that the average value of $\Delta\phi$ for all trains has changed by 3°. This average change, however, can easily be corrected by a phase feedback system in the damping ring.

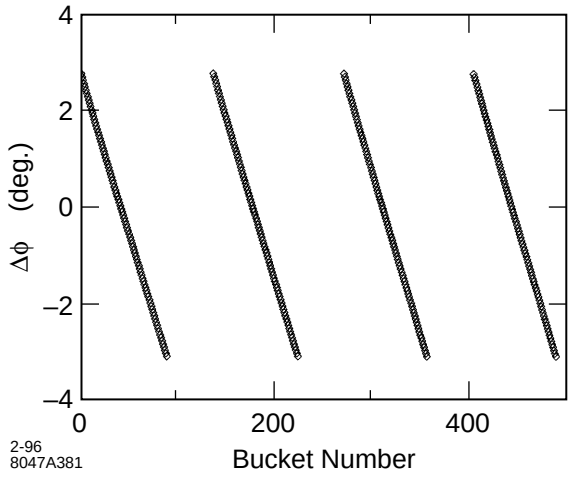


Figure 4-47. *The steady-state, bunch-to-bunch synchronous phase variation vs. bunch number for the NLC damping rings. Note that a more negative value of phase is more toward the front (toward the rf crest).*

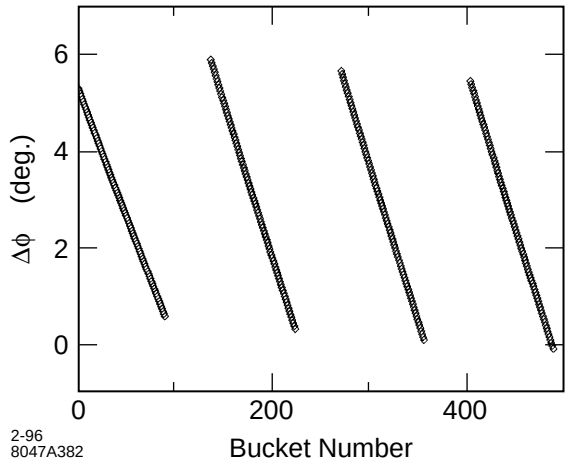


Figure 4-48. *The synchronous phase variation vs bunch number for the case when the bunches of the first train have only 95% of the nominal bunch population while the rest are unchanged.*

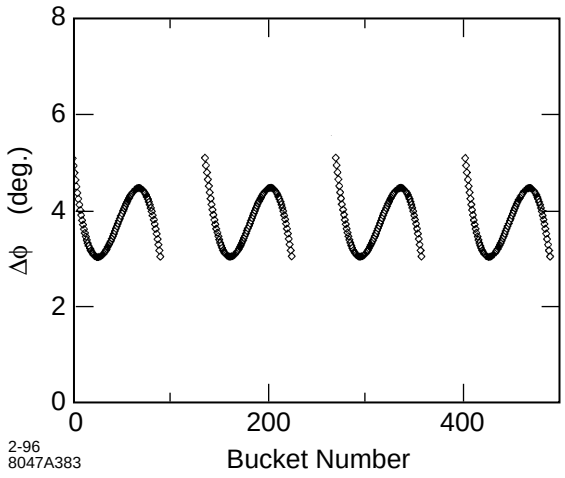


Figure 4-49. *The synchronous phase variation that can be achieved when one passive, lower harmonic cavity is used.*

Compensating the Variation in Synchronous Phase

Given a sufficient amount of extra power and bandwidth in the klystron, the synchronous phase variation discussed above can be compensated by properly varying the generator voltage as a function of time; this method will be studied in future work. However, here we will discuss a method of compensation described in Ref. [Kubo 1993], where the authors find that a properly tuned, passive, lower harmonic cavity can be used to compensate the phase variation due to bunch trains. The frequency of the compensation cavity must be equal to $\omega_{rf} - 2\pi n N_t/T_0$, with n an integer, N_t the number of bunch trains, and T_0 the revolution time. This method of compensation works quite well when the gap length is larger than the length of a bunch train. In our case, however, the gap length is only about half the train length. We will, therefore, also investigate the effect that can be obtained using two lower-harmonic cavities to compensate the synchronous phase variation.

Figure 4-49 gives the tracking results when an optimally-tuned, passive cavity with frequency $\omega_{rf} - 8\pi/T_0$ is included in the damping ring. The parameters of this cavity were optimized numerically. They are: $R/Q = 34\ \Omega$, $Q_0 = 3.43 \times 10^4$, $Q_L = 1.14 \times 10^4$, and $\psi = -85°$. We see from the figure that the maximum phase deviation has been reduced to $2.1°$, which is only a factor of 3 smaller than we had initially. Adding a second, passive cavity, with frequency $\omega_{rf} - 16\pi/T_0$, $R/Q = 18\ \Omega$, $Q_0 = 3.43 \times 10^4$, $Q_L = 1.0 \times 10^4$ and $\psi = -82°$ gives the results shown in Figure 4-50. Now the maximum phase deviation has been reduced to $0.65°$, which is a factor of nine smaller than we started with. Finally, note that the R/Qs of the compensation cavities are very small, so they should not contribute to an instability. One negative point, however, might be that the residual phase variation is nonlinear with bunch number, an effect which may have its own negative ramifications.

At this time, we have decided not to compensate the loading in the rings. The $\pm 3°$ L-band variation across the bunch train can be easily accommodated in the bunch compressors as described in Chapter 5. Furthermore, because the variation is linear across the bunch train, it may be straightforward to correct the resulting energy variation in the prelinac before the second compressor. This would allow a simpler design for the second compressor system.

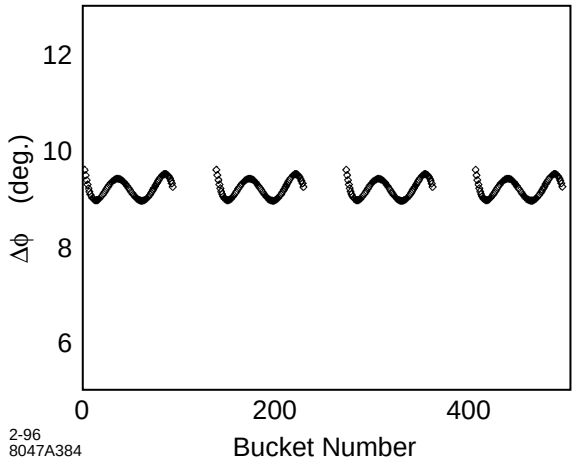


Figure 4-50. *The synchronous phase variation that can be achieved when two passive, lower harmonic cavities are used.*

4.4.5 Coupled Bunch Instabilities

The close spacing of bunches (1.4 ns) within a train in the main damping rings allows for the possibility that bunch wakefields will influence the motion of consecutive bunches. In general, such wakefields arise from narrow-band resonances in the rf cavities and from the finite conductivity of the vacuum chamber wall. Under certain conditions, the wakefields can generate a transverse or longitudinal coherent structure in the relative motion of the bunches which grows exponentially, leading to a loss of beam quality and/or loss of the beam itself. This is referred to as a coupled-bunch instability. The wakefields can also create static offsets in either the transverse or longitudinal position along the bunch train due to the nonuniform filling pattern of the ring. Furthermore, the restraints on beam size and bunch phase at extraction require control of beam oscillation transients from injection errors. These effects are generally proportional to the total beam current and thus are a concern for the 1.2-A design current for the main damping rings.

At the time of this study, the HOMs of the damping ring cavities were not available. Thus, we have estimated these various effects for the main damping rings assuming a scaled-down version of the PEP-II rf cavity as the primary source of narrow-band resonances, where the cavity size has been reduced such that the fundamental mode frequency has increased from 476 to 714 MHz. The strongest monopole and dipole higher-order-modes (HOMs) are given in Tables 4-18 and 4-19. It is also assumed that all HOMs are damped to $Q < 200$. The effects are calculated for two identical rf cells with a total voltage of 1 MV.

Longitudinal

In the longitudinal plane, coupled-bunch oscillations are driven only by the monopole HOMs of the rf cavities. Assuming a uniformly filled ring, the growth rate for dipole oscillations is given by

$$\frac{1}{\tau} = \frac{1}{2}\frac{I_0 \alpha_p f_{rf}}{(E/e)Q_s}\left[Z_{||}\right]_{\text{eff}} \tag{4.41}$$

where the effective impedance is

$$\left[Z_{\|}\right]^{l}_{\text{eff}} = \sum_{p=-\infty}^{p=+\infty} \frac{\omega_p}{\omega_{rf}} e^{-(\omega_p \sigma_\tau)^2} Z_{\|}(\omega_p) \ . \tag{4.42}$$

f_r (MHz)	R_s (kΩ)
3060.85	2.6
1941.55	1.8
2291.94	1.8
4563.48	0.66
1487.27	0.52
4150.84	0.52
3368.43	0.46
3554.05	0.36
5180.54	0.32

Table 4-18. *Strongest rf cavity monopole HOMs.*

f_r (MHz)	$R_\perp$ (kΩ/ m)
1193.2	85
2499.8	72
1988.2	36
2397.2	11.2
2310.6	9.7
4861.2	7.8
3755.4	7.0
3603.0	4.8
3423.6	3.8
4736.2	3.8
3346.5	2.6

Table 4-19. *Strongest rf cavity dipole HOMs.*

The effective impedance and the corresponding growth rate are shown in Figure 4-51. For $Q < 200$ the fastest growth rate falls approximately a factor of two below the radiation damping rate, indicating stable oscillations. The above applies strictly to a uniform fill pattern with every rf bucket filled. The effects of instabilities are typically lessened by nonuniform fill patterns. More detailed simulations are required to study transient effects including the gaps between bunch trains.

The growth rates are insensitive to tuning of the HOM frequencies because of the relatively large bandwidth of the HOMs as compared to the revolution frequency. Landau damping from a tune spread within the bunch is not effective because of the short bunch length. Coupled–bunch shape oscillations, *i.e.*, $m > 0$, are not driven by the HOMs at a significant level. Also, with the nominal detuning, the fundamental mode impedance does not drive any coupled–bunch instabilities at a rate comparable to the radiation damping rate.

Transverse

Coupled-bunch oscillations in the transverse planes are driven by dipole rf cavity HOMs and by the resistive-wall impedance. The HOMs given in Table 4-19 are assumed to be equally strong in the horizontal and vertical planes. The resistive-wall impedance results from the finite conductivity of the vacuum chamber wall and for a round pipe of

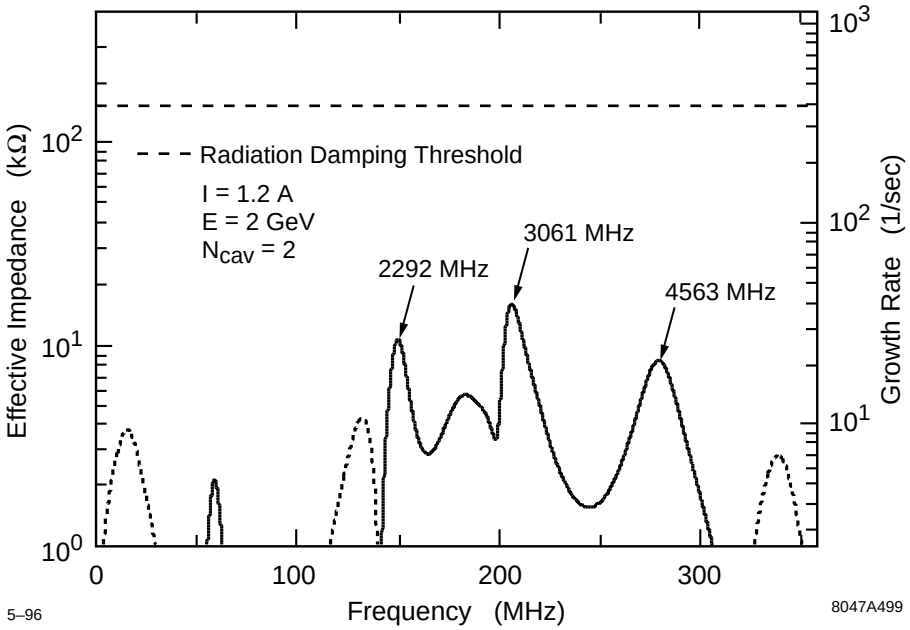


Figure 4-51. *Longitudinal coupled-bunch growth rates.*

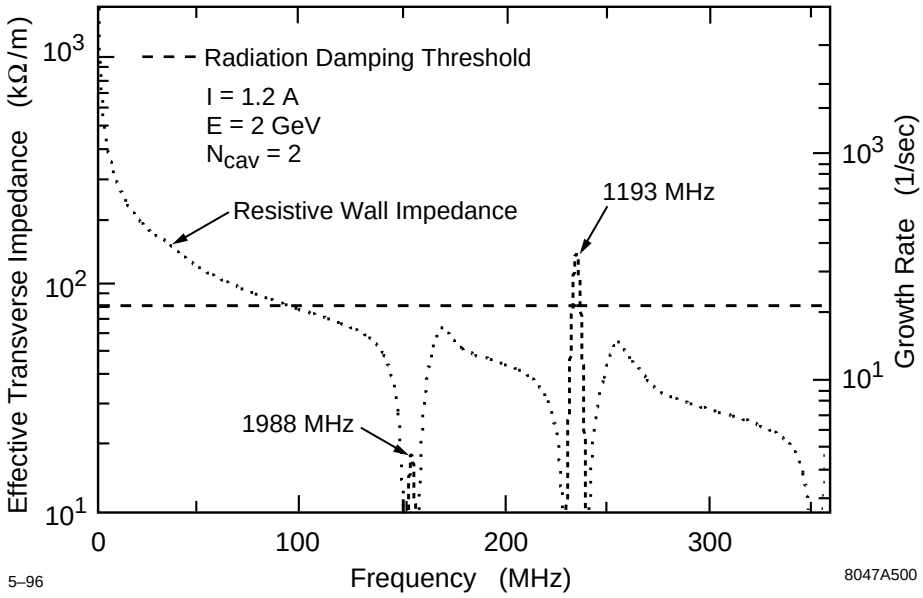


Figure 4-52. *Transverse coupled-bunch growth rates.*

radius b and resistivity ρ is given by

$$Z_{\perp}(\omega) = (1+j)\frac{cL}{\pi b^3}\sqrt{\frac{\rho\mu_0}{2}}\frac{1}{\sqrt{\omega}} \tag{4.43}$$

For a 1.25-cm-radius Al beam pipe, the net resistive-wall impedance is given as

$$Z_{\perp,RW} = (1-j)\frac{0.48}{\sqrt{n}}\ M\Omega/m \tag{4.44}$$

where $n = \omega/\omega_0$.

Assuming a uniformly filled ring, the coupled-bunch growth rate for dipole oscillations of the beam is given by

$$\frac{1}{\tau} = \frac{1}{2}\frac{I_0 f_0}{(E/e)}\beta_{\perp} j[Z_{\perp}]_{\text{eff}} \tag{4.45}$$

where the effective impedance is given by

$$[Z_{\perp}]_{\text{eff}} = \sum_{p=-\infty}^{p=+\infty} e^{-(\omega_p\sigma_\tau)^2} Z_{\perp}(\omega_p) \tag{4.46}$$

The effective impedance and the corresponding growth rates are shown in Figure 4-52. The resistive-wall impedance dominates mainly because of the small vacuum chamber. The large growth rate in excess of radiation damping will require a feedback system to damp transverse beam oscillations; this is described in Section 4.7. Higher-order head–tail modes are not driven at a significant level by the HOMs for $Q < 3000$.

There is a transient beam loading of the transverse impedance analogous to that of the longitudinal which could possibly result in a variable transverse offset along the length of the bunch train. This effect has not yet been studied in detail but it not expected to be significant.

4.4.6 Ion Trapping

The cross section for ionization of the residual gas is larger than that of elastic scattering or bremsstrahlung (compare Section 4.4.9). For carbon monoxide CO and a beam energy of 50 GeV, it is about 2 Mbarn [Rieke 1972]. The main effect of the residual gas on the beam is thus due to the creation of ions. The ions are produced at a rate of about $2\ p$[nTorr] s^{-1} or $6.5\ p$[Torr] m^{-1} ions per electron. The large bunch gap of 60 ns prevents ions from being trapped between bunch trains. Computer simulations may need to be performed to determine the fraction of ions which remains trapped during the gap but this is expected to be extremely small $\lesssim 10^{-3}$.

At the end of a single bunch train the longitudinal density of ions created by the train is 9×10^4 m^{-1} at a pressure of 10^{-8} Torr, assuming 90 bunches of 1.5×10^{10} electrons each. Particles in the trailing bunches then suffer a vertical tune shift of $\Delta Q_y \approx 0.02$. To limit this effect, we are designing the main damping ring vacuum system to achieve a pressure less than 5×10^{-9} Torr.

It has been predicted that the interaction of beam and ions also gives rise to a fast transverse beam-ion instability [Raubenheimer 1995a]. It is because of this instability that the present target for the Damping Rings is a vacuum level as low as 10^{-9} Torr. For an electron beam, the instability mechanism is the following. Ions created by the first bunches perturb the tail of the bunch train. The mutually driven transverse oscillations of beam and ions result in a

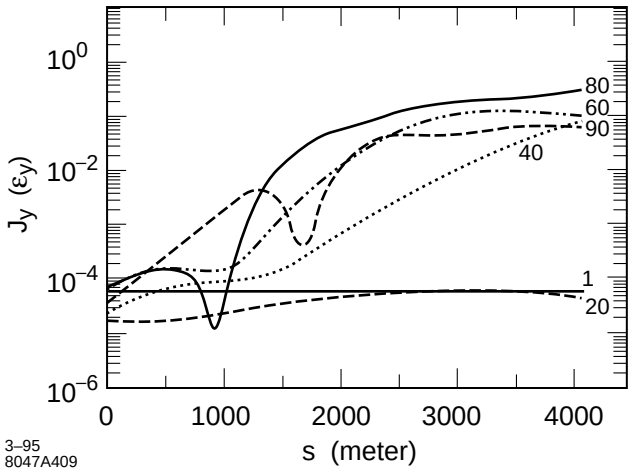


Figure 4-53. *Growth of the action for every twentieth of 90 bunches in the NLC damping ring for a pressure of* 10^{-8} *Torr, an atomic mass of 28 (carbon monoxide), and* 1.5×10^{10} *particles per bunch.*

fast multi-bunch instability. The growth time of this instability at the end of the bunch train is given by

$$\tau_{\mathrm{asym,e-}}^{-1}[\mathrm{s}^{-1}] \approx 5p[\,\mathrm{Torr}]\frac{N_b^{\frac{3}{2}} n_b^2 r_e r_p^{\frac{1}{2}} L_{\mathrm{sep}}^{\frac{1}{2}} c}{\gamma \sigma_y^{\frac{3}{2}} (\sigma_x + \sigma_y)^{\frac{3}{2}} A^{\frac{1}{2}} \omega_\beta} \tag{4.47}$$

where A denotes the atomic mass number of the ions, n_b the number of bunches, N_b the number of particles per bunch, L_{sep} the bunch spacing, r_p and r_e the classical proton and electron radius, respectively, c the velocity of light, and $\omega_\beta \approx 1/\beta_y$. The rise time τ_{asym} in Eq. 4.47 refers to the asymptotic equation (for large times t)

$$J_y \approx \frac{J_{y0}}{8\pi\sqrt{t/\tau_{\mathrm{asym}}}} \exp\left(2\sqrt{t/\tau_{\mathrm{asym}}}\right) \quad . \tag{4.48}$$

where J_y denotes the vertical action of the trailing bunch. The instability amplifies the Fourier component of the vertical bunch-to-bunch displacement which corresponds to the ion-oscillation frequency. Assuming the initial displacements are due to the finite number of electrons, the vertical action at time zero is given by $J_{y0} \approx \sigma_y^2/(N_b n_b 2\beta_y) \sim 3 \times 10^{-24}$ m. The rise time of Eq. 4.47 depends strongly on the number of bunches, the number of particles per bunch and the beam size. Due to the small vertical beam size and the rather large number of electrons, the growth times expected for the different NLC subsystems are exceedingly short.

In the electron damping rings, the predicted rise time for the trailing bunches is 100 ns at a pressure of 10^{-8} Torr. At 10^{-9} Torr the rise time would be 1 μs, which is still too short for a feedback system. However, when the instability grows from Schottky noise, about 200 rise times are required before the oscillation amplitude becomes of the order of the beam size. Moreover, decoherence and filamentation of the ions, not included in the derivation of Eq. 4.47, may lead to a further increase of the rise time by about a factor of two [Stupakov 1995]. Therefore, for the NLC damping rings a pressure of 10^{-9} Torr may be barely acceptable. An analytical treatment of the instability mechanism, computer simulations, possible remedies, and open questions are discussed in Ref. [Raubenheimer 1995a]. Figure 4-53 shows an example of the simulation results. The vertical action for different bunches of the train is depicted as a function of the traversed distance, for a pressure of 10^{-8} Torr. The rise times deduced from Figure 4-53 and the analytical estimate, Eq. 4.47, agree within a factor of two.

A positron bunch train deflects the equally charged ions and does not show the same ion-driven multibunch instability. However, single positron bunches may interact with the ionized electrons, which results in a similar instability, albeit

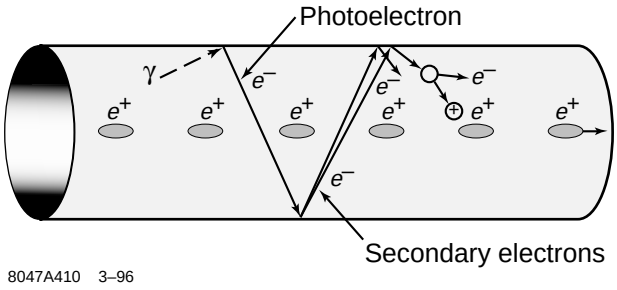


Figure 4-54. *Sketch of the beam-induced electron plasma.*

with larger rise times. In this case, the asymptotic growth rate at the bunch tail can be written as

$$\tau_{\mathrm{asym,e+}}^{-1}[\mathrm{s}^{-1}] \approx 7p[\,\mathrm{Torr}]\frac{N_b^{\frac{3}{2}} r_e^{\frac{3}{2}} \sigma_z^{\frac{1}{2}} c}{\gamma \sigma_y^{\frac{3}{2}} (\sigma_x + \sigma_y)^{\frac{3}{2}} \omega_\beta} \tag{4.49}$$

where σ_z is the rms bunch length. At a pressure of 10^{-9} Torr, the expected instability rise time for the positron Damping Ring is 120 μs. Since this is comparable to the synchrotron period (about 150 μs), the synchrotron motion will prevent the buildup of the instability.

Several possible remedies have been suggested for the multibunch instability [Raubenheimer 1995a]. First, additional gaps in a bunch train will alleviate the effect. Second, it may be possible to design lattices in which the ion frequency is strongly location-dependent. If the instability saturates at an amplitude of about one sigma, it is also conceivable to design damping rings in which the equilibrium emittance is about half the desired emittance. Preliminary simulation results indicate, however, that the saturation is not complete and that further amplitude growth at a smaller rate is possible. Finally, after extraction from the ring, fast kickers could realign the trailing bunches. Until the suitability of such schemes has been demonstrated, it will be assumed that in the damping rings of the NLC an average pressure of or below 10^{-9} Torr has to be maintained, to preserve the vertical emittance.

4.4.7 Positron-Electron Instability

Another instability that has been discussed is the positron-electron instability [Ohmi 1995]. This arises when an intense, positively charged bunched beam produces an electron cloud (or plasma) inside the vacuum chamber in which it travels. The basic mechanism is as follows: photons emitted due to synchrotron radiation hit the inside of the vacuum chamber and create photoelectrons. These, in turn, are accelerated transversely by the positively charged bunches and hit the wall on the opposite side of the vacuum chamber. Depending on their energy, these photoelectrons can be absorbed or can create secondary electrons. In this latter case, the secondaries are, in turn, accelerated, and the process is iterated, as sketched in Figure 4-54.

There are two effects from these electrons: they desorb gas molecules from the walls of the vacuum chamber and possibly ionize them, and they create an electron cloud. The gas desorption effect decreases in importance as the chamber gets "scrubbed" with increased use. But the electron cloud, of course, remains.

For high-enough current, a resonance occurs when the typical traversal time of the electrons across the chamber is equal to the time spacing of the bunches, leading to beam-induced multipactoring [Grobner]. But even for currents well below this threshold, as in the case of the NLC damping ring, the electron cloud still exists and couples the transverse motion of the bunches, leading to a potential coupled-bunch instability.

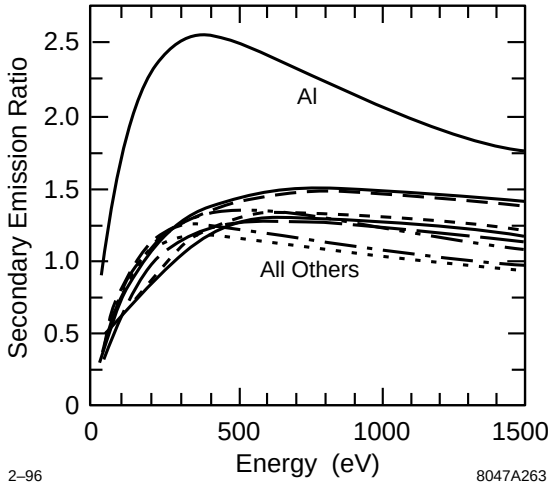


Figure 4-55. *The secondary emission ratio for various metals as a function of the incident electron energy in eV.*

Experimental evidence of such an electron-cloud instability has been reported at the KEK Photon Factory [Izawa 1995]. A related phenomenon has also been observed at CESR [Rogers 1995], in which case the electron cloud is not produced by the mechanism described, but by the trapping action of the combined electric and magnetic leak fields of the ion pumps. Although in this case the electron cloud occupies only a very small fraction of the vacuum chamber, it is enough to produce coupled-bunch motion. A theoretical explanation of the coupled motion based on the plasma idea has been put forward [Ohmi 1995], which predicts fairly fast growth rates for the instability. The predicted value for the onset of the instability is in rough agreement with the observations.

We have developed a simulation code [Furman 1995] and analytic calculations [Heifets 1995] to study this phenomenon at the NLC damping ring and the PEP-II positron beam. An important ingredient of the simulation is secondary electron emission. Figure 4-55 shows experimental measurements of the emission ratio for various metals (these curves are incorporated into the simulation code). Aluminum has the highest ratio, and that all other metals are not too different from each other.

Preliminary simulation results show that, for the nominal main damping ring parameters, an uncoated Al vacuum chamber would lead to a fast avalanche of electrons following injection into the machine. This avalanche would eventually stop due to the repulsive nature of the space-charge force of the electron cloud [Heifets 1995]. Even so, it is likely that the resultant instability would be difficult to control. On the other hand, if the vacuum chamber is coated with any of the other metals indicated in Figure 4-55, the simulation shows that the electron cloud density reaches saturation even if space-charge effects are not taken into consideration. The reason for this equilibrium state is a balance between the predominant absorption of the electrons upon hitting the walls, and the creation of photoelectrons following the passage of every bunch through the bending magnets.

For the nominal Main Damping Ring current and bunch spacing, the characteristic time of the plasma is < 10 ns. Thus a beam gap equivalent to a few bunch spacings makes the plasma disappear altogether. This implies that there will be no turn-to-turn coupling. If a leading bunch oscillates transversely, the electron plasma will act as a coupler and drive the trailing bunches. The amplitude of the driven bunches will grow linearly with time with a characteristic time scale estimated at 0.2 ms. Thus the feedback system will need to be able to damp instabilities of this time scale. However, we should point out that this is a subject of ongoing research and the correct answer may differ substantially.

There have been a number of solutions proposed for this effect. One such solution would place strips of ceramic, coated with thin conducting layer, along the edges of opening of the ante-chamber. Biased with a low DC voltage,

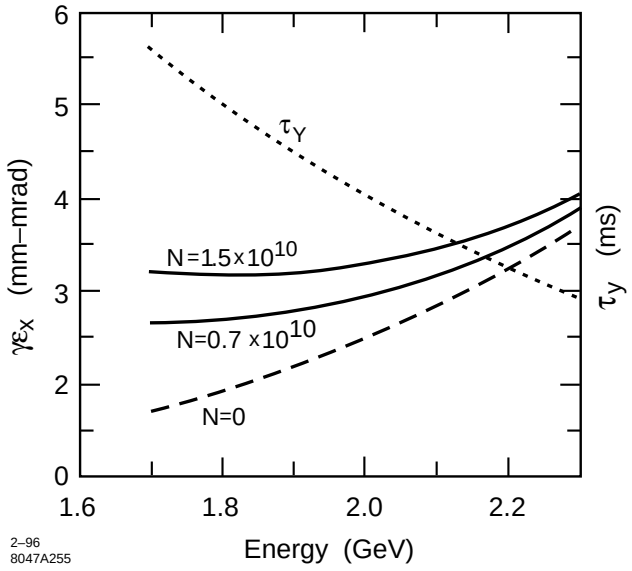


Figure 4-56. *Horizontal equilibrium emittance and vertical damping time versus ring energy for different single bunch charges.*

such a design may prohibit production of the photo-electrons and eliminate the photo-electron driven instability of the positron storage rings.

Finally, this effect may be important in the pre-damping ring as well as in the positron main damping ring. At this time, we have not evaluated the effect for those parameters.

4.4.8 Intrabeam Scattering

Intrabeam scattering is a very significant effect in the main damping rings because of the small normalized emittances; it has negligible effect in the pre-damping ring. In the main damping rings, intrabeam scattering has been calculated with ZAP [Zisman 1986]. Figure 4-56 is a plot of the horizontal emittance and the vertical damping time versus the ring energy. The vertical damping time is the dotted curve and there are three curves for the horizontal emittance: $N = 1.57 \times 10^{10}$ (solid), $N = 0.7 \times 10^{10}$ (solid), and $N = 0$ (dashes). The intrabeam scattering emittance growth was calculated by ZAP. Notice that the emittance decreases to a minimum as the ring energy is decreased to roughly 1.8 GeV. Unfortunately, the vertical damping time increases rapidly as the energy is decreased and, at an energy of 1.7 GeV, the damping time would only allow operation at 120 Hz.

At the nominal energy of 1.98 GeV for 180-Hz operation, ZAP estimates the horizontal emittance as roughly 3.30×10^{-6} m-rad at a charge of 1.5×10^{10} and 2.96×10^{-6} m-rad at a charge of 0.7×10^{10}. But, ZAP calculates the rms growth of the beam size. Because of the fast damping times, a significant fraction of the rms growth is actually contributions to the beam halo and not an increase of the beam core [Raubenheimer 1994]. This results in a correction for the growth of the core that, for the NLC main damping ring parameters, is roughly 30%. Thus, at 1.98 GeV, with the maximum beam charge, we expect the horizontal core emittance to increase to 3.1×10^{-6} m-rad; at lower charge, the growth is less.

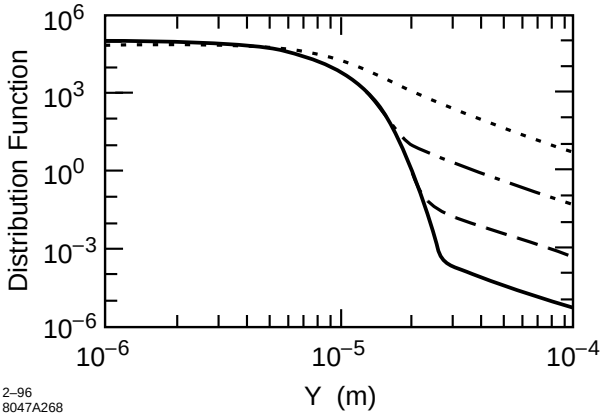


Figure 4-57. *Beam distribution in the NLC damping ring with CO gas pressures of 10^{-12} Torr (solid), 10^{10} Torr (dashed), 10^{-8} Torr (dot-dash), and 10^{-6} Torr (dotted).*

It is important to note that the scattering in the vertical plane will be less than in the horizontal plane [Raubenheimer 1993b]. This arises because intrabeam scattering emittance growth occurs due to changes in the longitudinal particle energy from the scattering. It is similar to the emittance growth from synchrotron radiation, except the intrabeam scattering occurs throughout the ring while the synchrotron radiation only occurs in the bending magnets. In a storage ring, the horizontal dispersion is minimized at the bends while the vertical dispersion, which is due to errors, is relatively uniform in magnitude around the ring. Thus, the intrabeam emittance growth has a much larger relative contribution to the horizontal emittance than the vertical emittance. We have not yet calculated the vertical growth for the NLC parameters, but experience with the KEK ATF damping ring suggests that the vertical emittance growth should be roughly 30% of the horizontal growth. In this case, the vertical emittance growth due to intrabeam scattering will not be significant.

4.4.9 Beam-Gas Scattering

The vacuum pressure in the NLC damping rings will be less than 10^{-8} Torr. This is low enough so that scattering with residual gas atoms does not limit the beam lifetime or dilute the vertical emittance. The tolerable vacuum level is determined from several different effects. First, elastic scattering on atomic nuclei may increase the equilibrium emittance. The cross section for this process is about

$$\sigma_{\rm scatt} \approx \sum_i 4\pi\alpha^2 Z_i^2 (1.4 a_0 Z_i^{-\frac{1}{3}})^2 \tag{4.50}$$

where α denotes the fine structure constant, a_0 Bohr's radius, and Z_i the atomic number. The sum is over all atoms of a molecule. For carbon monoxide, CO, a common molecule in accelerator vacuum systems, we have $\sigma_{\rm scatt} \approx 0.5$ Mbarn. At a pressure of 10^{-8} Torr, an electron undergoes no more than 5 collisions per second, and, therefore, during a vertical damping time of 4.8 ms only a small halo is created around the undisturbed beam core [Hirata 1992, Raubenheimer 1992]. Figure 4-57 shows the beam distribution due to gas scattering for several vacuum pressures [Raubenheimer 1992].

Second, the beam lifetime is reduced by bremsstrahlung in the field of the atomic nuclei [Piwinski 1985]. In this case the cross section is

$$\sigma_{\rm brems} \approx \sum_i \frac{4}{3} 4 r_e^2 Z_i^2 \alpha \ln \frac{E}{\Delta E} \ln \frac{183}{Z_i^{-\frac{1}{3}}} \tag{4.51}$$

where r_e is the classical electron radius, and $\Delta E/E \approx 10^{-2}$ the energy acceptance of the ring. Again for CO, the cross section is $\sigma_{\rm brems} \approx 6$ barn, which translates into a beam lifetime of 5 hours at a pressure of 10^{-8} Torr.

4.4.10 Lifetime Limitations

In a damping ring, it is unlikely that the beam lifetimes could limit the operation of the ring since the beams are stored for a very short time. But, poor beam lifetimes could make commissioning and studying the ring difficult. The primary beam lifetime limitations are excessively small apertures, beam-gas scattering, and the Touschek effect; the latter refers to large-angle collisions between particles within the beam that lead to particle loss.

With adequate design and tolerances, one should be able to avoid the first limitation. In addition, with reasonable vacuum pressures, the lifetime due to beam-gas scattering should not be significant. Assuming a pressure of 10^{-8} Torr in the NLC damping rings, the lifetime due to beam-gas scattering is roughly four hours; this is more than sufficient.

In contrast, the Touschek lifetime will tend to be severe in the main damping rings since it depends inversely upon the particle density in the beam. In the NLC main damping rings, the Touschek lifetime is roughly two minutes. While this is long compared to the operating beam storage time of 22 ms, it may be too short to study the properties of a stored beam. In this case, one can increase the vertical beam size or decrease the number of particles per bunch, thereby decreasing the beam density and increasing the lifetime.

4.4.11 PPDR Collective Limitations

Collective effects are not expected to be significant in the Positron Pre-Damping Ring. The PPDR has a longer bunch length and a larger vacuum chamber, reducing the contributions to the broadband impedance and the resistive-wall impedance. Thus, the growth rates of the coupled-bunch instabilities, that are driven by the resistive-wall impedance, should be well below the radiation damping rate. In addition, the longitudinal microwave threshold should be increased substantially due to the lower impedance and the larger momentum compaction. Finally, issues such as the Touschek lifetime and intrabeam scattering are not relevant because of the large beam emittances.

The one collective issue that may be significant in the pre-damping ring is the positron-electron instability, referred to as the "Ohmi Effect." This may be partially ameliorated by using an ante-chamber to absorb most of the radiation and by covering the inside of the vacuum chamber with titanium nitride or similar coating. In the future, we will verify the importance of the other effects and, hopefully, will have additional data with which to evaluate the "Ohmi Effect."

4.5 RF Systems

There are four distinct rf systems in the damping ring complex: the S-band and L-band energy compressors, and the pre-damping ring and main damping ring rf systems. As discussed in Section 4.2.9, the ring rf systems will be based

on 714-MHz cavities with HOM damping. The present design is based on the PEP-II cavity design but there are many laboratories around the world working on rf cavities with damped HOMs and we will gain from their experience.

In the pre-damping ring, the total rf voltage is 2.0 MV which leads to an energy bucket of $\pm 1.5\%$. Although this is not sufficient to capture the full $\pm 2\%$ energy spread, the L-band energy compressor will reduce the incoming energy spread by at least a factor of two (Section 4.3.3). The injection line rf systems are based on the electron and positron source linac rf systems while, as stated, the damping ring rf systems are based on the PEP-II rf systems.

In the next sections, we will describe the rf systems for the main damping rings and then the pre-damping ring. Next, we discuss the low-level rf control system for the main damping ring. The low-level rf controls are extreme important in these damping rings because of the high beam-loading and very tight stability tolerances on the extracted bunches. To achieve the stability, the low-level rf system requires many layers of feedback but, as has been realized in the SLC damping rings, it is extremely important to understand the interactions between these feedback loops. This is described in Section 4.5.3. Finally, we will outline the energy compressor rf systems; these systems are described in greater detail in Chapter 6.

4.5.1 Main Damping Rings RF Systems

The rf systems of the NLC main damping rings have to provide sufficient voltage to the beam to compensate for losses due to synchrotron radiation (from the dipole magnets and from the wigglers), losses into higher-order modes (HOMs) of the rf cavities, and parasitic losses to vacuum chamber components. The large synchrotron radiation losses, HOM, and parasitic losses result in an rf system with a large power throughput to the beam. The beam current in the cavity is a factor of 11 greater than the generator current needed to provide the gap voltage.

In addition, the large beam current makes the suppression of HOM impedances more important, to avoid uncontrollable coupled-bunch and single-bunch instabilities. Therefore, in addition to providing the required voltage, the systems should present a low-impedance environment to the beam to maximize instability threshold currents, and avoid the necessity for or minimize the power requirements of beam feedback systems.

The frequency of the damping ring rf systems is determined by the bunch spacing in the main linac—the bunches within a train are separated by 1.4 ns—thereby requiring a minimum rf frequency of 714 MHz. Four trains, 90 bunches each, occupy the damping ring at any one time, with 1.57×10^{10} particles per bunch and an average current of 1.2 A.

The accelerating voltage requirement of 1 MV is achievable using conventional copper cavity construction techniques. To alleviate the problems of coupled-bunch instabilities driven by cavity higher-order modes, some type of suppression of the HOMs is required. There are several suitable such monochromatic cavity designs currently being used or designed, which could find application in the NLC damping rings. For example, the ATF cavity being built for KEK uses a cylindrical body, to which end plates are brazed. These plates have nose cone structures and coupling apertures to which the damping waveguides attach, two on each end of the cavity [Sakanaka 1995]. A different design has been developed for DAΦNE, using a bell-shaped cavity structure with a rounded cavity and enlarged beam-pipe apertures to allow modes to propagate past the ends of the cavity. Long tapers connect the cavity to the smaller diameter beam pipe, resulting in a decreased R/Q value for many modes. Damping waveguides are attached to the cavity body, and to the adjacent tapers [Barry 1993].

For this study we choose a cavity based on a design developed for the PEP-II B-factory project. This cavity has a rounded body and reentrant nose cones, providing a large shunt impedance, and three waveguide loads mounted on the cavity body to efficiently suppress the HOMs. Tapers from the cavity nose cone region to the straight section vacuum chamber provide a smooth transition from the scaled PEP-II cavity to the small vacuum chamber, help to minimize the cavity broadband impedance, and also support additional damping waveguides.

Energy	1.98 GeV
Circumference	223 m
RF	714 MHz
Harmonic Number	532
σ_ϵ	0.09%
σ_z	3.9 mm
α	0.000465
$U_{s.r.}$	644 keV/turn
U_{HOMs}	7.5 keV/turn
$U_{parasitic}$	25 keV/turn
V_{RF}	1.0 MV
# Cavities	2
# Klystrons	2
Cavity Wall Dissipation	41.6 kW/cavity
Klystron Power	600 kW/klystron
Shunt Impedance	3.0 MΩ/cavity
Beam Current	1.2 A
Coupling Factor	11
Unloaded Q	25500
Synchronous Phase Angle	42.6°
Optimum Detuning at Full Current	148 kHz
Synchrotron Frequency	5 kHz
Loaded Q	2170
Energy acceptance	$\pm$1.5%

Table 4-20. *Rf system parameters for the NLC main damping rings.*

Other cavity designs are feasible, such as reducing the aperture at the nose cones to the nominal vacuum chamber dimensions, avoiding the use of tapers. This is expected to increase the strength of many HOMs, but the damping waveguides may still provide the required attenuation for these modes and the increase in broadband impedance may be acceptable. A more detailed optimization, including factoring in the cost of different options, will take place in the future.

The main parameters of the NLC damping ring rf systems are listed in Table 4-20. To determine the power requirements of the rf systems, we need to estimate the losses to parasitic modes in the vacuum chamber components. This we have done by scaling from measurements and calculations performed for other, recently designed or built, low-impedance storage rings. As discussed subsequently, the klystron power requirements, two klystrons each rated at 600 kW, appear to be within the capabilities of manufacturers at this frequency.

In addition, transients in the rf system excited by the gaps in the circulating beam and by injection timing mismatches cause a shift in synchronous phase and timing errors along the bunch train. These transients may be suppressed by the use of feedback systems around the rf klystrons and cavities. The required low-level rf system is described in Section 4.5.3.

Cavity Design

The determination of the number of cavities required involves a compromise between their power-handling capabilities and the beam impedance presented by the cavities. Power-handling capabilities of the rf window and coupler, and the

problems of heat dissipation in the cavity walls and apertures are the limitation to the power into each cavity. These problems have been extensively studied for the PEP-II B-factory rf systems, and we choose to follow their lead and use a cavity design based on the PEP-II B-factory cavities.

The basic cavity shape is a reentrant structure with nose cones, similar to the cavities currently used at the KEK Photon Factory, Daresbury SRS, and the ALS at Berkeley. A conventional copper construction is used, with aperture coupling to the feeder waveguide. Re-entrant nose cones are used to optimize the shunt impedance of the fundamental mode. We scale the dimensions of the PEP-II cavities to achieve a resonant frequency of approximately 714 MHz, with a beam-pipe aperture of 6.2-cm diameter at the nose cones. As discussed in Section 4.4.1, the large aperture at the nose cones reduces the broadband impedance of the cavity and is clearly advantageous. To mate the cavity to the adjacent beam pipe, we use long tapers, while maintaining a total cavity length (including tapers) of less than 0.5 m.

Figure 4-58 shows a half section of the cavity model illustrating the coax damping waveguides on the taper but not showing the damping waveguides on the cavity. Calculations with the electromagnetic design code URMEL give a shunt impedance of $R_s = 4.35\mathrm{M}\,\Omega, Q_0 = 36000$. For the B-factory cavities the addition of the damping waveguides reduced the shunt impedance and unloaded Q value by 30% (compared to the URMEL values). Using the same scaling, we calculate a shunt impedance of $R_s = 3.0\mathrm{M}\,\Omega$, $Q_o = 25500$, and $R/Q = 120\,\Omega$ for the NLC damping-ring rf systems.

The total loss parameter for each cavity is computed by ABCI to be $k_{\mathrm{total}} = 1.75\,\mathrm{V}/pC$ for a bunch length of 4.0 mm, of which the fundamental mode contributes 0.27 V/pC, giving the contribution from higher-order modes $k_{\mathrm{HOMs}} = 1.48\,\mathrm{V}/pC$ per cavity.

Three waveguides are attached to the cavity body to damp the higher-order modes. These waveguides are dimensioned to have a cut-off frequency above the frequency of the accelerating mode of the cavity, but below the frequency of the lowest HOM. Thus the power in HOMs may couple through the apertures in the cavity wall to the waveguides, where it is absorbed in a cooled lossy dielectric material. The location of these waveguides is at 30° from the center of the radius of the cavity body, this position being optimized to couple to the strongest HOMs in the cavity [Rimmer 1992]. The waveguides are located equidistantly in azimuth. The location of the lossy material within the waveguide is at a sufficient distance from the cavity aperture to exclude dissipation of the fundamental mode power from evanescent fields in the waveguide. Experience with the PEP-II cavities shows that Q values of HOMs are damped by two orders of magnitude in most cases, and three orders for targeted modes [Rimmer 1992].

Many higher-frequency cavity modes will spread into the tapers, where the delocalizing effect generally decreases their impedance through a reduction in the transit time factor. These modes may be further damped by locating additional damping waveguides on the tapers. There are many modes at relatively high frequencies, capable of driving within-bunch instabilities which are difficult to combat with feedback systems. Effective broadband damping eliminates these problematic modes. At this time, we are performing 2- and 3-D MAFIA studies of the cavity to optimize the placement of the waveguide on the taper and those on the cavity body.

The PEP-II cavities are designed for a maximum power dissipation of 150 kW. To maintain the same power per unit area in the damping ring cavities, we scale by the frequency ratio of the two rings' rf systems squared to obtain 67 kW per cavity maximum power dissipation in the NLC damping rings cavities. De-rating by 30% gives a conservative dissipation goal of 47 kW per cavity or less.

Higher-Order Mode and Parasitic Losses

Cavity HOM losses are determined from the loss parameter k_{HOMs} and the charge per bunch q; U_{HOMs} is the loss per turn to the cavity HOMs

$$U_{\mathrm{HOMs}} = k_{\mathrm{HOMs}} q \tag{4.52}$$

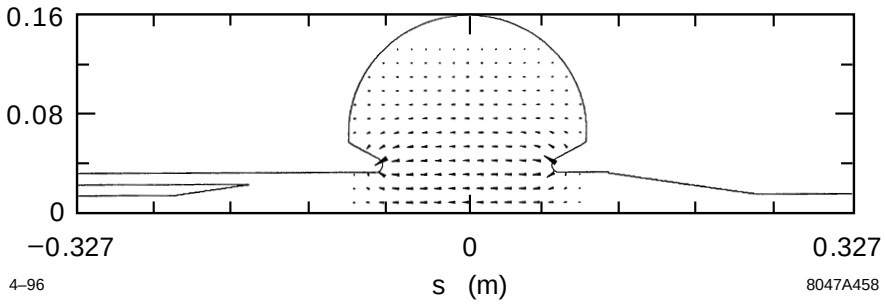


Figure 4-58. *Half section of the rf cavity, without cavity damping waveguides.*

giving $U_{\mathrm{HOMs}} = 3.7$ keV per cavity per turn.

We scale from measurements and calculations of other machines, parasitic losses to vacuum chamber components other than the rf cavities, such as bellows shields, BPMs, kicker magnets, and resistive wall, With careful attention to the design of components, using sophisticated electromagnetic design codes and careful measurements, the parasitic losses can be minimized in modern machines. We choose PEP-II and the ALS as examples of clean, low-impedance machines from which to obtain an estimate of the loss parameter for the NLC damping rings.

For the PEP-II B-factory, the total loss parameter for vacuum chamber components excluding the rf cavities has been estimated to be 6 V/pC, for a bunch length of 10 mm [PEP-II 1993]. Scaling the loss parameter with bunch length as k$\sim \rho^{-1.5}$, with machine circumference as 220/2200, and with the ratio of the vacuum chamber diameters as 6.2/2.8, gives $\mathrm{k_{parasitic}} = 6.9\,\mathrm{V/pC}$.

The ALS is a very similar machine to the NLC damping rings, in that the circumference and the bunch length are almost identical in the two machines. The vacuum chamber radius in the ALS is larger than in the NLC damping rings. Measurements of the ALS vacuum chamber components indicate a loss parameter of approximately 5.8 V/pC (excluding cavities) [Corlett 1993]. Scaling to the NLC damping ring vacuum chamber radius, we estimate $\mathrm{k_{parasitic}} = 15\,\mathrm{V}/pC$. Finally, the calculations in Section 4.4.1 give a $\mathrm{k_{parasitic}} = 5\,\mathrm{V}/pC$. Assuming a value of 10 V/pC, yields losses of $\mathrm{U_{parasitic}} = 25$ keV per turn. Of course, it should be noted that the parasitic losses are very small compared to the synchrotron radiation losses and thus the design is insensitive to the exact value.

Voltage Requirements

The voltage required to compensate for the synchrotron radiation loss, due to both dipole bend magnets and the wigglers, is $\mathrm{U_{SR}} = 644$ keV per turn. In addition, the vacuum chamber parasitic losses are $\mathrm{U_{parasitic}} = 25$ keV per turn, and the rf cavity HOM losses $\mathrm{U_{HOMs}} = 3.7$ keV per turn per cavity. For two rf cavities, the total loss per turn is 677 keV. In order to provide an energy acceptance of $\pm 1\%$, the cavity voltage must be at least $\mathrm{V_{rf}} = 0.85$ MV. To provide some contingency for additional energy acceptance and bunch length control, we take the cavity voltage requirement to be $\mathrm{V_{rf}} = 1.0$ MV.

Power Requirements

The cavity power dissipation is given by

$$\mathrm{P_c} = \frac{\mathrm{V_{RF}^2}}{2\mathrm{R_s}} \tag{4.53}$$

and each cavity dissipates 41.6 kW, based on two cavities, giving a total $\mathrm{P_c} = 83$ kW. This is within the estimated power-handling capability of 47 kW per cavity.

The power delivered to the beam is determined from the synchrotron radiation, HOM losses, and parasitic losses is

$$P_b = I_o(U_{\mathrm{SR}} + U_{\mathrm{HOMs}} + U_{\mathrm{parasitic}}) = I_o U_{\mathrm{losses}} \tag{4.54}$$

where $U_{\mathrm{losses}} = 677$ keV per turn, $\mathrm{I_o} = 1.2$ A for an average beam current, and the power to the beam $\mathrm{P_b} = 812$ kW.

The total rf power required $\mathrm{P_{rf}} = \mathrm{P_b} + \mathrm{P_c} = 895$ kW. Feeding each cavity independently would require two klystrons rated at least 450 kW. Providing a 20% overhead for rf feedback, and 10% for losses in transmission to the cavities, each klystron would be rated at 600 kW.

Coupling Factor and Loaded Q

To be able to obtain the condition of a matched cavity with full beam loading, the cavity coupling factor β_o is determined from

$$\beta_{\mathrm{o}} = \mathrm{I} + \frac{\mathrm{P_b}}{\mathrm{P_c}} \tag{4.55}$$

and we find $\beta_{\mathrm{o}} = 10.7$. The loaded quality factor of a cavity is then

$$\mathrm{Q_q} = \frac{\mathrm{Q_o}}{\mathrm{l} + \beta_{\mathrm{o}}} \tag{4.56}$$

and Q_l = 2170. The filling time $2Q_l/\omega = 0.97$ μs. The synchronous phase angle ϕ_{s} is determined from the total loss per turn U_{losses} and the rf voltage

$$\sin\phi_s = \frac{U_{\mathrm{losses}}}{\mathrm{V_{rf}}} \tag{4.57}$$

which gives ϕ_{s} = 42.6°. The optimum cavity detuning frequency $\delta\mathrm{f}$ is determined from

$$\delta_{\mathrm{f}} = \frac{\mathrm{f_{rf}}}{2\mathrm{Q_0}}\frac{\mathrm{P_b}}{\mathrm{P_c}}\cot\phi\ \mathrm{s} \tag{4.58}$$

where $\delta_{\mathrm{f}} = -148$ kHz, which is small compared to the orbit frequency of 1.36 MHz. Growth times of coupled-bunch modes driven by the accelerating cavity mode are likely to be long.

RF Parameters with Three Cavities

The possibility of using three rf cavities instead of two has been investigated, to reduce the power dissipation in the cavities, the throughput rating of the rf vacuum windows, and the power rating of the klystrons. Table 4-21 lists the parameters changed from the two-cavity case. Klystron power requirements are reduced to 350 kW per klystron, including 20% overhead for feedback systems. Window throughput power is also reduced by 40% to 350 kW. The cavity wall losses are reduced from 42 kW per cavity to less than 20 kW per cavity. Optimal cavity detuning increases to 222 kHz, and the coupling factor is increased to almost 16.

Transient Suppression Using Feedback or Feedforward

The nonuniform bucket population generates transients in the rf system, and the beam sees a nonuniform accelerating voltage. Each bunch in a train has its own individual synchronous phase angle, and the spacing between bunches changes to accommodate this. The use of sub-harmonic cavities, feedback around the rf system, and feedforward in the rf system are being evaluated to reduce the spread in bunches along the rf waveform.

# Cavities	3
U_{HOMs}	11.3 keV/turn
Cavity Wall Dissipation	18.5 kW/cavity
# Klystrons	3
Klystron Power	350 kW/klystron
Coupling Factor	15.7
Synchronous Phase Angle	42.8°
Optimum Detuning at Full Current	222 kHz
Loaded Q	1529

Table 4-21. *Rf system parameters with three cavities.*

Klystrons and accessories

The necessary rf power at 714 MHz will be provided by high-power klystrons. Depending upon the number of cavities and the overhead required for feedback and feedforward, the power could be supplied by three klystrons feeding three cavities, two klystrons feeding two cavities or one super-power klystron driving a pair of cavities. The choice will depend on economic as well as technical considerations. Klystrons exist with up to 1.2-MW CW output in the 350-508-MHz region and power supplies, circulators and loads have been developed to handle this power level. This technology can be readily adapted to the higher frequency and there is already R&D underway in the industry on 1-MW tubes in the 700-MHz frequency range for applications such as tritium production.

There is considerable commercial activity in this frequency range for the television broadcast industry and many mature products exist at the 50-70-kW CW power level. For one- or few-off applications it is possible to combine the output from several of these tubes to produce a higher-power supply such as that used at the SPS at CERN. This station used eight 60-kW TV klystrons (Phillips YK1190 series), producing 450 kW at 800 MHz (newer, higher-power tubes are now available). Given the relatively high volume of production of these transmitter tubes it may be cost-effective to build a high-power source this way with minimal R&D investment.

RF cavity windows

Approximately 600 kW per cavity (into two cavities) will be required to support 1.2 A of beam current at 1 MV total voltage, and overhead (in peak voltage) is required for feedback systems, so the rf window will be a critical component. Both coaxial and disk-type windows routinely operate on high-power klystrons at 1-1.2-MW CW, and cavity windows have been developed for PEP-II and other machines for up to 500 kW throughput. These technologies can be readily scaled to 714 MHz. Additional safety margin could be achieved by supplying the power through two windows per cavity which would provide a very reliable system. Using three cavities per ring would also reduce the power per window significantly.

4.5.2 Pre-Damping Ring RF Systems

The rf systems of the NLC pre-damping rings are based on “monochromatic” cavities as described for the main damping rings in Section 4.5.1. These systems serve to provide sufficient voltage to the beam to compensate for losses due to synchrotron radiation, losses to higher-order modes (HOMs) of the rf cavities, parasitic losses to

vacuum chamber components, and to provide a sufficient rf potential well to capture the injected beam. The resulting accelerating voltage requirement of 2 MV is achievable using four conventional copper cavities.

In order to determine the power requirements of the rf systems, we estimate the losses to parasitic modes in the vacuum chamber components by scaling from the losses estimated for the main damping rings, as described in Section 4.5.1. For the pre-damping ring, with smaller machine circumference and larger beam pipe diameter than the main damping rings, we estimate $\mathrm{k_{parasitic}} = 7$ V/pC.

We choose to use four cavities and two klystrons, with pairs of cavities fed from a single klystron. The klystron power rating of 440 kW appears to be within the capabilities of manufacturers at this frequency.

Table 4-22 lists the main parameters of the NLC pre-damping ring rf systems.

Voltage Requirements

The voltage required to compensate for the synchrotron radiation losses is $\mathrm{U_{s.r.}} = 371$ keV per turn. In addition, the vacuum chamber parasitic losses are $\mathrm{U_{parasitic}} = 17$ keV per turn, and the rf cavity HOM losses $\mathrm{U_{HOMs}} = 3.7$ keV per turn per cavity. For four rf cavities, the total loss per turn is 404 keV.

A larger cavity voltage is required to capture the injected beam, and at a nominal cavity voltage $\mathrm{V_{rf}} = 2.0$ MV, a sufficient energy acceptance of $\pm 1.46\%$ is obtained.

Power Requirements

The cavity power dissipation is given by

$$\mathrm{P_c} = \frac{\mathrm{V_{rf}^2}}{2\mathrm{R_s}} \tag{4.59}$$

and 41.6 kW is dissipated in each cavity, based on 4 cavities, giving a total $\mathrm{P_c} = 167$ kW. This is within the estimated power handling capability of 47 kW per cavity (see Section 4.5.1).

The power delivered to the beam is determined from the synchrotron radiation, HOM losses, and parasitic losses:

$$P_b = I_o(U_{s.r.} + U_{HOMs} + U_{parasitic}) = I_o U_{\mathrm{losses}} \tag{4.60}$$

$U_{\mathrm{losses}} = 404$ keV per turn, and for an average beam current $\mathrm{I_o} = 1.2$ A the power to the beam $\mathrm{P_b} = 485$ kW.

The total rf power required $\mathrm{P_rf} = \mathrm{P_b} + \mathrm{P_c} = 652$ kW. Feeding two cavities from one klystron requires two klystrons rated at 326 kW. Providing approximately 20% overhead for rf feedback, and 10% for losses in transmission to the cavities, each klystron would be rated at 440 kW.

Coupling Factor and Loaded Q

To be able to obtain the condition of a matched cavity with full beam loading, the cavity coupling factor β_o is determined from

$$\beta_o = 1 + \frac{\mathrm{P_b}}{\mathrm{P_c}} \tag{4.61}$$

and we find $\beta_o = 3.9$. The loaded quality factor $Q_l = 5190$, and the filling time $2Q_l/\omega = 2.3\ \mu$s.

Energy	1.98 GeV
Circumference	171 m
RF Frequency	714 MHz
Harmonic Number	408
σ_ϵ	0.0977%
σ_z	7.5 mm
α	0.00511
$U_{s.r.}$	371 keV/turn
U_{HOMs}	16 keV/turn
$U_{parasitic}$	17 keV/turn
V_{rf}	2.0 MV
# Cavities	4
# Klystrons	2
Cavity Wall Dissipation	42 kW/cavity
Klystron Power	440 kW/klystron
Shunt Impedance	3.0 MΩ/cavity
Beam Current	1.2 A
Coupling Factor	3.9
Unloaded Q	25500
Synchronous Phase Angle	11.7°
Optimum Detuning at Full Current	197 kHz
Synchrotron Frequency	31.5 kHz
Loaded Q	5190
Energy acceptance	$\pm$1.46%

Table 4-22. *Rf system parameters for the NLC pre-damping ring.*

The synchronous phase angle ϕ_s is determined from the total loss per turn U_{losses} and the rf voltage

$$\sin\phi_s = \frac{U_{losses}}{V_{rf}} \tag{4.62}$$

which gives $\phi_s = 11.7°$. The optimum cavity detuning frequency δf is determined from

$$\delta f = \frac{f_{rf}}{2Q_o}\frac{P_b}{P_c}\cot\phi_s \tag{4.63}$$

$\delta f = -197$ kHz, which is small compared to the orbit frequency of 1.75 MHz, therefore growth times of coupled-bunch modes driven by the impedance of the accelerating cavity mode are likely to be long.

4.5.3 Low-Level RF Systems

The low-level system includes various regulating feedback loops, controls, and protection interlocks. Most of these constituents are common to other storage ring and damping ring rf systems and therefore pose no unusual design challenges. The design goal is to ensure ease in diagnosis of rf-related problems while minimizing the complexity of the rf system: anticipated error sources are corrected for locally and regulation of each klystron-beam-cavity interaction is independently maintained. The function of the low-level rf system is

- to regulate the cavity voltage and beam phase, and
- to minimize the effects of transient beam loading caused by injection and extraction.

Cavity voltage and beam phase regulation is maintained by passive feedback while transient loading in the cavities will be minimized using feedforward at injection. Similar beam-loading compensation techniques will be tested and used at PEP-II [PEP-II 1993] and the KEK Accelerator Test Facility (ATF) damping ring [ATF 1995]. The feedforward algorithm, which is discussed in detail below, also allows beam injection at a high duty cycle.

The stability of the rf system and the particle beam in the main damping rings has been studied using a numerical model of the beam-cavity interaction with multiple feedback loops. The study has yet to be applied to the positron pre-damping ring, however, the low-level rf system designs are expected to be quite similar. Unique to the NLC damping rings are the high beam currents (about a factor of 10 increase compared to the SLC). Analyses of transient loading at injection have therefore been performed. The simulations will be extended to study synchronous phase transients due to the gaps between bunch trains, which are discussed in Section 4.4.4.

Low-Level Controls

A block diagram of the low-level control system is shown in Figure 4-59. The rf from the main NLC oscillator is phase-shifted using feedback from the interaction point (IP) collision time (as determined by the arrival time of e^+ in one linac and the e^- in the other). This global phasing of the front end (source, injector linac, etc.) is required since the single-turn injection/extraction scheme, which minimizes beam phase and bunch length variations in steady-state operation of the damping rings, disallows independent phasing of beam phase at injection to and extraction from the damping rings. The phase-shifted rf from the main oscillator is frequency-divided to 714 MHz, amplified as necessary, and used as the phase reference for the damping ring rf system. A separate voltage-controlled oscillator (VCO) is provided to allow the ring frequency to be varied (for measurement of the chromaticity, for example).

Beam Phase Loop. Deviations in damping ring timing are corrected using a phase feedback loop which uses the beam phase measured with a damping ring BPM and adjusts the phases of the source, prelinac, and damping ring. The bandwidth of this loop is limited by the minimum frequency for coherent synchrotron oscillations ($\approx$2 kHz at full current).

Referring to Figure 4-59, a common local oscillator provides rf input to the two klystrons. Each klystron powers a one single-cell cavity with higher-order mode damping for coupled-bunch modes. A circulator upstream of each cavity isolates the klystron from reflected power arising from cavity detuning or the sudden absence of beam. Additional protection is provided by an rf switch upstream of each klystron which is activated by standard protection interlocks. There are slow interlocks including those for cavity vacuum and for water flow and temperatures in the klystrons, cavities, and circulator loads. Fast interlocks protect against reflected power faults and arcs in the klystrons and circulators.

Five feedback loops are used to stabilize the beam and accelerator. These include loops to regulate the klystron output, the cavity resonant frequency, and the cavity phase and amplitude. In the heavily beam-loaded limit, the cavity amplitude and phase loops should ideally correct for beam-induced transients. Using feedforward, however, because the phasor voltage of the cavity is maintained independent of the beam current or the change in beam current, the gain requirements of these loops is significantly reduced. The precise gain and bandwidth requirements in the feedback design have yet to be determined.

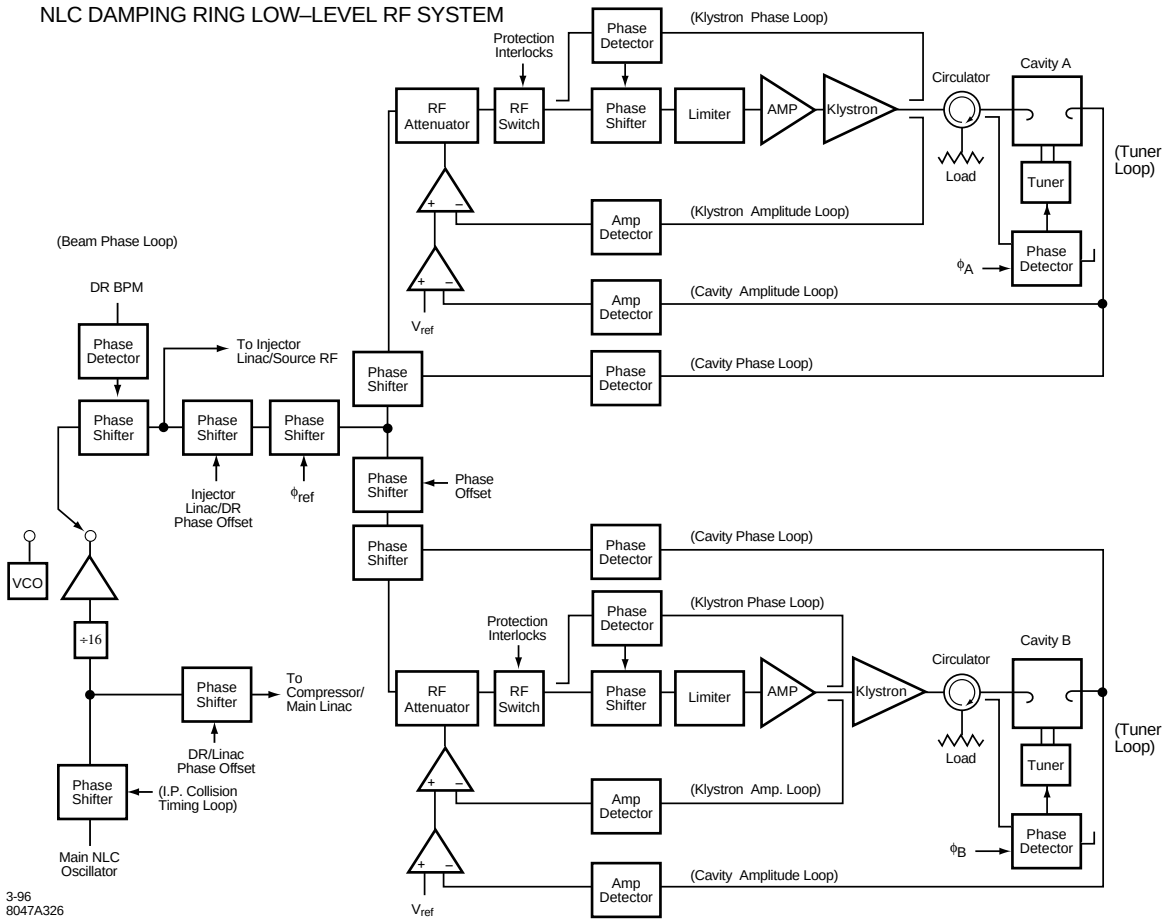


Figure 4-59. *Block diagram of the low-level control system for the NLC damping rings.*

Klystron Amplitude and Phase Loops. Slow, high gain feedback loops are used to correct for klystron high-voltage power supply ripple and other sources of phase or amplitude error. The input to both loops is a measure of the klystron forward power. The phase of this signal is compared with the input rf phase and the difference is used to shift the phase of the input rf. In the klystron amplitude loop, the rf signal is converted to DC with an amplitude detector. The difference in this amplitude from the error signal of the cavity amplitude feedback is used to adjust an rf attenuator upstream of the klystron.

Cavity Amplitude Loops. To regulate against slow changes in the cavity voltage, due to long-term thermal variations for example, the amplitude of each cavity is maintained at an externally specified level, $V_{\rm ref}$. The error signal is used to adjust an rf attenuator upstream of the klystron. The bandwidth is high and limited by the minimum frequency for coherent beam oscillations (≈ 2 kHz at full current).

Cavity Phase Loop. This loop maintains the phase of the net cavity voltage with respect to phase of the input rf by adjustment of an rf phase shifter. The bandwidth is limited to ≈ 2 kHz at full current.

Cavity Tuner Loops. A single loop per cavity is used to correct for thermal variations and reactive loading by the beam. For fast filling of the ring (see below), the tuner positions are fixed at externally specified angles, ϕ_a and ϕ_b, to minimize the reflected power from the cavities when operating at full current. The bandwidth of these loops is limited by the speed of the mechanical plungers ($\approx$1 Hz).

Feedforward Algorithm for Beam Injection

Because of the very high beam currents, beam-induced loading of the cavities at injection is a concern in the damping ring rf system design. Studies have shown that without compensation (either feedforward or direct feedback [Pedersen 1975]), beam-loading of the rf system could result in beam loss and instabilities in the regulation. A feedforward algorithm was devised to correct the injection transients. While the beam-loading limit was not a concern for the steady-state operation of the damping rings, direct feedback was also considered to help damp the transients.

Shown in Figure 4-60 is the parameter space for voltage regulation for the NLC main damping ring. The beam current I_b, which is twice the total DC current, is plotted as a function of the cavity tuning angle ϕ_z, which is the angle between the total current and total cavity voltage and is a measure of how far off resonance the cavity is being driven. The steady-state full current operating point is shown by the circle. The shaded area shows Robinson's region of instability [Robinson 1962] in the high-current limit. (The low current limit corresponds to $\phi_z > 0$.) The dotted curves are contours of constant loading angle, ϕ_l; along the line $\phi_l = 0$ the reflected power is minimized. The solid curve corresponds to a maximum total klystron output power of 1.2 MW. In the steady-state, the cavity voltage will be maintained at the nominal value of 1 MV for operation anywhere within the region bounded by the klystron power curve and the lower right quadrant of Figure 4-60. Operation outside this region results in loss of regulation of the cavity voltage, due to insufficient klystron power, and eventually beam loss. Transient loading or klystron nonlinearities may reduce the stable operating region. The klystrons are specified to deliver a maximum of 1.2-MW peak power to ensure linearity of the klystron response.

There are two possible schemes for injecting to full current. To minimize the total reflected power in the injection process, the tuning angle ϕ_z could be preset just before each train is injected such that $\phi_l = 0$ after the train is injected. This corresponds to moving along the $\phi_l = 0$ line in Figure 4-60. However, this scheme is time consuming since it is limited by the slow motion of the tuners. A faster scheme involves one preset of the tuners to ${\phi_z}^*$, when the beam is absent, followed by consecutive injection of all 4 trains. While the reflected power is not insignificant during initial detuning without beam, it is within the specifications of the circulators. In the interest of maintaining a high duty cycle and allowing for maximum flexibility and compatibility to the remainder of the accelerator, the fast injection scheme has been developed. (However this does not preclude the alternative.)

Two techniques for correcting for the beam-induced transients at injection have been studied. Based on recent experiences and success at the SLC, the use of direct feedback was considered first. While it is possible to inject all the trains, there are a number of disadvantages introduced by the loop:

- The required gain places a tight tolerance on the total delay (less than 210 ns for an open loop gain of 2.5) and hence on the delay time in the klystron.
- The required gain introduces coupled-bunch modes that are supported by the cavity.
- Because the transient is corrected $after$ it is detected, residual cavity voltage and phase oscillations still persist.

The second technique is based on pre-conditioning the cavity phase and amplitude based on knowledge of the incoming charge. Using this feedforward algorithm, the total cavity current and voltage as well as the phase angle between them are held constant during the injection process. The process is depicted using phasors in Figure 4-61. Here V_c is the total cavity voltage (which is the vector sum of the generator and beam voltages), I_c is the total cavity current, ${I_b}^{(k)}$

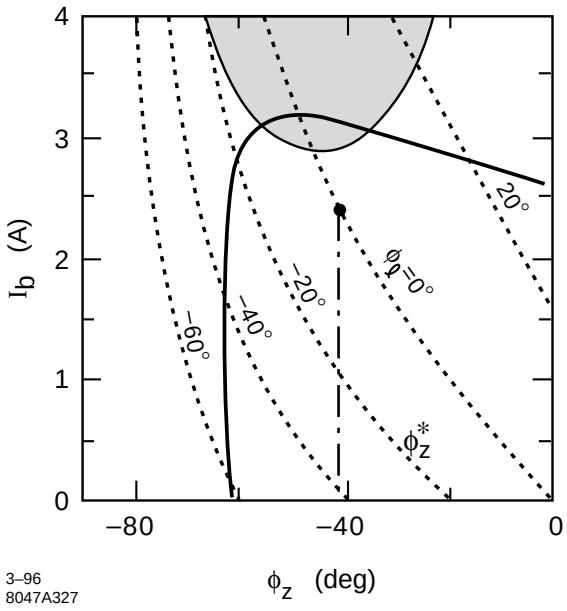


Figure 4-60. *Parameter space for voltage regulation for the main damping ring with 1.2 MW rf power. Here, the beam current $I_b = 2I_{dc}$ is plotted versus the tuning angle ϕ_z. The shaded region is that limited by Robinson instability and the dotted lines are contours of constant loading angle. The nominal operating point is shown at $\phi_l = 0$.*

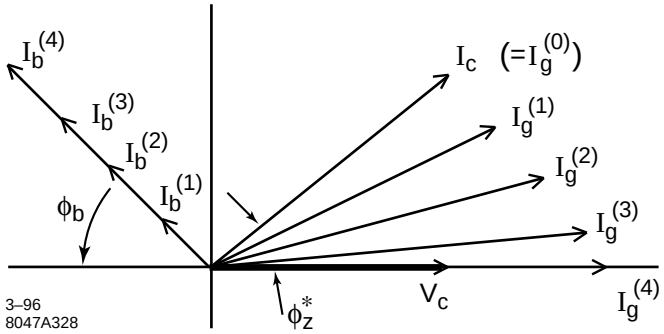


Figure 4-61. *Feedforward algorithm illustrated using phasors.*

and $I_g{}^{(k)}$ are respectively the beam and generator currents for train k, ϕ_b is the synchronous phase angle measured with respect to the crest, and $\phi_z{}^*$ is the cavity tuning angle for which the loading angle is zero at full current.

The principle of the injection procedure, using feedforward, is as follows (for the case of a zero-to-full-current ramp):

1. With beam off, deactivate tuner loops and detune for $\phi_l = 0$ at maximum current; that is, set the tuner position to $\phi_z{}^*$.
2. Set the external references (see Figure 4-59) $V_{\rm ref} = SI_g$ for the desired cavity voltage, where S is the transconductance of the klystron, and $\phi_{\rm ref} = \phi_z{}^*$.
3. Inject train 1.

4. Increment $V_{\rm ref}$ by $\Delta V_{\rm ref} = S\Delta I_g$ and $\phi_{\rm ref}$ by $\Delta\phi_{\rm ref}$ and inject train 2.
5. Repeat step 4 until all trains have been injected.
6. Reactivate tuner loops.

This procedure can be executed at maximum duty cycle so that the total time for injection is 4 cycles, or 22.2 ms at 180 Hz. Similar procedures are used for initial turn on with low beam currents or with low repetition frequencies. These are described in the section on the machine protection system (MPS) in Chapter 16.

The required changes in the reference voltage and reference phase are:

$$\Delta I_g = \frac{V_c/R + {I_b}^k \cos {\phi_b}^k}{\cos {\phi_b}^k} - \frac{V_c/R + {I_b}^{k+1} \cos {\phi_b}^{k+1}}{\cos {\phi_b}^{k+1}} \quad , \tag{4.64}$$

and

$$\Delta\phi = \tan^{-1}\left[\frac{\tan {\phi_z}^* + \frac{{I_b}^k R}{V_c} \sin {\phi_b}^k}{1 + \frac{{I_b}^k R}{V_c} \cos {\phi_b}^k}\right] - \tan^{-1}\left[\frac{\tan {\phi_z}^* + \frac{{I_b}^{k+1} R}{V_c} \sin {\phi_b}^{k+1}}{1 + \frac{{I_b}^{k+1} R}{V_c} \cos {\phi_b}^{k+1}}\right] \tag{4.65}$$

A potential systematic error associated with feedforward is that the cavity voltage will be regulated to the specified value, $V_{\rm ref}$; if the incoming charge is of different current than anticipated, then the amplitude feedback loop would regulate to an incorrect voltage.

Simulations of Low-Level rf System

Numerical simulations of the low-level rf system were performed by modifying existing software used in SLC rf system analysis [Minty 1995b]. In the simulations, the parameters of Table 4-20 are used. The model includes two (linear) klystrons, two rf cavities, and various feedback loops which include realistic nonlinear elements such as time delays and filters. At present, the cavity is modeled as a narrow-band resonator so that only beam interactions with the fundamental cavity mode are analyzed. Synchrotron radiation damping has not been taken into account.

The simulated response of the cavity voltage and beam phase to the injection of the first train is shown in Figure 4-62 for three cases: no transient loading compensation (solid), correction with direct feedback using a loop gain of 2.5, and correct phasing of the loop to minimize the transients (dot-dashed), and correction with feedforward (dashed). The results are summarized in Figure 4-63, in which the peak-to-peak cavity voltage and phase oscillations are plotted as a function of train number using direct feedback (a) or using feedforward (b). As can be seen, the correction of the beam-transients is substantially improved with the feedforward algorithm.

Figure 4-64 shows simulation results for the injection procedure outlined above for a zero to full current ramp. The feedforward references, as calculated from Eqs. 4.64 and 4.65, are also shown, along with the required generator and reflected powers. Notice that the Robinson damping time is shorter for the higher the current; this arises because of the fixed tuning angle. Also worth noting is the decrease in the coherent synchrotron oscillation frequency with current; as the beam current is increased the phase of the beam with respect to the generator crest becomes smaller resulting in a reduced restoring force for coherent phase perturbations. This effect must be taken into account in the feedback loop designs. Further analysis of stability tolerances imposed by the feedforward algorithm are in progress.

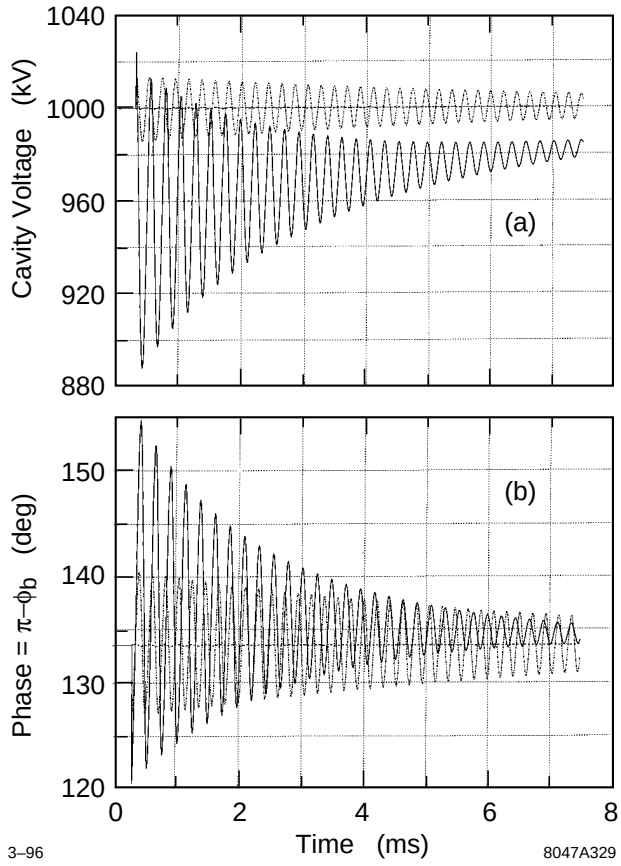


Figure 4-62. *Simulations of injection transients in the cavity voltage (a) and beam phase (b) at injection of the first train. The largest oscillations arise without feedback or feedforward; the direct rf feedback reduces the amplitude of both the voltage and phase oscillations but reduces the damping of the oscillations; the feedforward system reduces the oscillations to essentially zero.*

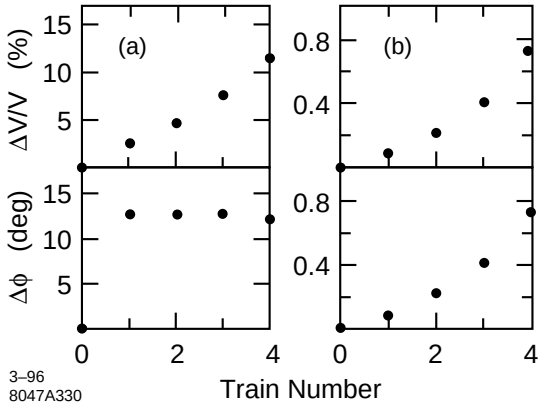


Figure 4-63. *Simulated peak-to-peak cavity voltage and phase oscillations as a function of train number. The transients are corrected either with direct feedback (a) or feedforward (b).*

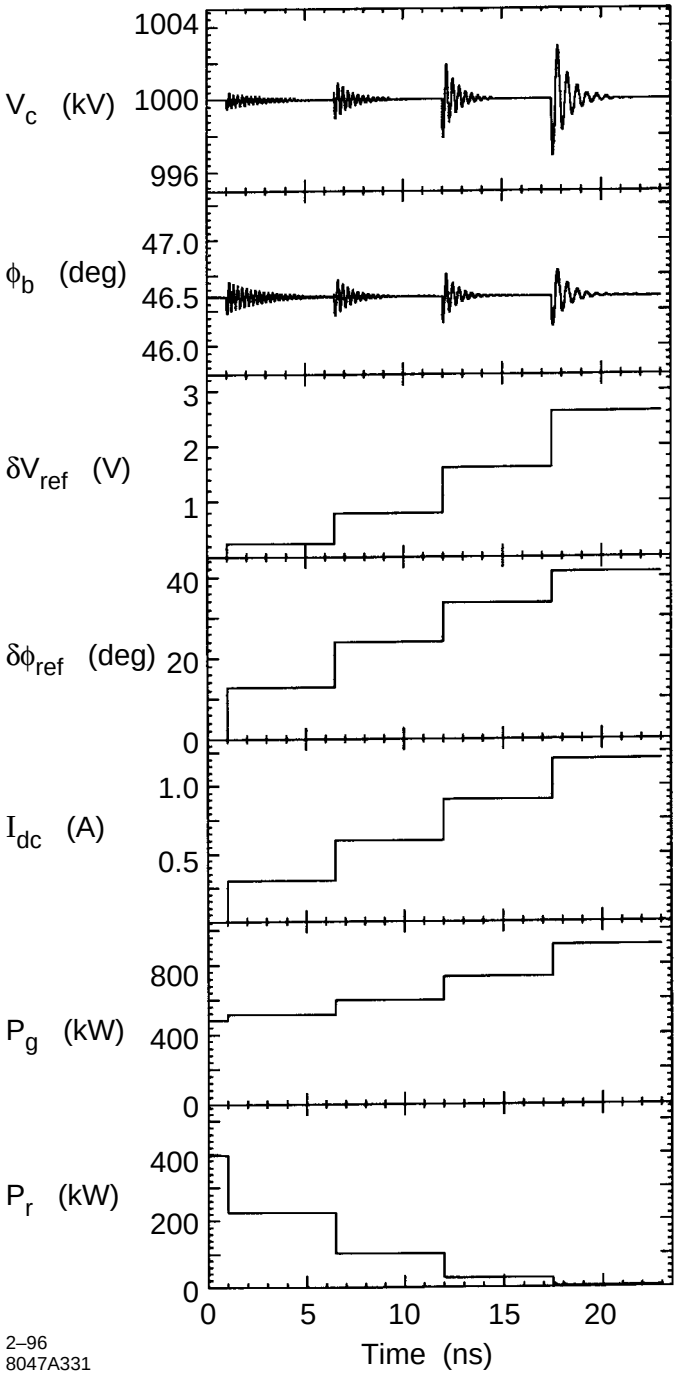


Figure 4-64. *Simulated current ramp from no beam to full current. Plotted are the cavity voltage V_c [kV], the beam phase ϕ_b [deg], the change in reference voltage $V_{\rm ref}$ [V], the change in reference phase $\phi_{\rm ref}$ [deg], the DC beam current I_b [A], the generator power P_g [kW], and the reflected power P_r [kW].*

	Electron	Positron
f_{rf} (MHz)	2856	1428
V_{rf} (MV)	42	80
L_{rf} (m)	2.5	4
Gradient (MV/m)	21	20
Structure Type	DDS	Detuned
Multibunch loading Comp.	ΔT	ΔT

Table 4-23. *Parameters of the rf systems for the electron and positron energy compressors.*

4.5.4 Energy Compressors

The electron and positron injection lines include rf sections that are used to reduce the incoming beam energy spread and energy jitter. Parameters of the energy compressors are given in Section 4.3.3. The electron compressor uses an S-band rf system while the positron system uses an L-band system; the L-band rf system is needed for the positron line because the beam emittance and beam size are too large for the apertures in the S-band structures.

The structures are shortened versions of those used in the electron and positron source linacs and are described in Chapter 6. The parameters for the injection line rf structures are listed in Table 4-23. In both cases, S-band and L-band, we use the ΔT (early injection) multibunch beam-loading compensation technique described in Chapter 6. This is the preferred method since it does not require additional off-frequency structures or klystrons. The technique can compensate for the full range of beam current by adjusting the klystron output.

4.6 Vacuum Systems

4.6.1 Main Damping Ring

The main damping ring vacuum system has been designed to attain a vacuum pressure less than 10^{-9} Torr. Because of the small vacuum chamber aperture and the large amount of synchrotron radiation, attaining such a pressure by simply pumping the beam chamber is unreasonable. Just to to attain a pressure of 10^{-8} Torr in the arcs without an ante-chamber would require placing a 400 L/s pump every 20 cm.

For these reasons, the arc and wiggler vacuum chambers have ante-chambers where the pumping is performed. At this time, we have a fairly advanced design for the arc chamber—this has allowed us to make accurate estimates of the ring impedance—and are working on the design of wiggler vacuum system.

Arcs

The NLC arc vacuum chamber is patterned after the Advanced Light Source (ALS) vacuum system. The vacuum chamber in each 4-m-arc cell will be constructed from two chamber sections and connected with a single bellows to the next arc cell chamber. The chambers will be machined in two halves from aluminum and then welded. Constructing

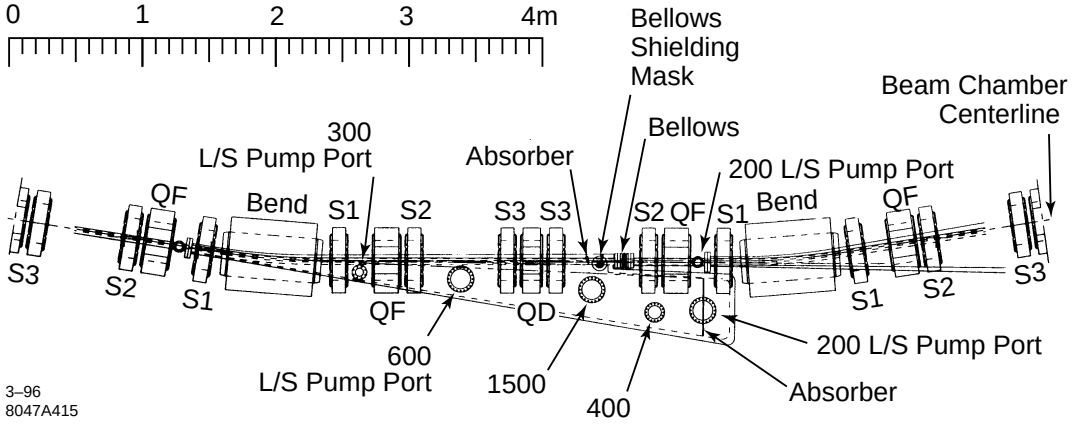


Figure 4-65. *Arc vacuum chamber in the NLC main damping ring.*

the chamber from machined-aluminum rather than extrusions readily allows for complicated shapes while maintaining very high construction tolerances.

The arc beam chamber is circular with a 12.5-mm radius and a 5-mm-high slot that couples it to an ante-chamber. Because of the small chamber radius, the 5-mm slot subtends a relatively large angle and allows 99.8% of the synchrotron radiation power to escape the beam chamber. This is actually a greater percentage of the radiation than escapes the ALS vacuum chamber which uses a 1-cm slot. Of course, because of the smaller dimensions, the alignment and construction tolerances will be tighter than those on the ALS chambers.

The chamber of a single arc section is illustrated in Figure 4-65 and a cross-section of the vacuum chamber inside a quadrupole magnet is shown in Figure 4-66. The chamber will be machined to accommodate the magnet poles and the arc magnets are designed as "C" magnets, allowing the ante-chamber to exit the sides of the magnets. The minimum wall thickness is 2.5 mm at the quadrupole poles, leaving 0.5 mm between the chamber and the magnet poles for alignment purposes. In the magnets, the 5-mm coupling slot is tapered to 8 mm in the ante-chamber and the wall thickness is increased to 11 mm.

The primary photon stop is at the end of the ante-chamber. The ante-chamber is designed with two vertical aperture constrictions to which reduce the flow of desorbed gas from the photon stop to the beam chamber. This allows for differential pumping which greatly reduces the pumping requirements. The vertical cross-section of the chamber in the photon stop region is shown in Figure 4-67.

The design of the beam chamber bellows is illustrated in Figure 4-68. The design is based on the ALS bellows. A smooth rf coupling within the bellows is achieved by maintaining the thin BeCu flex band under tension between either end of the internal plates. Thermal expansion of the beam pipe (1 mm for an 11°C in the 4-m chamber sections) is easily accommodated by fold-over fingers bridging a slot between the end of the flex band and the upstream end flange. During installation, when larger compression is required, the flex band is allowed to bow outward from the beam chamber centerline.

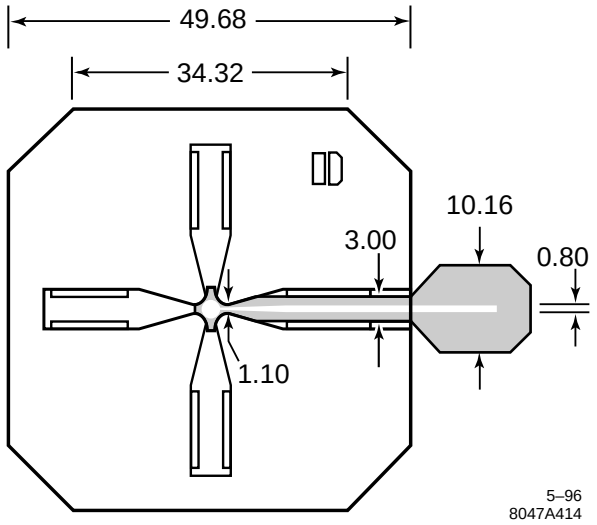


Figure 4-66. *Arc vacuum chamber in a quadrupole magnet.*

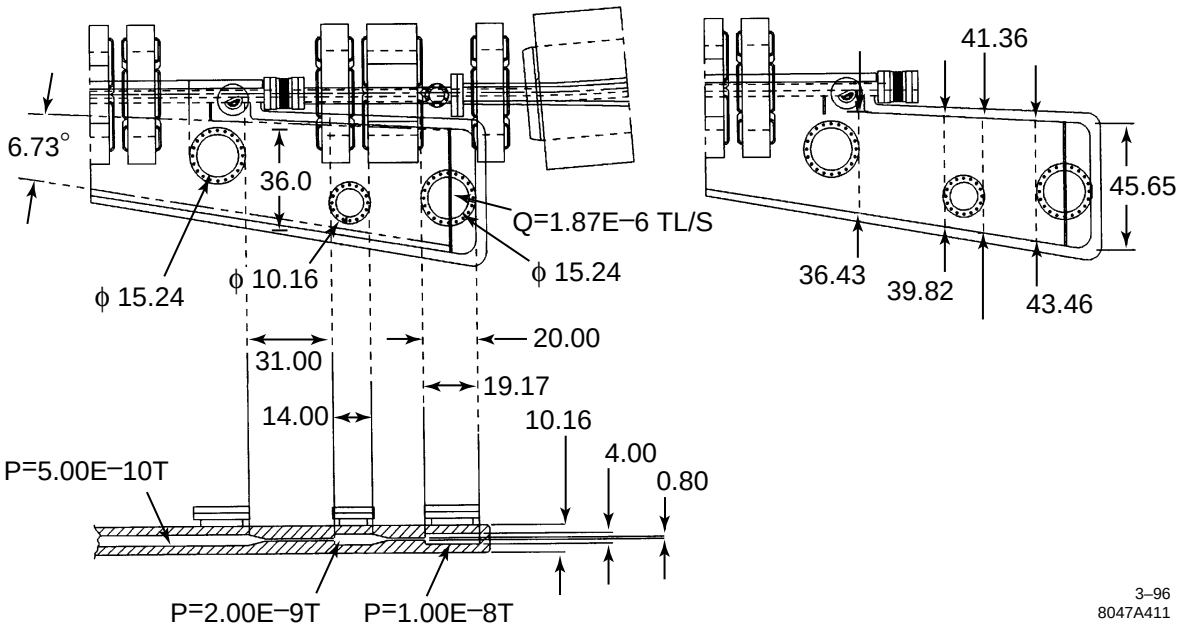


Figure 4-67. *End of arc cell chamber with photon stop and bellows and associated mask. The vertical ante-chamber profile shows aperture constrictions to reduce pumping requirements.*

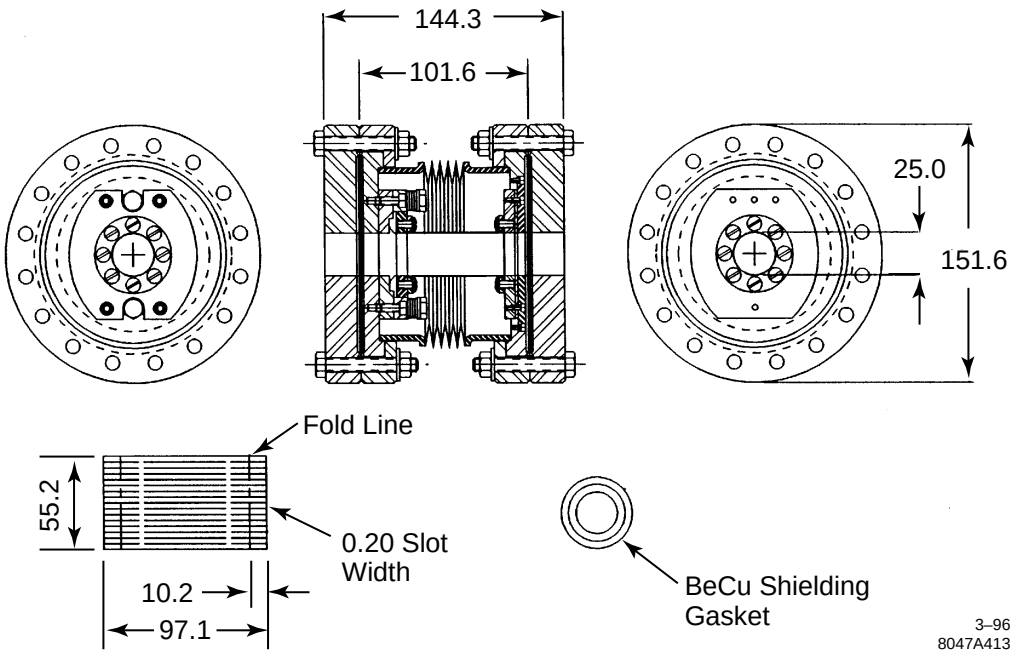


Figure 4-68. *Arc vacuum chamber bellows design.*

The rf liner in the bellows has the same dimensions and profile as the beam chamber, and, to simplify the design, the slot and ante-chamber are terminated just upstream. A mask which extends 2 mm into the beam chamber shields the bellows from the synchrotron radiation. The linear power density on the bellows mask and absorbers is kept just under 100 W/cm, which compares favorably with PEP and PEP-II values. The impedance of the slot termination, mask, and bellows have been calculated using MAFIA and are discussed in Section 4.4.1.

The initial vacuum design was based on NonEvaporable Getter (NEG) vacuum pumps that were supplemented with small ion pumps. Presently, we are considering using titanium sublimation pumps similar to those that are used in the ALS and will be used in the PEP-II Low Energy Ring. Assuming a desorption rate of 2×10^{-6} molecule per photon (thermal desorption is negligible), the total pumping per cell to attain a vacuum pressure of 10^{-9} Torr is 3200 L/s. The pressure profile in the beam chamber is illustrated in Figure 4-69.

Finally, it is anticipated that the arc vacuum system will baked once before installation. At this time, we have not decided if it is necessary to install the capability for *in-situ* baking. Catastrophic vacuum vents are rare. If one does occur, it is likely that the the chamber could be scrubbed using the intense synchrotron radiation from the beams. Of course, while the chamber is being scrubbed, the ring performance would likely be very poor, but the scrubbing should not take very long. This is the procedure that is used in the SLC damping rings and it was also the procedure used in the ALS to recover after the only unplanned vacuum vent experienced thus far.

Wiggler Regions

The wiggler straight section vacuum chamber is approximately 35-m long. It extends inside the arc from one side of the straight section to allow the absorption of the synchrotron radiation from the last wiggler magnets. The vacuum

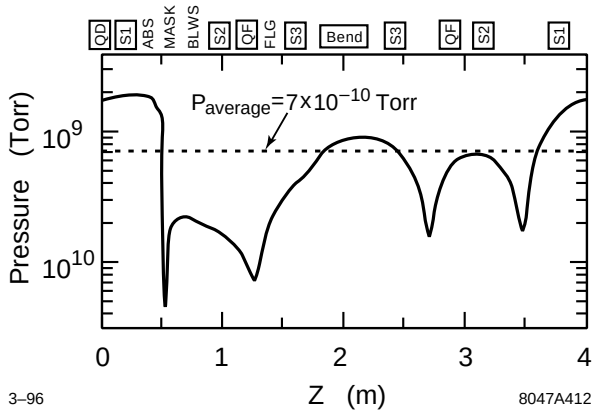


Figure 4-69. *Vacuum pressure in the arc beam chamber.*

chamber is aluminum and is designed to absorb approximately 400 kW of synchrotron radiation power. The cross section of the vacuum chamber together with a cross section of a wiggler magnet and the supporting frame is shown in Figure 4-70. As is seen from the picture, the vacuum chamber has a beam chamber that is roughly 45 mm by 16 mm and an antechamber on either side with a slot height of 10 mm. The antechamber slots are sloped at 0.25 radian angle, forming a water cooled dump to absorb the synchrotron radiation generated by the wiggler magnets. The vacuum pumping is performed by distributed Nonevaporable Getter (NEG) pumps which are located behind the dump surfaces at the end of each slot.

The design of the distributed NEG pumps is a departure from conventional NEG designs and follows the approach adopted for a vacuum chamber of the wiggler straight section of the Low Energy Ring of PEP-II B-Factory [Heim 1995]. Here, an SAES ST707 strip is cut into individual `wafers' using a laser sheet metal cutting machine. The cutting is performed under an argon purge to prevent oxidation. The NEG wafers are then slipped over a 9-mm diameter steel tube which is a combination of vacuum barrier and support rod. Between each wafer there is a 2-mm-thick stainless-steel wire-ring spacer. Each pump contains approximately 11 m of NEG strip per meter of length. This provides a distributed pumping speed of 385 liters per second per meter with a getter capacity of 1.1 Torr-liters per meter. The NEG pump is activated by a commercial tubular heater inserted inside the tube. The heater is outside of the vacuum space and it doesn't require an electrical feed-through. In addition, replacement of a failed heater does not require venting the wiggler vacuum chamber. The stainless steel rf screen separates the dump area from the NEG pumps. The purpose of the screen is to block rf TE modes from propagating down the slots and heating the NEG. The distributed NEG pumps will be supplemented by discrete ion pumps to pump hydrogen during NEG activation and methane during machine operation.

The wiggler vacuum chamber will be fabricated from two aluminum plates. Two pieces make up the top and bottom halves of the chamber. Each plate is machined on the outside first to form the water cooling passages and the magnet pole pockets. After the water cooling passages have been TIG closure welded, the interior aperture and dump surfaces are machined. A second TIG welding operation is required to join the top and bottom plates and to attach an aluminum-to-steel adapters to each end. Finally, the stainless steel end flanges are welded to the adapters.

The chamber will be vacuum baked at 150° for one or two days, then argon gas glow-discharge cleaned and, finally, coated with the layer of the titanium nitride to reduce photoelectron emission.

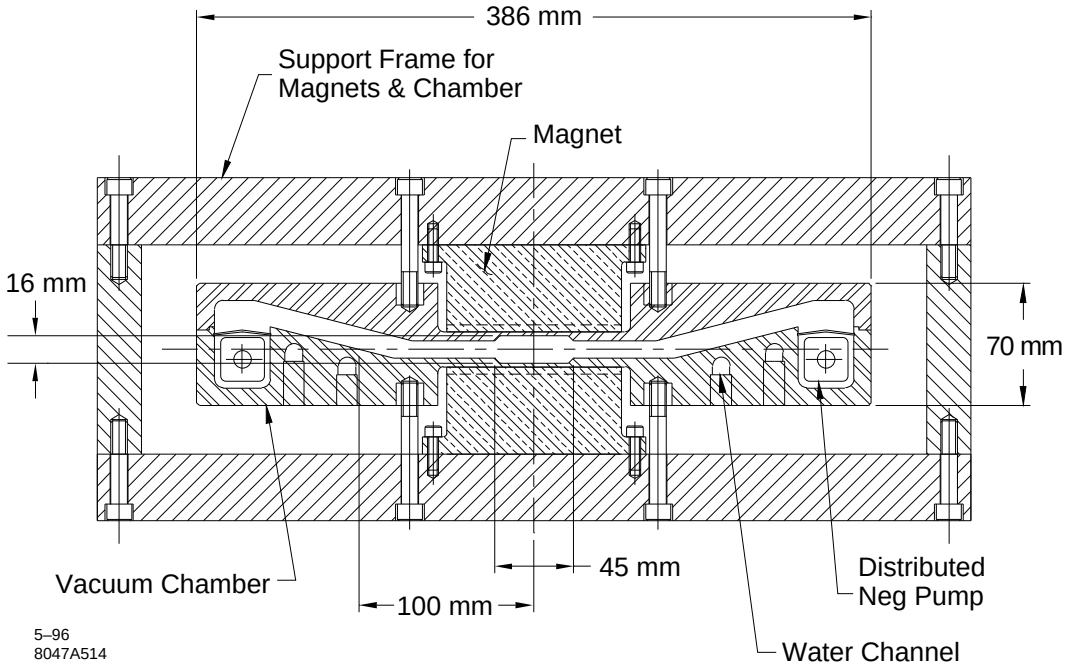


Figure 4-70. *Schematic of the cross section of the wiggler vacuum chamber, the cross section of the wiggler magnet, and the supporting frame.*

Injection/Extraction Region

At this point, we do not have a design for the vacuum system through the injection/extraction region of the main damping ring but we do not expect it to present significant difficulties. In concept, the arc ante-chamber will be terminated after the last matching bend. The circular 1.25-cm-radius beam chamber will continue for roughly 3 m to the rf cavities. In the present design, the aperture is increased before the cavities in a 15-cm taper to a 3.1-cm radius. After the two rf cavities, the chamber would taper back to the 1.25-cm radius to pass through the ceramic chamber of the extraction kicker system. Following the extraction kickers, the chamber would be increased to a 1.6-cm radius and then would become elliptical as the horizontal half width is increased to 2.5 cm to allow for the extracted beam. The chamber will maintain this elliptical profile between the two septa and then taper back down to pass through the injection kickers. Finally, the arc ante-chamber will resume at the matching bend.

Effort will be needed to maintain smooth transitions from one cross-section to the next. We will also need to consider how to minimize the impedance due to the septa which will be mounted in large vacuum canisters. Finally, we will need to design the coated ceramic chambers for the kickers and the numerous pumping ports that will be required to maintain the 1-nTorr vacuum pressure in this conductance-limited system.

4.6.2 Pre-Damping Ring

The vacuum system for the pre-damping ring has not been designed yet but the vacuum requirements are relatively loose; a pressure of 10^{-8} Torr will be sufficient. It has been assumed that the chamber would be constructed from

aluminum like the main damping ring chambers. Of course, since the tolerances are much looser, it would probably be formed by extrusion rather than be machined from aluminum stock.

As noted in Section 4.3.2, the chamber aperture in the arcs must have inner dimensions of 6 cm by 3.6 cm to accommodate the edge emittance of 0.09 m-rad and leave $\pm$2-mm clearance for steering. Although it is not required from the vacuum point of view, an ante-chamber will probably be used though the arcs because it eases the handling of the radiation power. In addition, the ante-chamber would reduce the number of secondary electrons which may cause the positron-electron instability discussed in Section 4.4.7. To further reduce the number of secondary electrons, the aluminum chamber could be coated.

4.6.3 Transport Lines

The vacuum system in the injection, extraction, and transfer lines has not been completed. The system must include valves to isolate the damping rings from the source linacs and the bunch compressors. The vacuum pumping is primarily needed to counter the thermal outgassing. In these sections, the vacuum pressure should be less than 10^{-8} Torr.

4.7 Feedback and Feedforward Systems

The damping rings will require numerous feedback and feedforward systems. Some of the rf feedback systems were described in the section on the low-level rf control, Section 4.5.3. This includes feedback loops around the klystrons and cavities as well as timing feedback loops from the bunch compressors and feedforward loops from the sources. In this section, we describe the transverse coupled bunch feedback systems and some of the transverse feedbacks needed to maintain the beam trajectory. In addition, we will review the requirements on some of the other global feedback and feedforward systems.

4.7.1 Coupled Bunch Transverse Feedbacks

As described in Section 4.4, the transverse coupled-bunch growth rates in the main damping rings require feedback damping. This section describes a bunch-by-bunch feedback for the main rings that is similar to that planned for the PEP-II rings.

The fastest growing coupled-bunch modes in the main damping ring are those driven by the transverse resistive-wall impedance. Because the pre-damping ring vacuum aperture is much larger than that in the main damping rings, the resistive-wall impedance is not expected to be significant and thus a coupled-bunch feedback system will probably not be necessary in the pre-damping ring; this will be verified in the future.

Main Damping Ring Feedback Design

The coupled-bunch growth rates calculated for the main damping rings indicate that a feedback system will be required to maintain stability of transverse beam oscillations. It is expected that the longitudinal oscillations will be stable and will not require a feedback system.

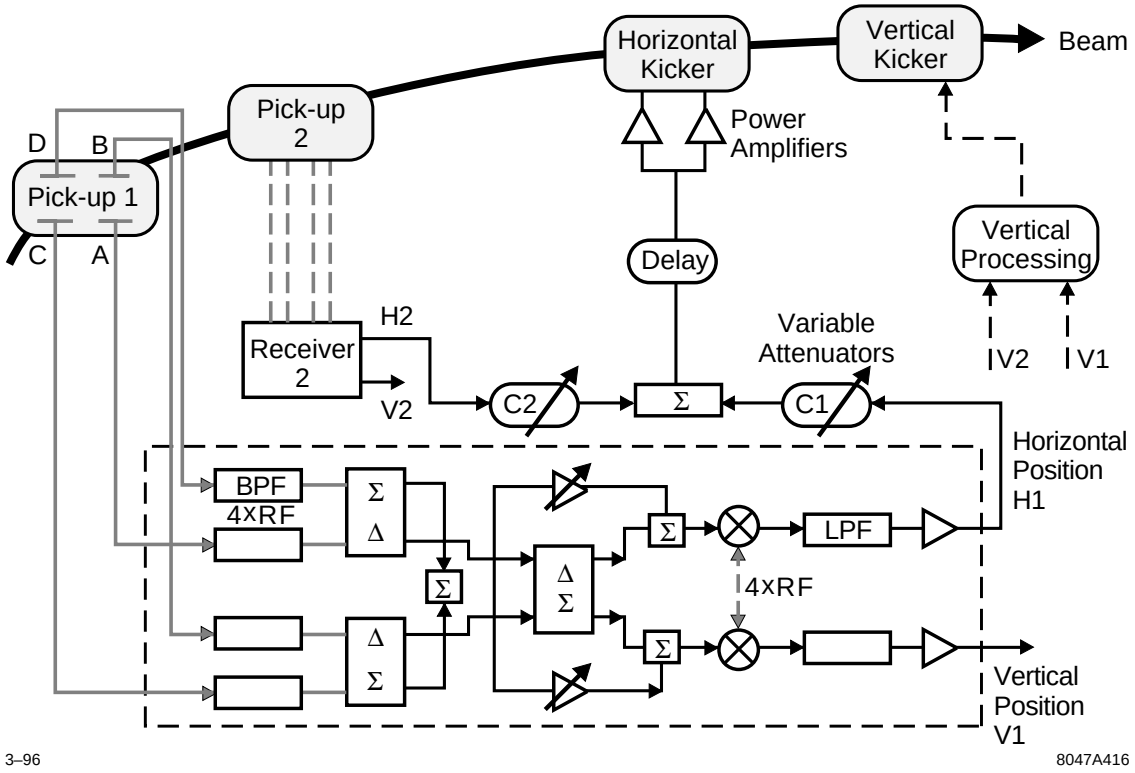


Figure 4-71. *Schematic of the coupled-bunch feedback system.*

As shown in the calculations of coupled-bunch growth rates, the transverse impedance excites a fairly broadband spectrum of beam modes, although the dominant modes are driven by the resistive-wall impedance up to about 50 MHz. We propose a bunch-by-bunch system that acts independently on each bunch, damping all possible coupled–bunch modes. For a bunch spacing of 1.4 ns, this implies that the minimum system bandwidth be at least 357 MHz. This bandwidth requirement poses the most difficult design restrictions on the system.

The transverse feedback system (TFB) can, in principle, be operated in any one of many frequency bands (0–357 MHz, 357–714 MHz, etc.). To minimize expensive wide-band high-power driver amplifier power, we chose to operate in the 0–357-MHz frequency band, where the impedance of a stripline-pair kicker is greatest. In addition, the kicker impedance is largest at the lower frequencies of the dominant coupled-bunch modes.

The design of the main damping ring TFB system follows the design of a system recently commissioned at the Advanced Light Source [Barry 1993] and one planned for PEP-II [Barry 1995]. The main damping ring TFB system would differ only in the frequency bandwidth and the total power required. A block diagram of the system is shown in Figure 4-71. Two sets of button pickups, located approximately 90° apart in betatron phase, are used to detect beam moments $I\Delta x$ and $I\Delta y$. The individual button signals are processed at 2.84 GHz ($4\text{x}f_{rf}$) to benefit from greater button sensitivity at higher frequency, combined via hybrids, and demodulated to baseband. The baseband-moment signals are combined in proportion according to the Twiss parameters at the pickups and the phase advance from pickups to kicker. The signal is delayed the appropriate amount for coincidence with the bunch at the kicker. Not shown in the figure is an element for rejection of an orbit offset at the pickup. The offset rejection can implemented in several ways, for example, as a notch filter.

As mentioned, the kicker is a 50-Ω stripline pair operated in difference mode with a length of one-half the bunch spacing (21 cm). The kicker impedance is given by

$$R_{k,\perp} = 2\ Z_L \left(g_\perp \frac{l}{b} \frac{\sin(\omega l/c)}{(\omega l/c)} \right)^2 \quad , \tag{4.66}$$

where $Z_L = 50\ \Omega$, l=21 cm, and b= 1.5 cm is half of the stripline separation, and $g_\perp$ is the coverage factor and assumed to be close to unity. This configuration yields a kicker impedance of 19.6 $k\Omega$ (at $\omega = 0$). The broadband power required is given by

$$P_{k,\perp} = \frac{V_k^2}{2R_{k,\perp}} \quad , \tag{4.67}$$

where V_k is the kicker voltage. Specification of this voltage is discussed in the following section on computer simulations.

Computer Simulations

This section describes computer simulations of the coupled-bunch oscillations and the corresponding feedback system used to damp the oscillations. The purpose of performing these simulations is twofold. The first is to determine the growth rates for the actual filling pattern which are difficult to calculate analytically. The second is to empirically determine how much power is required for the system to satisfactorily damp coupled-bunch oscillations. To this end, we can include effects such as injection beam transients, pickup and kicker noise, DC orbit offsets, etc.. In particular, the simulations include the limiting of the feedback response at larger beam oscillation amplitudes. The results described here represent a first attempt to find an adequate feedback voltage for damping injection beam offsets assuming a perfect feedback system; further optimization will be performed in the future.

The beam dynamics in this study are simulated using common tracking techniques at a single point in the ring, where difference equations are used to describe the discrete time evolution of the beam oscillations and wake voltages [Siemann 1983, Thompson 1989]. This approximation is valid for all of the effects we wish to study, especially since most of the wake voltages are localized at the rf cavities. One notable exception is the distributed wake of the resistive-wall impedance, which results from the skin effect on the inner vacuum chamber wall. The local approximation for the resistive-wall wake is valid when the corresponding growth or damping rate is slow compared to the betatron oscillation period. This condition is true for all cases studied here. Tracking is done in both transverse directions, which we generically label x. The turn-by-turn difference equations for x and $x^{'}$ for a linear lattice are given by

$$x_{i+1} = \cos 2\pi Q_x x_i + \beta_x\ \sin 2\pi Q_x x_i^{'}(1 - 2\lambda_x) \tag{4.68}$$

$$x_{i+1}^{'} = -\frac{1}{\beta_x} \sin 2\pi Q_x x_i + \cos 2\pi Q_x x_i^{'}(1 - 2\lambda_x) + \frac{\Re(\tilde{V}_{w,i+1})}{E} \tag{4.69}$$

where the tracking is assumed to be at a symmetry point with $\beta^{'} = 0$ and zero dispersion. The tunes and β functions are assumed to be constant as a function of amplitude and energy. Amplitude-dependent tunes generally lower the effective coupled-bunch growth rates but here we are restricting ourselves to a worst-case scenario with no nonlinearities.

The transverse wake voltage from dipole HOMs is found by summing the contributions from all previous bunch passages. Tracking of the wake voltage is done by treating it as a complex phasor, $\tilde{V}_w$. The difference equation for $\tilde{V}_w$ is given by

$$\tilde{V}_{w,i+1} = \tilde{V}_{w,i} e^{(j\omega_r + \frac{\omega_r}{2Q})\Delta t} - j2k_\perp q\ x_{i+1} \tag{4.70}$$

where $k_\perp$ is the transverse loss parameter, ω_r is the angular resonant frequency, and x is the transverse offset at the location of the HOM.

As mentioned above, the resistive-wall wake voltage is the only non-localized wake of concern. The approximate localized wake voltage for a round pipe is given by

$$V_{w,rw}(t) = qx\frac{lc}{\pi b^3}\sqrt{\frac{\mu_0\rho}{\pi}}\frac{1}{\sqrt{t}} \tag{4.71}$$

where l is the length of the vacuum chamber (usually the ring circumference), ρ the material resistivity, and b the chamber radius. The voltage for bunch passage must be remembered and individually recalculated each iteration. The resistive-wall wake is cutoff after four turns.

Main Damping Ring Feedback Systems

The TFB system measures a transverse position error and produces the appropriate angular kick to correct the error which is applied to the beam on a subsequent turn. Because the tracking in this simulation is done at only a single point in the ring, we are currently using a linear combination of the transverse position measured on the two previous turns to calculate the correct kick given by

$$V_{fb} = G_{fb}E\left(\frac{x}{\beta_x \tan\phi_x} - \frac{x_{-1}}{\beta_x \sin\phi_x}\right) \tag{4.72}$$

where G_{fb} is the TFB gain and usually ranges from 0.03–0.3, x and x_{-1} are the positions on the current and previous turns, and $\beta_\perp = 6.5$ m in the NLC simulations. The feedback voltage limits at the maximum amplifier output voltage.

A simulation result for the case of 4 trains of 90 bunches, filled to a total beam current of 1.2 A, is shown in Figure 4-72. The injected train is offset at injection by 500 μm; with the assumed value of $\beta_\perp = 6.5$ m, this corresponds to an injection offset that is $1.2\sigma_{x,y}$. Figure 4-72 shows the transverse offset of all bunches versus turn number, and Figure 4-73 shows the corresponding feedback voltage. The feedback voltage was limited to 1500 V for this case. The injected bunches saturate the feedback and damp linearly in time until the amplifier comes out of saturation, whereupon they damp exponentially. The stored bunches are excited to very small amplitudes but are quickly damped.

A maximum voltage of 1500 V corresponds to broadband power of 60 W. Although this seems like a rather modest requirement, it is reasonable considering the relatively small offsets at injection. We expect the power to scale roughly as the square of the offset so considerably more power is needed for larger offsets. Furthermore, we have not yet considered the effects of residual orbit offset and noise which also consume power.

4.7.2 Stored Orbit Feedbacks

In addition to the coupled-bunch feedback system just described and the rf feedbacks described in Section 4.5.3, we plan to implement transverse position feedbacks in the damping rings. These will be used to maintain the beam orbit to prevent variations in the extracted beam trajectory and to prevent emittance growth or dynamic aperture reduction due to orbit shifts. Similar feedback systems are installed in most synchrotron light sources where the stability requirements are severe because of the long distances from the radiation source point to the experimental stations.

Because the revolution rate in the rings is roughly 1.3 MHz, these feedback systems are not limited by the sample rate like most other beam-based feedbacks in the linear collider. Of course, they are still limited by the corrector slew rate which depends upon the corrector magnets, the power supplies, and the vacuum chamber. Since most destabilizing effects are relatively low frequency, we do not believe that the feedbacks need to respond to frequencies above 30 Hz. In this case, the correctors can be simple air-core dipole magnets mounted around the normal vacuum

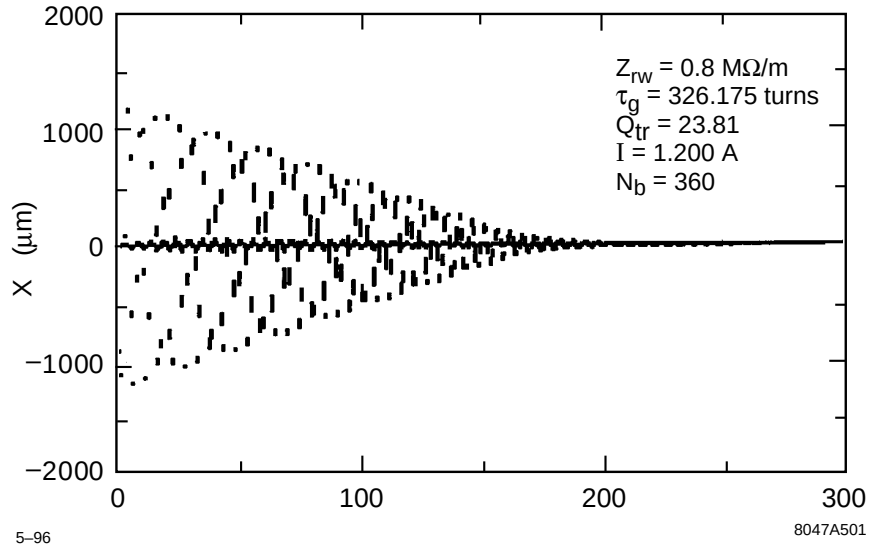


Figure 4-72. *Bunch position after a train is injected into a main damping ring with a 500-μm offset; it is difficult to resolve the individual bunches in the trains.*

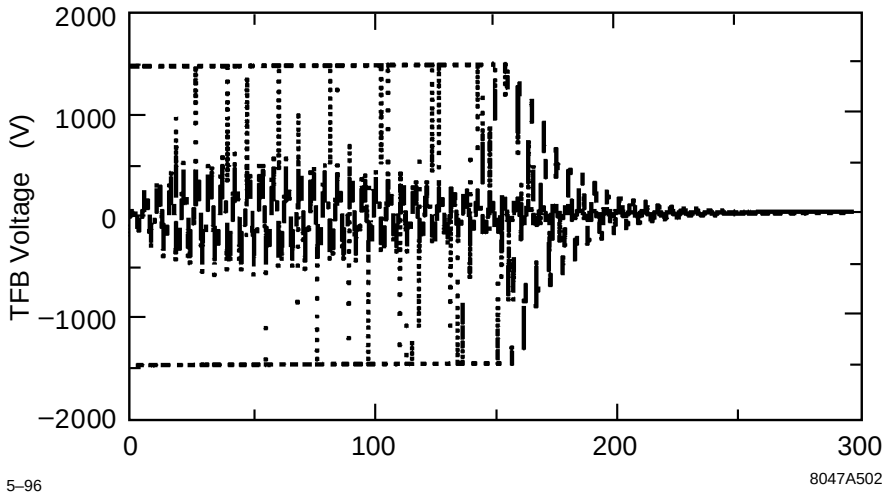


Figure 4-73. *Feedback voltage after a train is injected into a main damping ring with a 500-μm offset (1.2 sigma); notice that the feedback saturates at roughly 1.5 kV.*

chamber; they would be placed in the injection and extraction region where the vacuum chamber does not have an ante-chamber. Finally, although the feedback systems will be much more important in the main damping rings than in the pre-damping ring, they will likely be used to improve the operational stability in all three rings.

4.7.3 Injection Feedbacks

In addition, to the transverse position feedbacks in the damping rings, we will use fast beam-based feedbacks in the injection lines to stabilize the incoming trajectories, energy, and intensity:

- Bunch-by-Bunch Energy
- Bunch-by-Bunch Intensity
- Bunch Train Energy Spread
- Transverse Phase Space

These systems are similar to those used throughout the SLC and described in Appendix D.

4.7.4 Extraction Feedbacks

As in the injection lines, we will use fast beam-based feedbacks in the extraction lines to stabilize the outgoing trajectories. In addition, there will be a fast multibunch kicker system to correct for small (static) bunch-to-bunch position errors arising from the damping ring extraction kickers—these kickers are described in Chapter 7. There will also be a feedback for the damping ring phase based on the extracted beam trajectory through the bunch compressor; this is described in Section 4.5.3.

Finally, there will be an intensity feedforward system that will adjust the beam loading compensation in the linacs to fluctuations in the bunch train charge; a similar system is used in the SLC to stabilize the electron beam energy. The bunch train charge will be measured near the end of the store cycle. Without a feedforward system, the tolerance on variations in the bunch train charge is 0.3% to limit the energy deviation at the IP to 0.1%. It is expected an intensity feedforward system will reduce the bunch train intensity jitter tolerance by over a factor of five by adjusting the linac beam loading compensation to the measured beam charge. In addition, the feedforward system would automatically trigger the beam dumps, either at the end of the linac or at the end of the first or second bunch compressor, if the deviation in bunch train charge exceeds specified limits; such a system is needed to limit backgrounds in the detector at the IP.

4.8 Vibration and Stability

4.8.1 Vibration

Vibrations of the quadrupole magnets due to the ground motion will cause the beam trajectory to vary. The two dominant sources of vibrations are ground motion that is transmitted and possibly amplified by the magnet supports,

and a man-made noise, *e.g.*, generated by cooling water flow in the magnets [Turner 1995]. The latter, however, is more difficult to analyze, because much will depend on water cooling system and magnet supports design. Regardless, with proper design, the effect of cooling and other man-made sources should be minimal. For example, in the ALS synchrotron light source, the magnet cooling was measured to increase the magnet vibration by roughly 1 nm [Greene 1992]; motion at ten times this amplitude is still acceptable.

The effect of the vibrations on the beam trajectory depends crucially on the correlation from magnet to magnet. For high frequencies, when the wavelength of the ground waves is smaller than the distance between the quadrupole magnets, the magnet vibrations are uncorrelated. In this case, each magnet contributes independently in inverse proportion to the square of its focal length to the perturbation of the closed orbit. In the intermediate range corresponding to the wavelengths larger than the distance between the magnets, but smaller then the dimensions of the ring, one can expect resonant effects in the dependence of the closed orbit distortion on the frequency [Rossbach 1988]. At very small frequencies, when the wavelength becomes much larger than the size of the ring, the main components of the ground motion will be a rigid displacement and a tilt of the ring, which do not distort and displace the closed orbit relative to the magnets. However, the amplitude of the ground vibrations grows sharply when the frequency decreases.

The vertical perturbation of the closed orbit can be characterized by its deviation $y_B(s)$ from the ideal trajectory. We denote the value of this function at the location of ith quadrupole magnet by y_{Bi}. It is given by

$$y_{Bi} = \sqrt{\beta_i} \sum_i G_{ij} y_j \quad , \tag{4.73}$$

where β_i is the beta function and y_i is the displacement of the ith quadrupole, and the coefficients G_{ij} are computed from the known parameters of the lattice.

The matrix elements G_{ij} should satisfy the following sum rules,

$$1 = \sqrt{\beta_i} \sum_j G_{ij} \quad , \tag{4.74}$$

$$\left\{ \begin{array}{c} x_i \\ z_j \end{array} \right\} = \sqrt{\beta_i} \sum_j G_{ij} \left\{ \begin{array}{c} x_j \\ z_j \end{array} \right\} \quad . \tag{4.75}$$

The first of these two equations is due to the fact that if the entire ring shifts vertically as a rigid body by one unit, the closed orbit moves together with the ring by the same distance. The second equation expresses that the tilting of the ring in either x or z direction would also cause the same tilting of the orbit. These two sum rules are particularly important in the limit of low frequencies when the main components of the ground motion are the shifts and the tilts of the ring. Equations 4.74 and 4.75 guarantee that they do not produce distortions of the closed orbit relative to the magnets, and only the deformations of the shape of the ring would contribute to the closed orbit perturbation.

In order to eliminate rigid shifts and tilts of the closed orbit which go together with the ring and do not move the beam trajectory in the vacuum chamber, we consider the average square of the difference between the orbit and the magnet displacements,

$$\delta y_i^2 = \left\langle \left(y_{Bi} - y_i \right)^2 \right\rangle . \tag{4.76}$$

Using Eq. 4.74, we find that

$$\left\langle \delta y_i^2 \right\rangle = \sum_i G_{ij} G_{ik} \left[\left\langle y_i^2 \right\rangle - 2 \left\langle y_i y_j \right\rangle - \left\langle y_j y_k \right\rangle \right]. \tag{4.77}$$

The expectation value of the mixed product $\langle y_i y_j \rangle$ can be expressed in terms of the two-dimensional power spectrum of the ground motion, $P(\omega, k)$,

$$\left\langle y_i y_j \right\rangle = \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty P(\omega, k) \cos(k l_{ij}) \frac{dk}{2\pi} \quad , \tag{4.78}$$

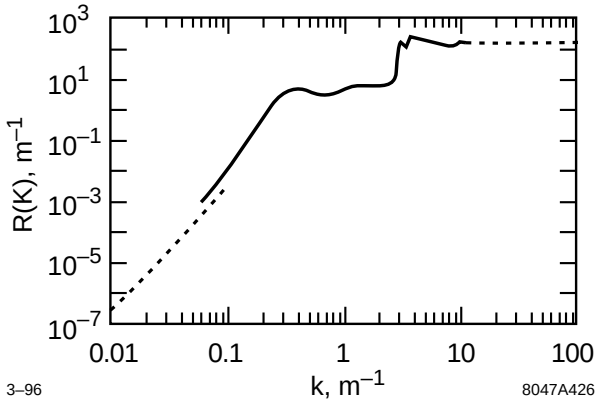


Figure 4-74. *Plot of the averaged response function for the NLC damping ring.*

where l_{ij} denotes the distance between the ith and the jth quadrupoles. We use the following representation for $P(\omega, k)$, $P(\omega, k) = P(\omega)\,\mu(\omega, k)$, where $P(\omega)$ is the integrated power density of the ground motion, and $\mu(\omega, k)$ describes the correlation of the ground motion at different points. In the model developed in Appendix C.2, this function is equal to

$$4 \Big/ \sqrt{\omega^2 / v(\omega)^2 - k^2} \quad , \tag{4.79}$$

where v denotes the velocity of the ground waves. In our calculations, we used a simpler model assuming that

$$\mu(\omega, k) = \begin{cases} 2\pi v(\omega)/\omega, & k < \omega / v(\omega) \\ 0, & k > \omega / v(\omega) \end{cases} \quad , \tag{4.80}$$

where $v(\omega)\ [\mathrm{m/s}] \approx 450 + 1900 \exp(-\omega/4\pi)$, and ω is measured in s^{-1} [Adolphsen 1995]. Qualitatively, this dependence tells that the contribution to the function $\mu(\omega, k)$ at the frequency ω comes, with equal weights, from the waves with the wavenumbers below $\omega / v(\omega)$ only. For the original model in Appendix C, the weights are not equal, with a larger contribution coming out from the vicinity of $k = \omega / v(\omega)$.

The final expression for δy_i^2 takes the form

$$\left\langle \delta y_i^2 \right\rangle = \beta_i \int_0^\infty \frac{d\omega}{2\pi} P(\omega)\, R_i\left(\frac{\omega}{v(\omega)}\right) \quad , \tag{4.81}$$

where $R_i(k)$ is the response function for the lattice at the location of the ith quadrupole. We average this function over the ring and denote it as $R(k)$; it is plotted in Figure 4-74.

For large ω, ($\omega > 2\pi \times 250\ Hz$), R approaches the high-frequency limit of uncorrelated quadrupole vibrations. It is approximately constant in this region with a value of about $R \approx 160\ m^{-1}$ shown by the horizontal dashed line in Figure 4-74. For very small frequencies, one can prove that $R \approx Ak^4$, where $A \approx 30\ m^3$; this asymptote is also shown in Figure 4-74 by a dashed line.

Assuming a spectrum of ground motion [Juravlev 1995]

$$P(\omega) \left[\frac{\mu m^2}{Hz}\right] \approx \frac{1.6 \times 10^{-3}}{\omega^4} \quad , \tag{4.82}$$

where ω is given in s^{-1}, we found the following value for δy_i upon integration in Eq. 4.81 from zero to infinity,

$$\sqrt{\langle \delta y_i^2 \rangle} = 0.14 \text{ nm} \quad . \tag{4.83}$$

This is negligibly small in comparison with a typical vertical size of the bunch of the order of 5 μm.

The high frequency limit of the function R can easily be translated into the tolerance for the uncorrelated motion of the quadrupoles in order the beam trajectory vary less then 10% of the vertical size of the beam. For the nominal emittance $\gamma\varepsilon_y = 2 \times 10^{-8}$ m-rad, the uncorrelated vibration of the quadrupoles should not exceed 18 nm.

Finally, if the trajectory motion is too large, one can use a global orbit feedback to stabilize the trajectory in the ring. Because of the high-revolution frequency, this is a straightforward feedback system (see Section 4.7).

4.8.2 Stability

In addition to the high frequency motion of the ground, there are alignment and electronic drifts that need to be controlled. To maintain the vertical emittance the closed orbit must not drift by more than 10 μm. This implies a tolerance on the magnet alignment and, more importantly, on the BPMs and BPM electronics. The primary sources of drift are thermal fluctuations and slow movement of the ground.

To estimate the effect of the ground fluctuations, we use the ATL-model which predicts that the rms difference in the displacement of any two points is given by $\Delta y^2 = A\,T\,L$ where A is a coefficient, T is the time, and L is the distance between the two locations. Using a coefficient of $A = 5 \times 10^{-7}\ \mu\text{m}^2 s^{-1} m^{-1}$, which is an upper limit on the value that has been measured at SLAC (see the discussion in Appendix C), we find that the stored orbit will vary by 10 μm after roughly 24 hours. This is not a significant limitation—it is reasonable to expect to correct the closed orbit every 24 hours. Furthermore, this trajectory variation would be substantially reduced by a global orbit feedback as discussed in Section 4.7.

Of course, it does not do any good to re-correct the closed orbit if the BPM readbacks are drifting due to either electronic variations or physics drifts. This is where the tolerance on the thermal stability is very important. To minimize the BPM drifts, we will anchor the BPMs to the magnets so that they are constrained transversely and temperature in the ring tunnel will be maintained to a fraction of a °C, while that in the electronics racks will be stabilized to 1°C. With this thermal regulation, we would hope to limit the BPM drifts to less than 10 μm over 24 hours.

4.9 Alignment and Supports

The main damping rings and downstream systems all have very tight alignment and field tolerances to prevent dilution of the transverse emittances. In all cases, these tolerances will be attained using beam-based alignment techniques where the alignment accuracy depends on beam measurements and not mechanical systems; the techniques to be used in the damping rings are described in Section 4.9.3.

To attain the desired vertical emittances and dynamic aperture, the magnets in the main damping rings must be aligned to the level of 50∼100 μm in the vertical plane and $100 \sim 200\,\mu$m in the horizontal plane; these tolerances are specified in Sections 4.3.8 and 4.3.9. In addition, the vacuum chamber must be aligned at the level of ∼500 μm to prevent synchrotron radiation from hitting the inside of the beam chamber. Of course, these tolerances refer to the alignment over a relatively short distance (1 or 2 cells) and not globally.

To attain these tolerances, we plan to rely on beam-based alignment of the quadrupoles and sextupoles. Unfortunately, one cannot do the same for the combined-function bending magnets; these need to be aligned using conventional techniques to 200-μm rms (see Section 4.3.8).

In the downstream systems (X-band linacs, final focus, etc.), the alignment will be implemented using remote magnet movers. In the damping rings the tolerances are sufficiently loose that there is some question as to whether magnet movers are required; once the alignment is known, one could mechanically perform the differential moves to attain the desired alignment. Mechanical alignment could only be considered if the alignment is stable at the 50-μm level for months—experience from the ALS and calculations using the "ATL" model would suggest that this is the case. Regardless, the mover question will be determined in the future; for this design, we have assumed magnet movers.

4.9.1 Girders and Supports

The girder and support systems are still being designed. The concept is similar to that adopted in the collimation section and the final focus. The supports would be mounted on a low concrete girder attached directly to the floor. The magnet movers would then be mounted on the girders and aligned using shims. Finally, the magnets would be mounted on the movers.

The movers could be similar to the linac movers, described in Chapter 7. These are based on the SLAC Final Focus Test Beam (FFTB) magnet movers and have a range in excess of ± 1 mm in x and y with a 0.3-μm step size. Of course, the damping rings do not such high precision and a simpler design would probably be adopted.

4.9.2 Mechanical Alignment

Mechanical alignment will be used to pre-align the quadrupoles and sextupoles in the damping rings and to align the elements that cannot be aligned using beam-based alignment, *i.e.*, the bending magnets, vacuum chamber, and rf cavities. The required alignment is 200 μm, which is well within the capabilities of modern equipment.

4.9.3 Beam-based Alignment

Attaining the required magnet alignment, in the main damping rings, through purely mechanical systems will be difficult. For this reason, we plan to use beam-based alignment to determine the positions of the quadrupoles and sextupoles with respect to the beam. Quadrupole beam-based alignment is being used at MAX-lab [Rojsel 1994], LEP [Barnet 1994], ALS [Robin 1995] and SSRL. In all these rings, it is possible to adjust the fields in each quadrupole individually, either with separate power supplies or with additional trim windings. All techniques vary the quadrupole field and measure the orbit variation; the magnitude of the orbit change is proportional to the quadrupole-to-beam offset.

The actual algorithm used to deduce the offsets is somewhat different in each of the laboratories, but once the quadrupole-to-beam offset is determined, it is referenced to the adjacent BPMs. The quadrupole-to-BPM offsets are then used to either re-align the magnets or to steer the beams through the center of the quadrupoles.

How accurately one can measure these quadrupole-to-BPM offsets depends upon several quantities: the reproducibility of the BPMs, stability of the correctors and other magnets, and physical motion of the machine. At the ALS, the

quadrupole-to-BPM offset can be measured with a repeatability that is better than ± 5 microns [Robin 1995]. This is at the level of the BPM noise and is well within the necessary tolerances required in the NLC.

Sextupole beam-based alignment is being used at KEK [Kamada 1994] and at DESY [Herb 1995]. In these measurements, a local orbit bump is made in a section of the ring containing one sextupole. All the sextupoles in that family are then varied as a whole and the change in the resulting orbit is measured. The amplitude of the orbit bump is changed and the measurement is repeated. The horizontal deflection then depends quadratically upon the orbit offset while the vertical deflection is linear in the offset:

$$\theta_x = 0.5 K_2 l (\Delta X + x_m)^2 \qquad \theta_y = K_2 l \Delta X y_m \quad , \tag{4.84}$$

where $K_2 l$ is the integrated sextupole strength, ΔX is the orbit offset, and x_m and y_m are the sextupole misalignments. By fitting the resulting oscillations, the sextupole center was inferred a resolution of several hundred microns in TRISTAN [Kamada 1994].

This process will be simplified in the NLC damping rings because all of the sextupoles will have individual power supplies and the BPMs will have 1-μm resolution. To estimate the resolution of technique, we assume that the orbit is moved in five steps by a total of ± 250 μm in the sextupole while measurements are made of the closed orbit with the sextupole on and off; it is necessary to turn the sextupole on and off to isolate the individual magnet—alternately, one could use a remote magnet mover to move just the sextupole to be aligned. By using all of the ring BPMs to fit the resulting betatron oscillations to a single kick at the sextupole location, the effective BPM resolution should decrease to roughly 0.1 μm. Of course, systematic errors will degrade this effective resolution and thus we will assume a 0.5-μm effective resolution. In this case, one should be able to determine the vertical alignment of the sextupole with a resolution of roughly 20 μm; this is clearly sufficient, although higher resolution could be obtained with a larger orbit step size.

Unfortunately, this technique will be slow. To align all of the sextupoles in each of the main rings will probably take many hours. There are several possible directions in which one may try to improve the speed of the measurement. First, putting an additional winding on four of the six poles of the sextupole would provide a quadrupole field [Raubenheimer 1995b, Kikuchi 1995]. This would allow one to use the quadrupole beam-based alignment techniques to find the quadrupole center which should be close to the sextupole magnetic center; this should be significantly faster. Second, one could vary multiple sextupoles at once and fit for a number of kicks at the same time; this will degrade the resolution but it will still probably be sufficient. Third, one could vary the sextupole strength and measure the tune variation. For example, turning off a sextupole in which the beam is horizontally offset by 50 μm will cause tune changes of roughly $\Delta\nu \approx 0.001$—easily measurable. Unfortunately, applying this technique to the vertical alignment is more difficult, although it may be possible by operating on the difference coupling resonance and measuring the resulting tune separation. All such combined approaches require additional research.

At this point, we have not estimated the frequency with which beam-based alignment would have to be used. If it is required often (daily), the alignment system must be highly automated and the magnets would need to be mounted on movers. Fortunately, the alignment tolerances are relatively loose (compared to the rest of the collider) and it is expected that alignment would only need to be performed at the beginning of a run. Experience at the ALS suggests that the alignment is stable to better than 50 μm over periods of six months or more [Robin 1995].

One of the important problems that needs to be resolved is how the BPMs are mounted to the magnets. At the ALS, the BPMs are mounted in the vacuum chamber and this is free move transversely. Thermal fluctuations then cause the BPM positions to drift by as much as 50 μm. Designs have been suggested that allow the BPMs to be mounted in the chamber but will constrain the transverse movement at the magnets; this is clearly desirable.

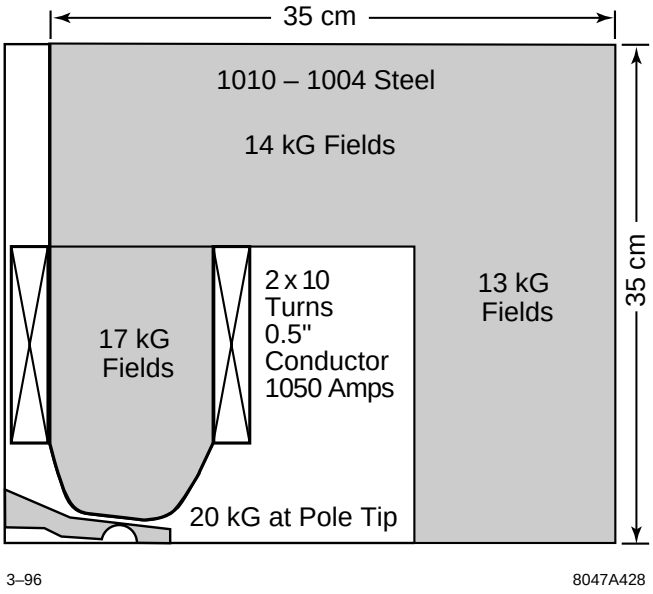


Figure 4-75. *Main damping ring "C" bend magnet cross section.*

4.10 Magnet Design and Power Supplies

4.10.1 Main Damping Rings

Bending Magnets

The required field parameters of the bending magnets are listed in Table 4-24 for nominal energy of 1.98 GeV. The magnet is a high-field combined-function magnet with a 15-kGauss field and a relatively strong quadrupole gradient. The full aperture at the beam pipe location is 3.2 cm and the required good field region is ± 0.5 cm.

To determine the feasibility of the bending magnets, we completed a first pass at the design of an "H" magnet using the 2D computer code POISSON and the 3D code TOSCA. Although the "H" style magnet seemed possible, the required fields and the geometry naturally lend themselves to a "C" style magnet. Thus, we have modeled a similar design with POISSON based on the "C" geometry; this magnet is illustrated in Figure 4-75 with a blowup of the pole in Figure 4-76.

In both cases, the required fields are attained using standard 1010 magnet iron with peak fields in the pole tip of 19.6 kGauss. The fields are sufficiently high that saturation effects cause the relative quadrupole and higher multipole components to vary as a function of excitation. The deviation of the magnetic field from a simple dipole field with a linear gradient is plotted in Figure 4-77 for three different excitation currents. The variation of the relative quadrupole is $\sim$1% but it, as well as the variation in the sextupole component, is easily compensated; the higher-order multipoles are not significant.

The magnet is a sector magnet and, because the sagitta is substantial (1.3 cm), the magnet pole must be curved along the beam trajectory. At this time, we have not considered construction techniques although it seems most likely that

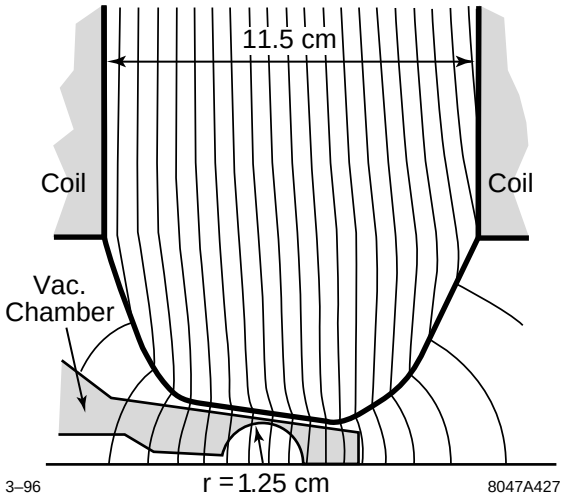


Figure 4-76. *Main damping ring "C" bend magnet pole.*

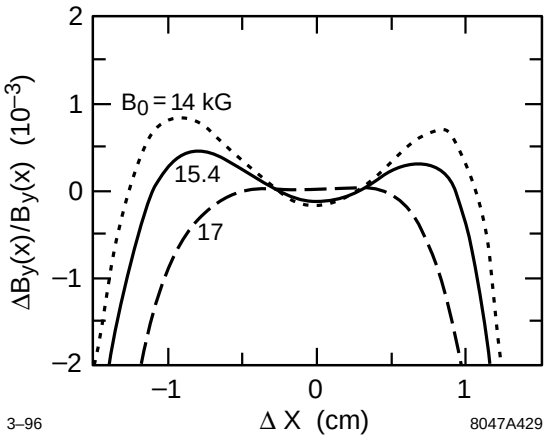


Figure 4-77. *Main damping ring "C" bend magnet multipoles versus excitation for ring operation from 1.8 GeV to 2.2 GeV.*

the magnets would be machined with solid cores to accommodate the curvature; the KEK ATF dipole magnets are being constructed in a similar fashion. We have estimated that a construction tolerance of 25 μm is needed to attain the required field quality.

The ends of the poles will be chamfered at a 60° angle. For convenience, we plan to use removable end pieces on the magnets. Here, the last 4 cm (or so) of the pole is removable to allow straightforward optimization of the fringing fields and future modifications; similar pole-ends are used on the SLC damping ring bending magnets. At this time, we have not studied the fringing fields at the magnet ends of the "C" magnet design. Although the "H" magnet design and other previous experience with combined function magnet design [Raubenheimer 1993a] leads us to believe that they should not be difficult, in the future, we will model the magnet in 3D and include the fringing fields in the dynamic aperture calculations.

Dipole Magnet Designs:	2
Main Bends:	
Number	38
Magnetic Length	68.4 cm
Half Gap	1.6 cm
Amp-turns	40,900
$B_{0\,\mathrm{eff}}$	15.3 kGauss
$B_1\ (dB_y/dx)$	126 kGauss/ m
Matching Bends:	
Number	4
Magnetic Length	40 cm
Half Gap	1.6 cm
Amp-turns	34,500
$B_{0\,\mathrm{eff}}$	13.0 kGauss
$B_1\ (dB_y/dx)$	107 kGauss/ m

Table 4-24. *Dipole magnet parameters for main damping rings at 1.98 GeV.*

It is also important to include the fringing field when calculating the radiation damping rate [Rivkin 1985]. To this extent, we have quoted the field in terms of $B_{0\,\mathrm{eff}}$ which is defined as:

$$B_{0\,\mathrm{eff}} \equiv \sqrt{\int_{-\infty}^{\infty} ds \frac{B_0^2}{L_B}} \tag{4.85}$$

and is the quantity that determines the synchrotron radiation damping. In general, the peak field in the center of the dipole will be slightly greater than $B_{0\,\mathrm{eff}}$ because of the fringe fields.

The main dipole magnet coils will be constructed from two layers, with 10 turns per layer, of 0.5-in square water-cooled conductor to provide a total of 45,000 A-turns for operation at 2.2 GeV. The matching bends have half the integrated field and gradient of the main bending magnets. To reduce the fringing fields, we increased the length and decreased the fields. Since the magnets have the same ratio of field to gradient as the main bending magnets, they would be constructed with the same cross section as the main bends but powered with 38,000, rather than 45,000 A-turns, for 2.2-GeV operation. At this time, we have not considered whether an additional trim coil will be needed to account for the different saturation behavior of the matching the main bending magnets.

All of the bends in each ring will powered by a single 100-kW power supply rated at 800 V and 1250 A. The stability of the main power supply is not very critical. Only low-frequency variations penetrate the vacuum chamber. Thus, the primary effect of bend power-supply fluctuations are slow shifts of the stored beam energy. This has three effects: (1) the bunch compressor converts the energy changes to changes in the beam phase and therefore energy fluctuations at the IP (this effect is discussed further in Chapter 5), (2) in regions of dispersion, the close orbit shifts, and (3) the betatron tunes change by an amount proportional to the uncorrected chromaticity.

In general, the most severe limitation is that on the extracted beam energy. To limit the fluctuations to 1/20 of the rms beam energy spread, the power supplies must regulate better than $\Delta I/I = 5 \times 10^{-5}$. This will cause tune variation at the level of 0.002 and 0.001 in the horizontal and vertical planes and it will cause the beam energy to fluctuate at the IP by roughly 0.05%. The horizontal and vertical orbit shifts at the extraction point, which are proportional to the residual dispersion, will be much less than 1/10 of the respective beam sizes.

Magnet	Quan.	K1 (m^{-2})	B_{pole} (kGauss)	Mag. Len. (cm)	Aper. (cm)
QFH (Arc Quad)	72	7.9	7.9	20	1.5
QDH (Arc Quad)	36	-6.0	6.0	15	1.5
QFHM (Arc Matching)	4	7.5	7.5	20	1.5
QDHM (Arc Matching)	4	-7.1	7.1	15	1.5
QM2 (Eta Matching)	4	8.9	8.9	20	1.5
QM3 (Inj/Ext Matching)	2	-5.3	5.3	15	1.5
QM4 (Inj/Ext Matching)	2	6.7	6.7	20	1.5
QM5 (Inj/Ext Matching)	2	-2.0	2.7	15	2.0
QM6 (Inj/Ext Matching)	2	2.0	2.7	15	2.0
QFX (Inj/Ext cells)	2	2.9	5.9	15	3.0
QDX (Inj/Ext cells)	2	-2.3	3.1	15	2.0
QM3W (Wiggler Matching)	2	-5.3	5.3	15	1.5
QM4W (Wiggler Matching)	2	6.3	6.3	20	1.5
QM5W (Wiggler Matching)	2	-0.4	0.8	15	3.0
QM6W (Wiggler Matching)	2	0.4	0.8	15	3.0
QFW (Wiggler cells)	2	2.8	5.6	15	3.0
QDW (Wiggler cells)	2	-1.0	2.0	15	3.0

Table 4-25. *Quadrupole magnets for NLC main damping ring vers. 6.11 at 1.98 GeV.*

Quadrupole Magnets

The required parameters of the quadrupole magnets are listed in Table 4-25 for the at the nominal energy of 1.98 GeV. Five basic magnet designs are needed:

1. A 20-cm magnetic length with a maximum 10.3-kGauss pole-tip field (at 2.2 GeV) and a 1.5-cm radius aperture for the dispersion matching.
2. A 20-cm magnetic length with a maximum 9.3-kGauss pole-tip field (at 2.2 GeV) and a 1.5-cm radius aperture for the arc QFs.
3. A 15-cm magnetic length with a maximum 8.2-kGauss pole-tip field (at 2.2 GeV) and a 1.5-cm radius aperture for the arc QDs.
4. A 15-cm magnetic length with a maximum 5-kGauss pole-tip field (at 2.2 GeV) and a 2.0-cm radius aperture for the injection and extraction region.
5. A 15-cm magnetic length with a maximum 7-kGauss pole-tip field (at 2.2 GeV) and a 3.0-cm radius aperture for the wiggler region.

The larger-radius quadrupoles are needed in the wiggler and injection and extraction regions to accommodate the wiggler region vacuum chamber and the injected and extracted beam trajectories. The five families are summarized in Table 4-26.

Quadrupole Magnet Designs:	5
Arc QFs:	
Number	80
Magnetic Length	20 cm
Inscribed Radius	1.5 cm
Max. $B_{\rm pole}$	9.3 kGauss
Main Coil	15 turns
Max. Current	410 A
η QFs:	
Number	4
Magnetic Length	20 cm
Inscribed Radius	1.5 cm
Max. $B_{\rm pole}$	10.3 kGauss
Main Coil	17 turns
Max. Current	410 A
Arc QDs:	
Number	44
Magnetic Length	15 cm
Inscribed Radius	1.5 cm
Max. $B_{\rm pole}$	8.2 kGauss
Main Coil	15 turns
Max. Current	350 A
Inj/Ext Quads:	
Number	6
Magnetic Length	15 cm
Inscribed Radius	2.0 cm
Max. $B_{\rm pole}$	5 kGauss
Wiggler Quads:	
Number	10
Magnetic Length	15 cm
Inscribed Radius	3.0 cm
Max. $B_{\rm pole}$	7 kGauss

Table 4-26. *Quadrupole magnet parameters for NLC main damping rings at 2.2 GeV.*

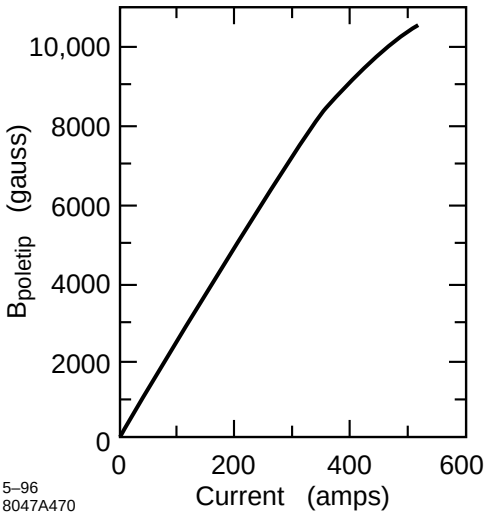


Figure 4-78. *Pole-tip field versus current for symmetric arc quadrupole design.*

As discussed in Section 4.3.9, we have assumed that the quadrupoles would be constructed in two pieces, as was done for both the ALS and APS synchrotron light sources; this will reduce the random multipoles arising from construction errors. Thus, we assumed an uncorrelated pole-to-pole error (significantly reduced relative to four-separate-pole construction) of 0.77 mils (20 μm) rms in the pole placement in the transverse plane and the two-pole correlated contribution which can characterized in terms of a relative translation (in the transverse plane) and rotation of the two halves. The former we assumed to be 1.5 mils (40 μm) in each of the two dimensions, and the latter to be 1.1 mr, corresponding to a linear displacement of 8.9 mils (225 μm) at a 20-cm arm length. The resulting multipoles are described in Section 4.3.9.

The design of the arc quadrupole magnets is further complicated by the vacuum ante-chamber. The ante-chamber is 3-cm high and in excess of 25-cm wide as illustrated in Figure 4-66. Thus, it is not reasonable to design the magnet around the chamber. We are pursuing two approaches of accommodating the ante-chamber: first, a symmetric design where iron is removed from the mid-plane on both sides of the magnet, and, second, the asymmetric case where iron is only removed from one side of the midplane. The former approach has been adopted at the APS storage ring and latter was adopted for the ALS ring. In both cases, the systematic magnet multipoles were corrected using shims to levels below our tolerances (see Section 4.3.9).

These two cases have been modeled with POISSON and the pole-tip field versus current for the symmetric case is plotted Figure 4-78.

The straight-section quadrupoles have larger apertures but are thought to be more straight-forward and we have not yet modeled them. In the injection/extraction magnets, there is no ante-chamber and the large bore wiggler magnets were sized to encompass the wiggler section ante-chamber easily; we will re-size these magnets after the wiggler vacuum chamber design is completed.

Each quadrupole will have an independent power supply which adds substantial flexibility to the lattice design and simplifies the beam-based alignment procedures since an additional trim winding is not needed on each quadrupole. The disadvantage of the independent supplies may be the additional cost.

The primary effect of quadrupole power-supply fluctuations are changes in the betatron tunes and the orbits. We have calculated tolerances on the power supplies assuming independent power supplies which have uncorrelated

Magnet	Quan.	K2 (m^{-3})	$B_{\rm pole}$ (kGauss)	Mag. Len. (cm)	Aper. (cm)
S1	76	-220	2.1	9	1.7
S2	76	+460	4.5	9	1.7
S3	76	-420	4.1	9	1.7

Table 4-27. *Sextupole Magnets for NLC Main Damping Ring at 1.98 GeV.*

Sextupole Magnet Designs:	1
Arc Sextupoles: Number	216
Magnetic Length	9 cm
Inscribed Radius	1.7 cm
Max. $B_{\rm pole}$	5.5 kGauss
Main Coil	10 turns
Trim Coil	29 turns
Max. Current	260 A

Table 4-28. *Sextupole magnet parameters for NLC Main Damping Ring.*

fluctuations. The tolerance to limit the fluctuations in the tunes to 0.001 is $\Delta I/I = 2 \times 10^{-4}$. Further assuming an rms orbit displacement in the quadrupoles of 500 μm, this power-supply tolerance will cause horizontal and vertical orbit fluctuations of 0.5% and 5% of the respective beam sizes. Finally, if the quadrupoles were on string supplies or if the power supply fluctuations are correlated, the magnitude of the orbit errors does not change (assuming random orbit offsets) but the sensitivity to the tune variation increases by roughly a factor of 10, making the tolerance $\Delta I/I = 2 \times 10^{-5}$.

Sextupole Magnets

The required parameters of the sextupole magnets are listed in Table 4-27 for the nominal energy of 1.98 GeV. All the sextupole magnets have similar parameters and would likely be constructed from the same design; these parameters are summarized in Table 4-28.

Like the quadrupole magnets, the design of the sextupole magnets is complicated by the vacuum ante-chamber. Because we would like to use trim coils in the sextupoles to implement dipole and possibly skew quadrupole correction, we are only pursuing the asymmetric, where iron is only removed from one side of the mid-plane, approach of accommodating the ante-chamber; this philosophy has been adopted at the ALS ring.

This magnet has been modeled with POISSON and a plot of the pole-tip field versus current appears in Figure 4-79. In addition, a drawing of the whole magnet is shown in Figure 4-80.

In addition to the main coils, we are planning to use trim windings on some magnets as dipole correctors. The present magnet modeling has include a single trim winding capable of producing pole-tip fields of roughly 600 Gauss; this is more than sufficient for the dipole correction.

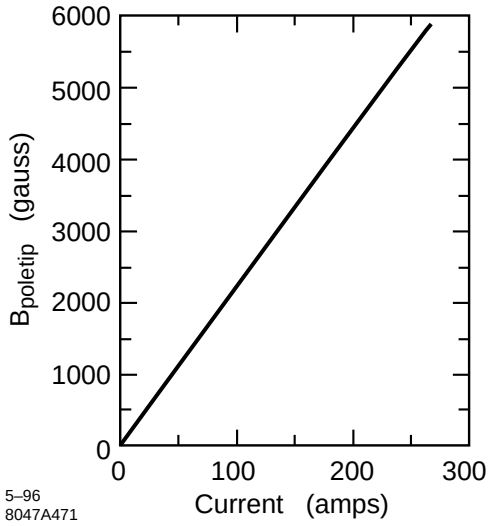


Figure 4-79. *Pole-tip field versus current for asymmetric arc sextupole design.*

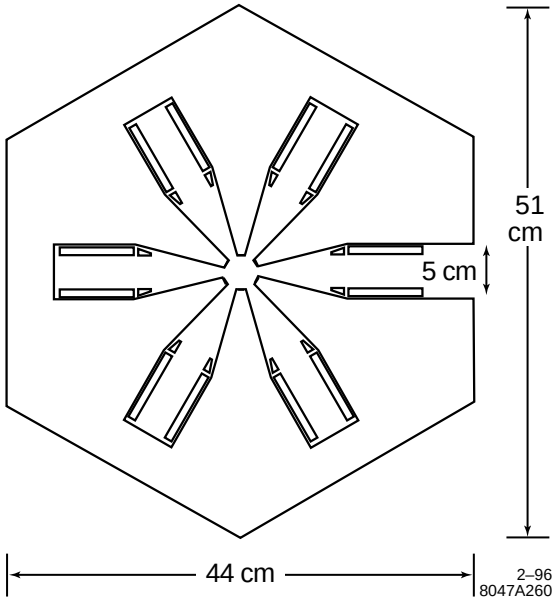


Figure 4-80. *Main damping ring sextupole design.*

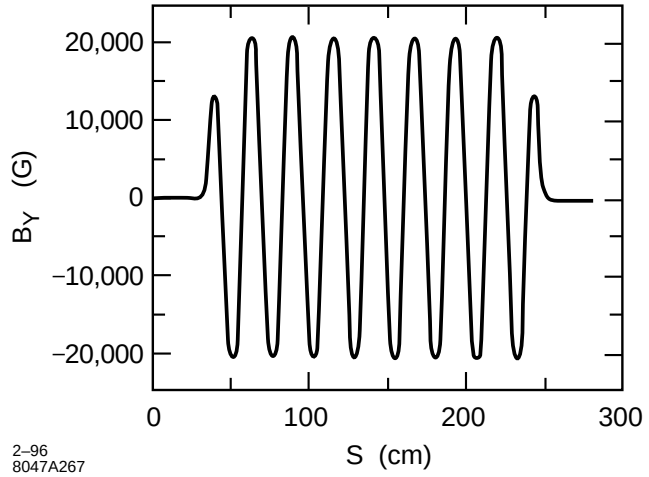


Figure 4-81. *Measured wiggler field in SSRL BL-9 wiggler [Baltay].*

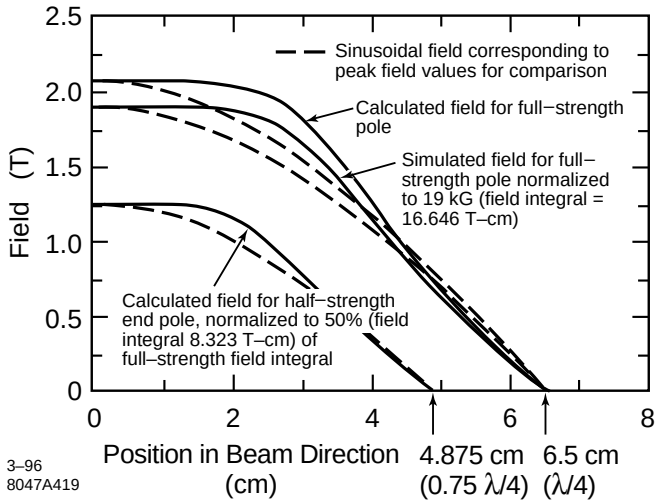


Figure 4-82. *Field from SSRL BL-9 wiggler using shaped poles [Baltay].*

Wiggler Magnets

The main damping rings require roughly 25 m of strong wiggler to attain the desired damping times. The two parameters that are relevant for a damping wiggler are the integral of $B_y^2(s)$, which determines the damping, and the quantum excitation which is determined by the wiggler period and the field. To achieve the desired vertical damping time of 4.63 ms, the integral of B_y^2 over the length of the wiggler must be 6200 kGauss^2-m, and to constrain the emittance dilution, the quantum excitation due to the wiggler should be less than that of the arcs.

We have chosen to consider a relatively short, high-field, short-period device. If we assume a sinusoidal field with a peak field of 22 kGauss (close to the saturation of Vanadium Permendur), then we need a length of 25.6 m and a period of roughly 25 cm. Using simple scaling laws, it is thought that such a device could either be constructed as an electro-magnetic wiggler or as a permanent magnet hybrid wiggler.

The SSRL Beam-Line 9 wiggler nearly meets our requirements [Baltay]. It is a hybrid wiggler with a length of 2.14 m, a 26-cm period, and a peak field of 20.5 kGauss. Because the wiggler poles were designed to optimize the total flux, the field is not sinusoidal and the integral of B_y^2 averaged over one period is within 1% of the required field in the main damping rings. When the matching poles at the ends of the wiggler are included, the field is within 3% of the required field. The measured field profile along the wiggler is plotted in Figure 4-81 and a comparison between the normal sinusoid field pattern and that of the wiggler is plotted in Figure 4-82 [Baltay].

Our goal in choosing the wiggler parameters was to minimize the length of wiggler needed and thereby the cost. It was thought that, for a given technology, the costs scaled with length. At a later time, it may be decided that this was not an optimal choice. In that case, we could increase the length, decreasing the peak field and increasing the wiggler period. For example, it might be more reasonable to consider a 30-m wiggler with a peak field much less than 20 kGauss, and a 30-cm period. Alternately, we could increase the ring energy—the required peak field scales as $1/\sqrt{\gamma}$.

Injection and Extraction Septa

For the injection and extraction scheme shown in Figures 4-13 and 4-14, there are 3 m of space for the septa. As an initial design, we have assumed a DC current sheet septa that has a 1-cm vertical aperture and is composed of two pieces: an 83-cm segment and a 1-m section; the two sections are separated by 15 cm. The initial section has a 5-mm-thick blade and a 2-kGauss field while the second section has a 15-mm-thick blade which would consist of three turns of conductor and a 6-kGauss field. Both sections would be powered in series with roughly 1800 A, producing a total bend angle of 115 mr. The current density in the blades is 36 Amps/mm^2 which is roughly one third the current density in the SLC damping-ring septa blades. The cooling channels have not been studied yet but the requirements are relatively loose; a coolant flow of roughly 1 GPM will limit the temperature rise to 6°C in the first section and 30°C in the second segment.

The dipole field quality along the injected and extracted beam trajectories is less than 2×10^{-4} over a radius of 3 mm; this is roughly 2 sigma of the injected beam. Finally, both septa require back-leg coils to decrease the field on the close orbit. A back-leg trim of 15 A-turns on the first segment will reduce the field on the closed orbit to much less than 0.5 Gauss with an insignificant gradient, and a back-leg trim of 150 A-turns will reduce the field from the second segment to similar levels.

The tolerance on the septum field is determined from the extracted beam angular divergence at the septum which is roughly 8 μr. Thus, to limit the extracted beam fluctuations to less than 1/10 of the beam size, the septum field must be stabilized to roughly $\Delta I/I \sim 5 \times 10^{-6}$. This field tolerance should be eased by a factor of 10 by powering a compensation bend in series with the septum as is discussed in Section 4.3.1.

The compensation bending magnet has not been designed. It is required to have an integrated field of 7.7 kGauss-m. In the present design, we have assumed a magnetic length of 80 cm and a field of roughly 10 kGauss. The length and design of the magnets will be adjusted so that the current needed to excite these compensation magnets will be equal to the current needed for the septa so that both can be powered in series.

Correction Coils

The ring will need dipole, quadrupole, and skew quadrupole correctors. The quadrupole correctors are required to optimize the dispersion match through the arcs and wiggler, and the skew quadrupole correctors are needed to correct the betatron coupling and vertical dispersion. The number, placement, and strength of the quadrupole and skew quadrupole correctors has not been determined. With independent power supplies, additional quadrupole trims will not be necessary and, with magnet movers on the quadrupoles or sextupoles, additional skew quadrupole trims will not be needed.

Each standard arc cell will require one vertical dipole corrector and two horizontal dipole correctors which we have nominally placed at the S1 and S3 sextupoles. These correctors should have maximum strengths of 33 G-m which corresponds to a 0.5-mr deflection. Simulations of the trajectory correction show that the typical strength of the correctors is roughly 120 μr. Assuming uncorrelated variation, the corrector power supplies must regulate at the level of $\Delta I/I \lesssim 1 \times 10^{-4}$ to limit the vertical orbit jitter to one tenth of the vertical beam size.

As stated, in the arcs, we plan to use trim windings to the sextupoles to generate dipole fields as is done in the ALS. This choice saves space along the beam line, and makes it relatively easy to generate horizontal magnetic field for the vertical steering despite the vacuum ante-chamber. Also sextupole trims generate a more uniform dipole field than trim winding on the quadrupoles. Additional trim windings on the sextupoles could be used to generate skew quadrupole fields for correction if needed.

In the wiggler straight section, we plan to use trim windings on QFW and QDW quadrupoles for the horizontal and vertical correction. These quadrupoles have larger bore and smaller corrector strengths are required so we are not concerned about the sextupole fields introduced. These correctors should have maximum strengths of 16 G-m, which corresponds to a 0.25-mr deflection. The corrector power supplies must also regulate at the level of $\Delta I/I \lesssim 1 \times 10^{-4}$ to limit the vertical orbit jitter to one-tenth of the vertical beam size. In addition, we could add trims to the ends of each wiggler sections for additional horizontal correction. No estimate of the required strengths has been made.

In the injection and extraction straight section, we need additional orbit correction to optimize the trajectory through the septa. At this point, we have not studied these needs and have only assumed correctors located at the focusing and defocusing quadrupoles that are similar to the wiggler straight section. It will be straightforward to add additional dipole correctors in this region since the vacuum chamber will not include the ante-chamber and will be relatively clear.

4.10.2 Pre-Damping Ring

Bending Magnets

The bending magnets for the pre-damping ring are combined-functions magnets with a 17.5-kGauss effective field and a gradient of 33 kGauss/m; the required good field region is ± 2 cm. The vertical aperture is 4.4 cm which is 6-mm larger than the vacuum chamber height, leaving roughly 2.5 mm for the vacuum chamber wall. Parameters of the magnets are listed in Table 4-29.

At this time, we have not modeled the magnet and may have to modify the parameters subsequently. Because the magnet is a high-field gradient magnet, the fields may be difficult to attain without significant saturation. Although the magnet could be designed to operate with a saturated pole, the saturation will cause the relative field components to vary as a function of the excitation. This may make changing the ring energy difficult. If necessary, we could consider using additional trim magnets to provide the bending or we could construct the pole from Vanadium Permendur which saturates at a higher magnetic field.

Quadrupole and Sextupole Magnets

The quadrupoles and sextupoles in the pre-damping ring have larger apertures than in the main damping ring but similar construction issues apply. The arc magnets will need to accommodate the vacuum ante-chamber and thus will probably be constructed with an asymmetric design. The injection and extraction region quadrupoles have larger

Magnet	Quan.	$B_{0\,\mathrm{eff}}$ (kGauss)	dB_y/dx (kGauss/m)	Mag. Len. (cm)	$\frac{1}{2}$ Gap (cm)
BB (Main Bends)	28	17.5	33	80	1.9
BM (Matching Bends)	4	17.5	0	40	1.9

Table 4-29. *Pre-damping ring bending magnets for Vers. 3.4.*

Arc QFs:	
Number	42
Magnetic Length	26 cm
Inscribed Radius	3.5 cm
Max. B_{pole}	8.8 kGauss
Arc QDs:	
Number	56
Magnetic Length	15 cm
Inscribed Radius	3.5 cm
Max. B_{pole}	5.1 kGauss
Inj/Ext Quads:	
Number	6
Magnetic Length	20 cm
Inscribed Radius	4.0 cm
Max. B_{pole}	4.3 kGauss
Arc Sexts:	
Number	112
Magnetic Length	9 cm
Inscribed Radius	3.0 cm
Max. B_{pole}	2 kGauss

Table 4-30. *Quadrupole and sextupole magnet parameters for NLC pre-damping rings at 2.2 GeV.*

apertures. Although we have not yet started designs of the magnets, they are not expected to be difficult. Table 4-30 lists the magnets and requirements.

Corrector Magnets

The dipole correction fields will be generated in the arcs using trim windings on the sextupoles as is planned for the main damping rings. In the straight sections, the dipole fields would be generated with trim windings on the quadrupoles.

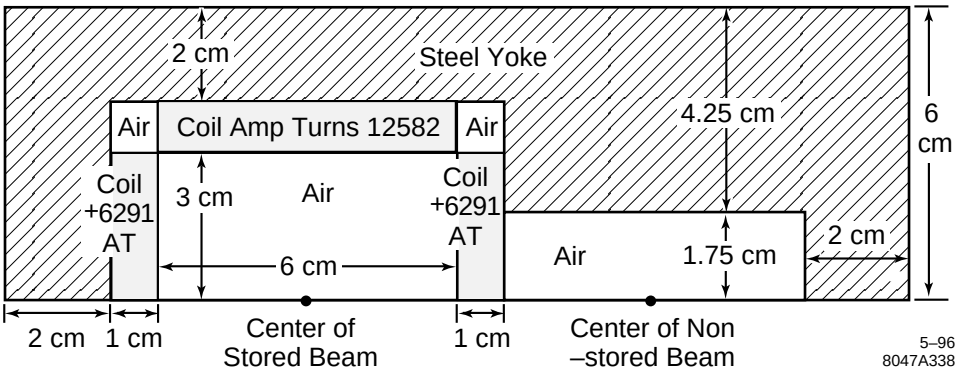


Figure 4-83. *Upper half of a single Panofsky quadrupole for pre-damping ring injection and extraction regions; the present design uses a dual Panofsky quadrupole.*

Injection/Extraction Septa

In the present design, the injection and extraction septa are vertically deflecting Lambertson septa. This is not considered a liability in the pre-damping ring because the vertical emittances are still large. A Lambertson septum can provide strong magnetic fields with greater reliability than a current-sheet septum. The fields in the septa are assumed to be 5 kGauss with a 2.3-cm-thick blade. At this time, no design exists but these parameters are quite relaxed.

Panofsky Quadrupoles

To ease the requirements on the injection and extraction kickers, we plan to use a dual Panofsky quadrupole as the horizontally focusing quadrupole next to the two septa. This provides horizontal focusing of both the stored and injected or extracted beams but does not deflect their trajectories. The dual quadrupole is 50-cm-long and has a gradient of 60 kGauss/m. The apertures are 6.2 cm by 6.2 cm with a 2-cm-thick central blade. The required current density is roughly 20 A/mm^2 which is a factor of six less than that in the SLC damping ring septa; these currents should be straightforward to attain. A preliminary single quadrupole design is illustrated in Figure 4-83.

4.10.3 Injection and Extraction Kickers

Damping ring injection is accomplished by bringing the beam toward a ring straight section at a large enough angle to clear the conventional magnets, bending it nearly parallel to the ring orbit with a DC septum magnet, and bending it onto the ring orbit with a pulsed kicker magnet at the point where the injected and stored orbits cross. Extraction is accomplished by analogous components, in the reverse order. The overall NLC design optimization requires that several trains of bunches must be damped simultaneously. The rise and fall times of the kicker magnetic field must be less than the gap between bunch trains, and the field must be flat for the duration of the train. The NLC parameters allow 60 ns for kicker rise and fall times, and require a 130-ns flattop.

These parameters are similar to those of the SLC damping rings. The SLC damping rings have revolution periods of 120 ns and each ring stores two bunches. The e^+ ring kickers extract and inject one bunch each cycle (the other bunch remains for more damping), while the e^- ring kickers inject and extract both bunches each cycle (one for collisions, one to produce more positrons). The SLC e^+ kickers demonstrate injection and extraction with 60-ns rise and fall

Parameter	Value
Rise time	60 ns
Falltime	60 ns
Flattop	130 ns
Repetition rate	180 Hz
Beam Energy	2.0 GeV
Injected Invariant Emittance	1.0×10^{-4} m-rad
Extracted Invariant Emittance	3.0×10^{-6} m-rad
Insertion Lattice	90° FODO
Insertion Quad Spacing	2.9 m
Beta at Kicker	5.0 m
R_{12} from kicker to septum	6 m
Kick Angle	2.5 mr
Injection Kick Tolerance	$\pm 3.5 \times 10^{-3}$
Extraction Kick Tolerance	$\pm 0.5 \times 10^{-3}$

Table 4-31. *NLC main damping ring kicker requirements.*

times without disturbing other stored bunches. The SLC e^- kickers demonstrate injection and extraction with 60-ns rise or fall times of 60-ns-long bunch trains (albeit containing only two bunches) on a single pulse.

The SLC kickers are implemented with thyratron pulsers, discharging coaxial cables, or oil-filled Blumleins through cables into transmission line magnets with resistive terminations. The magnets use ferrite flux returns and distributed capacitance with loaded epoxy dielectric. The kicker beam pipes are ceramic, with a thin metallic coating to decouple the magnets from the beam at high frequencies. This basic topology, with some parametric changes, is also appropriate for the NLC kickers.

However, several iterations of design over several years were required to produce kicker systems that were operationally reliable and met the SLC performance specifications. Some additional development and demonstration will be required for the NLC. The large aperture (the result of the large emittance) in the NLC e^+ pre-damping ring poses quantitatively larger difficulties than in the SLC case. On the other hand, some of the NLC parameters are more favorable than for SLC, and allow straightforward solutions to the new problems.

Main Damping Rings

The e^- and e^+ main damping rings are identical, and there is little difference in the requirements for injection and extraction, so the same kicker system may be used in all four cases. (The e^+ pre-damping ring kickers are discussed elsewhere, since the requirements are somewhat different.) The requirements are listed in Table 4-31. The tolerances correspond to a centroid jitter of 10% of the beam sigma.

For a transmission line magnet, the field rise time is the sum of the electrical pulse rise time and the magnet transit time. We budget half of the 60-ns rise time to the electrical pulse, leaving 30 ns for the magnet transit time, as for the SLC. The kick angle and beam energy determine the product of field and length. Combined with the horizontal aperture, this determines the total magnetic flux. Combined with the transit time, this determines the voltage, independent of length. The height is fixed by the aperture requirement, then the length essentially determines the required current, or equivalently the magnet and pulser impedance. For the SLC, the available length is short, resulting in an impedance of 12.5 Ω. This requires a large capacitance per unit length, and degrades the thyratron rise time. The kicker length of 1.2 m is chosen for the NLC to produce an impedance of 50 Ω, requiring 16 times less capacitance per unit length

Parameter	Value
Electrical transit time	30 ns
Horizontal aperture	30 mm
Vertical aperture	30 mm
Length	120 cm
Field	139 Gauss
Current	332 A
Voltage	16.7 kV
Inductance	1.51 μHenry
Capacitance	596 pF
Impedance	50 Ω

Table 4-32. *NLC main damping ring kicker magnet parameters.*

than for the SLC. Despite the higher impedance, the magnet voltage is less than half that required in the SLC (this is indirectly the result of the lower emittance). The SLC kicker magnets are 9-cell LC transmission lines with ferrite flux returns, and a similar construction (albeit physically longer) would have adequate bandwidth to meet the NLC requirements. The kicker magnet parameters are listed in Table 4-32.

Kicker Pulsers

Kicker pulsers very similar to those used in the SLC would probably be adequate for the NLC. For three of the four kickers in the SLC (all but e^- extraction), a simple pulser circuit with a charge-line discharged by a thyratron is operationally adequate, although it does not strictly rise from 0.0% to 100.0% or fall from 100.0% to 0.0% in 60 ns. The coherent damping time of the SLC damping rings is fast enough that the midstore kicks in the e^+ ring due to rise and fall times in excess of 60 ns does not degrade the stored beam emittance (although it is operationally preferred to adjust the ring tune so the two midstore kicks tend to cancel). The rise time of the e^- injection kicker and the fall time of the e^- extraction kicker are irrelevant, because there are no other bunches in the ring at those times. Any kick on the turn after e^- injection from excessive fall time damps away rapidly. The case of SLC e^- extraction is more complex and is discussed below.

Since in the NLC, rise or fall times in excess of the 60-ns nominal would only affect bunches that remain in the rings, slight precursors or tails would presumably be tolerable. If this is the case, simple thyratron cable discharge pulsers would be adequate for the NLC main damping rings. With a continuous coaxial charge line, the pulse is intrinsically flat, with only a very slight droop from attenuation in the charge-line cable, plus any residual ringing from the thyratron switching.

Should the flatness of the pulse not meet the specification, it is possible to install an identical kicker magnet in the extraction line outside the ring, at a point related by a $\pm I$ transport matrix to the ring kicker. If the pulser feeds both magnets in parallel, and the cable delays are set appropriately, kicker pulse errors induced at the first magnet will be canceled by the second magnet. It is not unreasonable to expect an order of magnitude relaxation of the flatness tolerance by this technique.

Since the system impedance is higher than for SLC, the thyratron rise time should be somewhat better for the NLC. If the rise and fall times from a simple thyratron pulser are not adequate, several strategies are available. The budget for thyratron rise and fall time can be increased by reducing the magnet transit time. The rise and fall time of the thyratron pulse can be improved, at the expense of more complexity. Or thyratrons could be replaced by other switches, such as spark gaps or solid-state devices.

Cutting the magnet length in half gains 15 ns, but requires twice the current and voltage. Two half-length magnets fed in parallel also gains 15 ns, with no increase in voltage but a doubling of (total) current. Reducing the magnet capacitance per unit length by a factor of four would also gain 15 ns, but would double the impedance and thus the voltage. There is room in the lattice for two such magnets fed in parallel, which could operate at the same voltage and (total) current as the original design but still gain 15 ns of transit time.

In the SLC e^- ring, both bunches are extracted on a single pulse, so a degraded rise time results in one of the two bunches receiving a kick on the turn before extraction. The thyratron pulser is augmented by a saturating ferrite pulse-sharpening filter that removes thyratron prepulse (produced by the intermediate grids in the multistage thyratrons), and part of the slow early rise of the current waveform. This device does not degrade the pulse flatness, but does degrade the fall time. Also, a peaking capacitor in parallel with the charge-line is used to make the current overshoot initially, which makes the field reach the 100% level more quickly, but both degrades the flatness and the fall time (a smaller capacitance value would improve the rise time with minimal flatness degradation, but for the SLC rise time is more important). Since the "bunch train" only has two bunches in the SLC, a single pulse shape adjustment is adequate to make the kick magnitude equal at the two times of interest. It is implemented by oil-filled, motor-adjustable, variable-impedance insertions in the charge-lines.

All of the above pulser modifications degrade the fall time. For SLC e^- extraction, any fall time degradation is irrelevant and is not corrected. For the NLC, one could "use" bunch trains in such an order that the (trailing) train that is disturbed by a slow fall time is the one that will remain in the ring the longest for additional damping. If pulser modifications are used to improve the rise time for the NLC, another thyratron could also be added to the pulser, which would be fired to define the end of the pulse, circumventing the degradation in fall time. This technique is standard for long pulse kickers in proton machines.

Spark gaps, since they operate at high pressure rather than near-vacuum, have much higher dI/dt than thyratrons. With laser-triggering to reduce time jitter to tolerable values, and adequate engineering for erosion reduction and/or rapid replacement, spark gaps could be a thyratron alternative with better rise time and fall time.

Solid-state FETs are now available with rise times better than 10 ns, up to about 1000 V and about 100 A, and have the advantage that they can be turned off as well as on. They can be run in parallel for higher currents, and can be stacked in series for higher voltages. A series-switched FET pulser to produce a 20-kV pulse would require about 40 stages (since the charge voltage is twice the pulse voltage), each stage with several FETs in parallel to handle the current. It is not trivial to trigger a FET stack in such a way that none of the FETs is ever over-voltaged and destroyed, but it is not beyond the state of the art. The number of stages can be reduced by using a transformer, although more parallel FETs are needed in each stage, and the transformer loses some of the rise time improvement.

If the magnet is divided into enough segments fed in parallel (about 20), the voltage required is low enough that a few FETs in parallel could drive each segment directly. The engineering challenges with this approach are that the impedance becomes so low that the FETs must be located very close to the magnet, and the magnet is no longer a transmission line, so it is more difficult to obtain a flat pulse shape.

Pre-Damping Ring

The e^+ pre-damping ring must accept and rapidly reduce the large emittance of the undamped positrons. The pre-damping function is separated from the final damping largely due to the tradeoff between damping rate and final emittance, but the separation has the side benefit of making the kickers less difficult. The requirements for the injection kicker are listed in Table 4-33; the extraction kicker is similar except the required kick angle is 6.7 mr instead of 8 mr. The large emittance results in a large kick angle which require very high kicker power levels, but also relaxes the tolerances compared to the main damping rings.

Parameter	Value
Rise time	60 ns
Fall time	60 ns
Flattop	130 ns
Repetition rate	180 Hz
Beam Energy	2.0 GeV
Injected Invariant Emittance (edge)	6.0×10^{-2} m-rad
Extracted Invariant Emittance (rms)	1.0×10^{-4} m-rad
Insertion Lattice	90° FODO
Insertion Quad Spacing	2.9 m
Beta at Kicker	5.0 m
R_{12} from kicker to septum	6 m
Injection Kick Angle	8.0 mr
Injection Kick Tolerance	$\pm 10 \times 10^{-3}$
Extraction Kick Tolerance	$\pm 1.6 \times 10^{-3}$

Table 4-33. *NLC e^+ pre-damping ring injection kicker requirements.*

Parameter	Value
Electrical transit time	30 ns
Horizontal aperture	62 mm
Vertical aperture	45 mm
Number of feeds	6
Length per feed	28.9 cm
Field	308 Gauss
Current	1104 A
Voltage	18.4 kV
Inductance per feed	0.854 μHenry
Capacitance per feed	3070 pF
Impedance	16.6 Ω

Table 4-34. *NLC e^+ pre-damping ring kicker magnet parameters.*

Achieving a 30-ns magnet transit time in a single magnet would require 80 kV, which is far outside of the SLC experience, so the magnet is divided into four segments driven in parallel at roughly 20 kV. To keep the total length reasonable, the magnet impedance is chosen to be 16.6 Ω. This is 50% higher than the SLC kickers, but requires roughly the same capacitance per unit length. The pre-damping ring kicker magnet parameters are listed in Table 4-34.

The pre-damping ring kicker pulser parameters are essentially identical to those for the main ring kicker pulsers, apart from the lower impedance level. However, six such pulsers are needed for injection, and five for extraction. The lower impedance makes the rise time and fall time requirement somewhat harder to meet, although by using two thyratrons in parallel for each pulser, the impedance per tube can be increased back to the same as for the main ring kickers. (The SLC e^- pulsers demonstrate the feasibility of parallel thyratrons.) The same types of magnet modifications and pulser modifications discussed in connection with the main ring kickers could be used for the pre-damping ring kickers, but the more relaxed tolerances may make them unnecessary.

4.10.4 Transport Lines

The magnets used in the injection, extraction, and other transport lines are straightforward and will not be described.

4.11 Instrumentation for the Main Damping Rings

The emphasis for the damping ring instrumentation is to provide sufficiently high-resolution monitoring to meet the performance requirements for emittance and stability. In particular, the main damping rings require precise control of the beam position in the ring magnets as well as high-resolution beam size measurements to monitor the damping process. The beam-based alignment procedures also require high resolution BPMs to determine small orbit shifts as magnets are varied and to track sources of beam jitter. Finally, the multibunch stability requirements demand accurate bunch-to-bunch measurements of bunch intensity, position, and phase.

4.11.1 Beam Position Monitors

The BPM system is scaled from the PEP-II BPM system using "button" BPM electrodes and BPM processors with a 10-MHz bandwidth that will average over roughly twenty bunches in a train. Each BPM module will contain on-board DSP and memory to allow for multi-turn averaging and other data manipulation. Using the present PEP-II ADC's, the maximum measurement rate would be once every $7\ \mu$s (roughly once every 10 turns).

The button-style BPMs have the advantage of lower beam impedance and simpler construction compared to stripline monitors. We have assumed buttons with an 8-mm diameter in a 1.25-cm-radius beam pipe as described in Section 4.4.1.

Three BPMs are required per arc cell to ensure coverage of all betatron phases and provide some redundancy in monitoring. This makes for 114 arc BPMs, plus eight matching section BPMs, nine wiggler section BPMs and seven straight-section BPMs, giving a total of 138 BPMs per ring. The arc BPMs will be centered in each of the S1 and one of the S3 sextupoles rather than in the quadrupoles, since the orbit tolerances are tightest in the sextupoles.

The tightest resolution requirement is in monitoring the position of the stored orbit since it impacts the vertical equilibrium emittance and is needed for the beam-based alignment of the magnets. Here, 1-μm resolution is required with a 1-μm rms stability over a few hours, and 10-μm stability over 24 hours. Similar requirements are found elsewhere in the NLC, but the stored orbit measurement is made easier by measuring over many turns and applying filtering and signal processing techniques in the BPM electronics to extract the beam signal. The proposed system should have a single turn resolution of

$$\sigma \sim 1\ \mu\mathrm{m}\sqrt{1+\left(\frac{0.5\,\mathrm{A}}{I_{\mathrm{train}}}\right)^2} \tag{4.86}$$

for the average position of twenty adjacent bunches. This is more than sufficient for normal operation and for low current studies the 1-μm resolution could be obtained by averaging over many turns.

Of course, the stringent BPM stability requirement might only be met if remote sensing of magnet positions is also implemented in the ring. BPMs reveal relative changes between the beam and the magnet holding the BPM, but are blind to thermal and seismic drifts in magnet position. The stability issues were discussed in Section 4.8.2.

Specifications for single-turn measurement resolution are considerably looser $\sim 50\,\mu$m as it is used mainly for diagnosing injection errors. The more important criterion for the single-turn measurements is the resolution of individual bunches within each train. A small number (minimum two) of dedicated BPMs will be needed to resolve bunch-to-bunch injection errors. Similarly, bunch-to-bunch errors in the stored beam, which arise from instabilities, can be detected by the same two high-bandwidth BPMs. The bunch spacing in the train suggests a 1-GHz bandwidth is required to discriminate individual bunch motion. The high-bandwidth processor can be used in conjunction with a resonant stripline BPM, turned to have notches in the response at the revolution harmonics thereby enhancing the betatron sidebands.

The limited redundancy in the specified number of arc BPMs is rapidly consumed by the number of special-purpose BPMs: two BPMs are required for single-turn phase space reconstruction; sample-and-hold BPM sum signal is used for peak current monitoring of transient bunch length instabilities; high-bandwidth signal path ($\sim$30 GHz) is required from one BPM for spectral analysis of bunch length instability phenomena; several additional pickups are used for phase monitoring and rf feedback systems.

4.11.2 Beam Size Monitors

Measurement of the beam size in the ring is required to monitor the damping process and to verify that the desired equilibrium emittance is attained. Imaging of visible synchrotron radiation gives an immediate diagnostic of the injected beam size and condition. The gated cameras used in the SLC damping rings can resolve turn-by-turn changes in beam size but will not distinguish individual bunches in the NLC train.

Unfortunately, the fully-damped beam is too small to resolve by visible light. The diffraction-limited resolution is given by [Hofmann 1982]

$$\Delta y = 0.6\gamma\sqrt{2\lambda\lambda_c} \tag{4.87}$$

where $\lambda_c = \frac{4\pi}{3\gamma^3}$, giving at best a 26-$\mu$m resolution, whereas the vertical equilibrium beam size in the arc dipoles is 6 μm.

The visible synchrotron light will be an excellent diagnostic for the beam before it reaches the equilibrium beam size, but measurement of the final spot would require at least a factor 20 shorter wavelength. This would require a monitor operating in the XUV to soft X-ray regime, for which there is no easy design solution.

An alternative to a beam size monitor is to measure the equilibrium beam size with a scanning "laser wire" as proposed at many other points along the NLC. The scanning process probably loses the turn-by-turn beam size information, but this should not be an issue for the stored beam. It may also be possible to design a laser scanner that could scan the beam size quickly compared to the damping time, *i.e.*, $\lesssim 1$ ms.

4.11.3 Bunch Length Measurement

Streak camera measurements of the visible synchrotron radiation yield bunch length and bunch distribution data on a turn-by-turn basis and can further discriminate between individual bunches in the train. Additional confirmation of bunch length can be made with the extracted beam profile at a high-dispersion location downstream of the bunch compressor; this is discussed further in Chapter 5. Finally, transient bunch length behavior in the rings can be observed on the BPM signals, as mentioned above.

4.11.4 Polarization Measurement

Unfortunately, it may be difficult to measure the polarization in the ring because of backgrounds due to synchrotron radiation and particle loss. Regardless, the polarization of the electrons can be measured in the extraction line using a Compton laser polarimeter. Depolarizing resonances are only likely to be encountered in the ring if the energy is changed to achieve shorter damping times for faster repetition rate operation.

4.11.5 RF Diagnostics

The damping ring bunches should be phase-locked to the main linac rf system at the level of 0.1°. Beam loading introduces a phase offset across the bunch so high-bandwidth systems must be used to detect the phase of individual bunches. Stable operation of the rf system under the heavy beam loading will require extensive monitoring of phase and power levels throughout the klystron and feedback systems; this is discussed further in Section 4.5.3.

4.11.6 Ion Measurements

Ion effects can be a serious performance limitation and may work in an unpredictable way. The onset of ion accumulation may be observed with gated beam loss monitors which detect electrons scattered from the ion cloud although this would likely be difficult due to backgrounds. The effect of the ions may be to cause tune shifts in the beam during the store time. Dynamic tune changes such as these can be monitored by gated excitation of the betatron tunes.

4.11.7 Conventional Diagnostics

This would include numerous betatron tune measurement systems based on turn-by-turn BPM analysis, spectrum analyzers, beam excitation systems, etc.. Also included are machine-protection systems that are routinely monitored, for example, during machine repetition rate changes.

4.12 Conclusions and Comments

In this chapter, we have described the damping ring complex for the NLC. The damping rings are designed to produce beams with emittances of $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad at a repetition rate of 180 Hz. With the present electron and positron injector designs, this requires damping the incoming electron emittance by over four orders of magnitude and the incoming positron emittance by roughly six orders of magnitude.

The damping is performed in three rings: a pre-damping ring for the positrons which damps the very large incoming positron emittance to roughly the same size as the incoming electron beam, and two main damping rings that damping the electrons from the injector and the positrons from the pre-damping ring to the desired final values. All three rings use high-field combined-function bending magnets and the main damping rings also have 26 m of wiggler magnets to attain the desired damping rates.

The design of the pre-damping ring is relatively straightforward—the primary difficulties are related to the very large incoming beam emittance, *i.e.,* the large transverse and energy apertures that are required and the injection and extraction systems. We presented a design with solutions for both of these issues. Although we have not yet started detailed engineering studies for the components of the ring, the tolerances and requirements appear reasonable. The alignment tolerances are well within the capabilities of conventional alignment practices (although we will include the ability to use beam-based alignment) and most magnet requirements are not dissimilar from those of the PEP-II LER or the ALS synchrotron light storage ring.

The main damping rings are more difficult and thus most of our studies have concentrated on them. Fortunately, the rings have a number of similarities with the present generation of synchrotron light storage rings and with the new flavor factories being constructed. In addition, many of the design and physics questions will be studied in the ATF prototype damping ring that is being constructed at KEK. Thus, many of the accelerator physics questions that may arise in the NLC damping rings will be studied and verified in these other rings.

These similarities have also simplified the design process and provide confidence that most of the components could be constructed. In particular, the beam currents are less than those in the PEP-II LER and we have adopted rf cavities and a multibunch feedback system based on those being constructed for the LER. Likewise, the vacuum system is based on those of the ALS and PLS synchrotron radiation sources and the magnets designs are similar to those in these and other radiation sources. Although many of these components are the present state-of-the-art, we will benefit from operational experience of these other rings before the construction of the NLC is started. This will allow appropriate design modifications to ensure the required reliability for the NLC damping rings.

One exception are the combined-function bending magnets which have higher fields than in these other rings. We have performed some initial studies which show that the fields are attainable, but we still will need to build and measure a prototype before we can have complete confidence in the design. Fortunately, this is relatively inexpensive. If indeed the specifications cannot be met, it is straightforward to redesign the rings with looser requirements on the bending magnets although this is at the expense of some of the operating range.

Finally, the damping rings must provide a stable reliable platform for the rest of the collider. One of the lessons from the SLC is that we cannot expect optimal performance if all the components are operated at their limits. Thus, throughout the damping ring designs, we have provided substantial margin in the design performance over that actually required for the NLC. In particular, we have designed for higher currents than needed as well as apertures in excess that required. Similarly, we have provided for beam-based alignment and tuning systems that will provide much better accuracy than is actually needed. All of these overheads in the design should yield improved operations and reliability.

References

[Adolphsen 1995] C. Adolphsen, private communication (1995).

[Baltay] M. Baltay, private communication (1995).

[Bane 1988] K. Bane, "The Caculated Longitudinal Impedance of the SLC Damping Rings", *Proc. of the 1st European Part. Acc. Conf.*, Rome, Italy, 637 (1988).

[Bane 1989] K. Bane and R. Ruth, "Bunch Lengthening Calculations for the SLC Damping Rings", *Proc. of the 1989 Part. Acc. Conf.*, Chicago, IL, 789 (1989).

[Bane 1993] K. Bane and K. Oide, "Simulations of the Longitudinal Instability in the SLC Damping Rings", *Proc. of the 1993 Part. Acc. Conf.*, Washington DC, 3339 (1993).

[Bane 1995a] K. Bane and K. Oide, "Simulations of the Longitudinal Instability in the New SLC Damping Rings", *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Bane 1995b] K. Bane *et al.*, "High-Intensity Single Bunch Instability Behavior in the New SLC Damping Ring Vacuum Chamber", *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Bane 1995c] K. Bane and K. Kubo, "Bunch-to-Bunch Variation in the Synchronous Phase in the NLC Damping Rings", NLC–Note 19 (1996).

[Barnet 1994] I. Barnet *et al.*, "Dynamic Beam Based Alignment", CERN–SL94–84 (BI), (1994).

[Barry 1993] W. Barry *et al.*, "Design of the ALS Transverse Coupled Bunch Feedback System", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC (1993).

[Barry 1995] W. Barry *et al.*, "Design of the PEP-II Transverse Coupled Bunch Feedback System", *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Bartalucci 1993] S. Bartalucci *et al.*, "Analysis of methods for controlling multibunch instabilities in DAFNE", LNF–93/067 (P), (1993).

[Berg 1995] J. Scott Berg, "Transverse Coupled Bunch Mode Coupling and Growth Rates for the NLC Main Damping Ring", NLC–NOTE–10 (1995).

[Boussard 1975] D. Boussard, CERN LABII/RF/INT/75–2 (1975).

[Boussard 1983] D. Boussard and G. Lambert, "Reduction of the Apparent Impedance of Wide-Band Accelerating Cavities by RF Feedback", *IEEE Trans. Nucl. Sci.* **30**, 2239 (1983).

[Brinkmann 1990] R. Brinkmann, "A Study of Low Emittance Damping Ring Lattices", DESY–M–90–09 (1990).

[Chao 1995] A. Chao, B. Chen, and K. Oide, "A Weak Microwave Instability with Potential Well Distortion and Radial Mode Coupling", presented at the 1995 Part. Acc. Conf., Dallas, TX (1995).

[Corlett 1993] J.N. Corlett and R.A. Rimmer, "Impedance Measurements of Components for the ALS", *Proc. of the 1993 Part. Accel. Conf.*, Washington, DC (1993).

[Fox 1993] J. Fox *et al.*, "Feedback Control of Coupled Bunch Instabilities", *Proc. of the 1993 Part. Accel. Conf.*, Washington, DC (1993).

[Furman 1995] M. Furman, "The Positron Beam Instability Code POSINST", to be documented (1995).

[Grobner] O. Gröbner, "Bunch Induced Multipactoring", *Proc. Xth Int. Conf. on High Energy Acc.*, Serpukhov, USSR, 277 (1977).

[Greene 1992] E. Greene, T. Warwick, "Environmental Vibration Measurements of the ALS Floor and Storage Ring", LSBL–144 (1992).

[Halbach 1969] K. Halbach, *Nucl. Instr. and Methods* **74**, 147 (1969).

[Halbach 1985] K. Halbach, *J. Appl. Phys.* **57**, 3605 (1985).

[Haïssinksi 1973] J. Haïssinski, "Exact Longitudinal Equilibium Distribution of Stored Electrons in the Presence of Self-Fields", *Il Nuovo Cimento* **18B**, No. 1, 72 (1973).

[Heifets 1994a] S. Heifets, "Bunch-to-Bunch Phase Variation in the NLC DR", NLC–NOTE–6 (1994).

[Heifets 1994b] S. Heifets, "Study of the NLC DR Impedance and Collective Effects: the First Attempt", NLC–NOTE–9 (1994).

[Heifets 1995] S. Heifets, "Study of an Instability of the PEP-II Positron Beam", PEP-II AP Note 95.21 (June 7, 1995).

[Heim 1995] J. Heim *et al.*, "Wiggler Insertion of the PEP-II B-Factory," *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Herb 1995] S. Herb, G.B. Jaczko, and F. Willeke, "Beam-Based Calibration of Beam Position Monitor Offsets in the HERA Proton Ring using Strong Sextupole Fields", (1995).

[Hirata 1992] K. Hirata and K. Yokoya, "Nongaussian Distribution of Electron Beams due to Incoherent Stochastic Processes", *Particle Acclerators* **39** 147–158; and KEK–Preprint–91–212 (1992).

[Hofmann 1982] A. Hofmann and F. Meot, "Optical Resolution of Beam Cross-section Measurements by Means of Synchrotron Radiation", *Nucl. Instr. and Methods* **203**, 483–493 (1982).

[Izawa 1995] M. Izawa, Y. Sato and T. Toyomasu, "The Vertical Instability in a Positron Bunched Beam", *Phys. Rev. Lett.* **74**, 5044 (1995).

[JLC] KEK JLC damping ring design.

[Juravlev 1995] V.M. Juravlev *et al.*, "Seismic Conditions in Finland and Stability Requirement for the Future Linear Collider" (1995).

[Kamada 1994] S. Kamada, "Overview on Experimental Techniques and Observations", *Proc. of the Workshop on Nonlinear Dynamics in Particle Accelerators: Theory and Experiments*, Arcidosso, Italy (1994).

[ATF 1995] "ATF Accelerator Test Facility Design and Study Report", KEK Internal 95–4 (1995).

[Kikuchi 1995] M. Kikuchi *et al.*, "Beam-Based Alignment of Sextupoles with the Modulation Method", *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Klatt 1986] R. Klatt and T. Weiland, "Wake Field Calculation with Three-Dimensional SCI Code", *1986 Linear Acc. Conf. Proc.*, SLAC, 282 (1986).

[Krejcik 1993] P. Krejcik *et al., Proc. of the 1993 Part. Acc. Conf.*, Washington DC (1993).

[Kubo 1993] K. Kubo, *et al*, "Compensation of Bunch Position Shift Using Sub-RF Cavity in Damping Ring", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC, 3503 (1993).

[Lambertson 1995] See, for example, G.R. Lambertson, "Update of RF Reaction During Beam Gap", PEP–II Tech Note 85 and CBP Tech Note 73, (1995).

[Mattison 1995] T. Mattison, "Injection/Extraction Kickers for the NLC Damping Rings", talk given to NLC damping ring group 1/10/95.

[Minty 1995] M.G. Minty, W.L. Spence, "Injection Envelope Matching in Storage Rings," *Proc. of 1995 Part. Acc. Conf.*, Dallas, TX (1995).

[Minty 1995b] M.G. Minty and R.H. Siemann, "Heavy Beam Loading in Storage Ring RF Systems", SLAC–PUB–95–6801, submitted to *Nucl. Instr. and Methods* , (September 1995).

[Moshammer 1993] W. Moshammer and T. O. Raubenheimer, presentation at LC93.

[Murphy 1987] J. B. Murphy, "Storage Ring Lattice Considerations for Short Wavelength Single Pass FELs", *Proc. of the ICFA Workshop on Low Emittance e^+e^- Beams*, BNL–52090, Brookhaven, NY, 197 (1987).

[Ohmi 1995] K. Ohmi, "Beam and Photo-Electron Interactions in Positron Storage Rings", KEK Preprint 94–198, (February 1995).

[Oide 1990] K. Oide and K. Yokoya, "Longitudinal Single Bunch Instability in Electron Storage Rings", KEK Preprint 90–10, (1990).

[Oide 1994] K. Oide, "A Mechanism of Longitudinal Single Bunch Instability in Storage Rings", KEK Preprint 94–138, (1994).

[Pedersen 1975] F. Pedersen, "RF Cavity Feedback", *IEEE Trans. Nucl. Sci.* NS–22, no. 3 (1975), and NS-32, no. 3 (1985).

[Pedersen 1992] F. Pedersen, SLAC-400, 192 (1992).

[PEP-II 1993] LBL–PUB 5379, SLAC–418, in PEP–II Conceptual Design Report, June 1993.

[Piwinski 1972] A. Piwinski, DESY Report 72/72 (1972).

[Piwinski 1985] A. Piwinski, "Beam losses and lifetime", CERN Accelerator School, Gif-sur-Yvette, CERN**85-19** (1985).

[Raubenheimer 1988] T. O. Raubenheimer, L. Z. Rivkin, R. D. Ruth, "Damping Ring Designs for a TeV Linear Collider", *Proc. of the 1988 DPF Summer Study, Snowmass '88*, Snowmass, CO, 620 (1988).

[Raubenheimer 1991] T.O. Raubenheimer, "Generation and Acceleration of Low Emittance Flat Beams for Future Linear Colliders", Ph.D. Thesis, Stanford University, SLAC–387 (1991).

[Raubenheimer 1992] T.O. Raubenheimer and P. Chen, SLAC–PUB **5893**, in "Ions in the Linacs of Future Linear Colliders", Proc. of LINAC92, Ottawa, vol. 2, 630–632 (1992).

[Raubenheimer 1992] T. Raubenheimer, "Emittance Growth due to Beam Gas Scattering", KEK Report **92–7** (1992).

[Raubenheimer 1993a] T.O. Raubenheimer, R. Early, T. Limberg, and H. Moshammer, "A Possible Redesign of the SLAC SLC Damping Rings", *Proc. of the 1993 Part. Acc. Conf.*, Washington DC, (1993).

[Raubenheimer 1993b] T.O. Raubenheimer, K. Egawa, M. Kikuchi, K. Kubo, S. Kuroda, K. Oide, S. Sakanaka, N. Terunuma, J. Urakawa, "The Vertical Emittance in the ATF Damping Ring", *Nucl. Instr. and Methods* **A335:**1 (1993).

[Raubenheimer 1994] T.O. Raubenheimer "The Core Emittance with Intrabeam Scattering in e^+/e^- Rings", *Particle Acclerators* **45:**111 (1994).

[Raubenheimer 1995a] T.O. Raubenheimer and F. Zimmermann, "A Fast Beam-Ion Instability: Linear Theory and Simulations", *Phys. Rev. E* **52**, 5487 (1995).

[Raubenheimer 1995b] T.O. Raubenheimer, private communication (1995).

[Rieke 1972] F.F. Rieke and W. Prepejchal, "Ionization cross sections of gaseous atoms and molecules for high energy elelctrons and positrons", *Phys. Rev. A* **6**, (1972).

[Rimmer 1992] R.A. Rimmer, "RF Cavity Development for the PEP-II B-Factory", Proc. Intl. Workshop on B-Factories, KEK, Tsukuba, Japan (1992).

[Rivkin 1985] L. Z. Rivkin, "Damping Ring for the SLAC Linear Collider", Ph.D. Thesis, Cal. Tech. (1985).

[Robin 1995] D. Robin, G. Portmann, and L. Schachinger, "Automated Beam Based Alignment of the ALS Quadrupoles", NLC Note 18, (1995).

[Robinson 1962] K.W. Robinson, CEA report, CEAL–1010 (1962).

[Rojsel 1994] P. Rojsel, "A Beam Position Measurement System Using Quadrupole Magnets Magnetic Centre as the Position Reference", *Nucl. Instr. and Methods* in *Phys. Rev.* 343, (1994).

[Rogers 1995] J. Rogers, "Photoelectron Trapping Mechanism for Horizontal Coupled Bunch Mode Growth in CESR", CBN 95–2 (1995).

[Rossbach 1988] J. Rossbach, "Closed-Orbit Distortions of Periodic FODO Lattices Due to Plane Ground Waves", *Particle Acclerators* **23**, 121 (1988).

[Sakanaka 1995] S. Sakanaka *et al.*, "Design of a high power test cavity for the ATF damping ring", *Proc. of the Part. Acc. Conf.*, Dallas, TX (1995).

[Sands 1976] M. Sands, "Beam-Cavity Interactions", Orsay report, 2–76 (1976).

[Siemann 1983] R. Siemann, "Computer Models of Instabilities in Electron Storage Rings", in *The Physics of Particle Accelerators*, AIP Conf. Proc. **127**, 431 (1983).

[Stupakov 1995] G. Stupakov, T.O. Raubenheimer and F. Zimmermann, "A Fast Beam-Ion Instability: Effect of Ion Decoherence", *Phys. Rev. E* **52**, 5499 (1996).

[Thompson 1989] K.A. Thompson, "Transverse and Longitudinal Coupled Bunch Instabilities in Trains of Closely Spaced Bunches", *Proc. of the 1989 Part. Acc. Conf.*, Chicago, IL (1989).

[Thompson 1991] K.A. Thompson, "Simulation of Longitudinal Coupled-bunch Instabilities", ABC-24, (1991).

[Teng 1985] L.C. Teng, "Minimum Emittance Lattice for Synchrotron Radiation Storage Rings", LS–17 Argonne (1985).

[Turner 1995] J.L. Turner *et al.*, "Vibration Studies of the Stanford Linear Collider", *Proc. of the 1995 Part. Acc. Conf.*, Dallas, TX, (1995).

[Wilson 1994] P. Wilson, private communication (1994).

[Zisman 1986] M. Zisman, S. Chattopadhyay, J. Bisognano, "ZAP Users Manual", LBL–21270 (1986).

Contributors

- Karl Bane
- Lou Bertolini
- John Byrd
- Yunhai Cai
- John Corlett
- Dick Early
- Miguel Furman
- Sam Heifets
- Kwok Ko
- Patrick Krejcik
- Tom Mattison
- Bobby McKee
- Michiko Minty
- Cho Ng
- Tor Raubenheimer
- Robert Rimmer
- David Robin
- Ben Scott
- Steve Smith
- Bill Spence
- Cherril Spencer
- Gennady Stupakov
- Perry Wilson
- Sasha Zholents
- Frank Zimmermann

5

Bunch Compressors and Prelinac

Contents

5.1 Introduction

This chapter describes the NLC bunch compressors, which compress the 4-mm-long bunches from the damping rings to the $100 \sim$150-μm bunch lengths required in the main linacs and final foci. It also describes the S-band prelinacs that accelerate the beams from the damping rings at roughly 2 GeV to an energy of 10 GeV and the spin rotators that provide control over the particle spin orientation. These systems are part of the low-energy injector complex that prepares the beams for injection into the X-band main linacs. Thus, it is essential that they provide a stable and reliable platform for the tuning of the linacs and final foci. In the next paragraphs, we briefly summarize the status of each of these subsystems before delving into greater detail in the body of the chapter.

The first items in the compressor systems are the spin rotators which are located at the exit of the damping rings. We plan to install a spin rotator on the electron side and leave space for an installation on the positron side. This is to allow for a future upgrade to polarized positrons or polarized electrons which would be wanted for γ-γ or e^--e^- collisions. The rotators are based on pairs of solenoids separated by a horizontal arc to allow full control over the spin orientation; this is similar in concept to the original SLC spin rotator system. Solenoids were chosen to control the spin because the other alternate, namely a snake, requires a vertical bending chicane which must be unreasonably long to prevent synchrotron radiation from increasing the vertical emittance. Of course, the problem with solenoids is that they couple the beam which, with flat beams, leads to an increase in the projected vertical emittance; this is the reason the original solenoid system is not presently used in the SLC. Our present design uses pairs of solenoids which are optically separated so that the coupling is fully canceled. There are very tight tolerances on the quadrupole fields between the sextupoles but an extensive skew correction system has been included to ease the tolerances as well as correct any residual coupling from the damping rings.

Next, the NLC bunch compressors must compress the bunches from the damping rings with 4-mm lengths to the $100 \sim$150-μm bunch lengths required in the main linacs and final foci. To perform the compressions, a two-stage compressor system has been designed: the first stage follows the damping rings and the second stage operates at a beam energy of 10 GeV at the exit of the 8-GeV S-band prelinacs.

The two-stage system has a number of advantages over a single stage compressor. In particular, it keeps the rms energy spread less than roughly 2%, it is optically more straightforward, and the bunch lengths are more logically matched to the acceleration rf frequency so that the energy spread due to the longitudinal wakefields can be canceled locally. The disadvantage of the two-stage design is that it is more complex and lengthy than a single-stage compressor. The first stage rotates the longitudinal phase by $\pi/2$ while the second stage performs a 2π rotation. In this manner, phase errors due to the beam loading in the damping rings and energy errors due to imperfect multibunch energy compensation in the 8-GeV S-band prelinacs do not affect the beam phase at injection into the main linac.

One of the driving philosophies behind the NLC compressor design has been to utilize naturally achromatic magnetic lattices wherever the beam energy spread is large. In particular, the optics is chosen so that quadrupoles are not placed in regions of large dispersion and strong sextupoles are not needed. This choice arises from experience with the SLC bunch compressors which are based on second-order achromats where quadrupoles are located in dispersive regions and strong sextupoles are used to cancel the chromatic aberrations. Unfortunately, the SLC design is extremely difficult to operate and tune because of the large nonlinearities and the sensitivity to multipole errors in the quadrupoles; over the years additional nonlinear elements have been added (skew sextupoles and octupoles) to help cancel the residual aberrations but the tuning is still difficult. To further facilitate the tuning in the NLC design, we have explicitly designed orthogonal tuning controls and diagnostics into the system. We feel confident that these considerations will make the system relatively straightforward to operate and tune.

Finally, the rf acceleration throughout the compressor system, *i.e.*, the bunch compressor rf and the 8-GeV prelinac, is performed with relatively low frequency rf. In particular, the first bunch compressors use L-band (1.4 GHz) rf and the prelinac and second bunch compressor use S-band (2.8 GHz) acceleration. Although these systems are, in general,

longer and more expensive than higher frequency accelerator systems, we feel that they are needed at the low beam energy. In particular, we have designed the systems so that they have small beam loading and relatively loose alignment tolerances. This is important to provide the reliability and stability that is desired in the bunch compressor system. Of course, such lightly-loaded systems are inefficient and could not be used to accelerate the beams to very high energy.

To minimize the long-range transverse wakefields, all of the accelerator structures are Damped-Detuned Structures (DDS), similar to those developed for the X-band main linacs. In addition, the beam loading is primarily compensated using the ΔT (early injection) method. This has the advantage of canceling the bunch-to-bunch energy deviations locally within an accelerator structure thereby minimizing the transverse emittance growth due to dispersive and chromatic effects. Details of the multibunch beam loading compensation, the accelerator structures, and power sources is presented in Chapter 6.

The design described in this chapter appears to meet, or even to surpass, all requirements; the requirements are detailed in Section 5.2. At the same time, the present design still allows for improvements and optimization. First, because the designs were being made concurrently, we overestimated the errors in the beams from the damping rings that we should expect—assuming a more reasonable input beam could significantly change the optimization and potentially decrease the length and complexity of the compressor system. Second, an option of simplifying the second stage to only perform a $\pi/2$ rotation in longitudinal phase space should be explored. This also could provide a substantial saving in length and S-band rf components, but it will require tighter tolerances on the phase variation in the damping rings and the energy compensation in the 8-GeV S-band prelinac. Third, another path, that should be further investigated, is the possibility of optically compensating the longitudinal nonlinearities. The present design uses additional rf systems in the second compressor to compensate the longitudinal nonlinearities introduced in the dispersive regions. It would certainly be desirable to compensate these effects locally without having to resort to the additional rf systems.

In the next section (Section 5.2), we review the requirements of the system and the assumed beam from the damping ring. We then provide an overview of the design choices that were made and some of the alternate options. Next, in Section 5.3, we describe the various components of the compressor in more detail and in Sections 5.4 and 5.5, we discuss the longitudinal and transverse dynamics which are relevant to the design, and the tuning schemes that are needed to attain the desired performance. The tolerances and tuning schemes are described in Section 5.6, and the chapter concludes with a summary and an outlook on further studies.

Finally, it should be noted that, at the time of this study, not all of the damping ring or the low frequency rf system parameters were finalized. Thus, in many of the dynamics studies, performed for the bunch compressors, we intentionally over-estimated the expected errors to provide a margin. The example of the synchronous phase offsets has already been mentioned. Similarly, we assumed SLAC S-band wakefields when calculating effects of the accelerator structures. Because the present design calls for structures that have roughly twice the group velocity, the SLAC transverse single bunch wakefield is roughly a factor of two too strong and the longitudinal wakefield is roughly a factor of $\sqrt{2}$ too strong. Thus, the system will perform even better than has been calculated. We will note sections that have these errors.

5.2 Requirements and Design Options

In this section, we will describe the design goals of the bunch compressor systems and beam requirements. We will also outline the design choices that were made and some of the alternate options.

5.2.1 Design Goals

The NLC bunch compressor system is designed with six goals:

- The compressor system must compress a bunch with $\sigma_z \approx 4$ mm from the damping rings to the $\sigma_z = 100\ \mu$m $\sim$ $150\ \mu$m needed in the main X-band linacs and the final foci. This requires compressing the bunch length, and correspondingly increasing the energy spread, by a factor of 30–40.
- The bunch compressor system should be able to compensate for bunch-to-bunch phase offsets induced by the beam loading in the damping rings; if not corrected, these phase errors would become energy errors at the end of the linacs. This requires that the compressor system rotate the longitudinal phase space by $\pi/2$, translating the phase errors into injection energy errors.
- The bunch compressor system should provide a trombone-like arm, reversing the direction of the beam before injection into the main X-band linacs. This allows extension of the linac lengths to upgrade the beam energy at the IP without significant modification to the injection systems. It also allows for abort systems to dump the beam and feed-forward systems to correct the beam trajectory, phase, and energy before injection into the main linacs.
- The compressor system must include a spin rotator system that can rotate the polarization vector of the electron beam from the vertical plane to any arbitrary orientation. This is easiest at lower beam energy.
- The compressor system must be able to correct for coupling mismatches from the damping rings; dispersive mismatches from the damping rings will be corrected in the ring extraction line, described in Chapter 4.
- The compressor systems are the injectors into the main linac and they need to provide a stable and reliable platform to allow tuning of the linac and final focus systems. The system should have reasonable tolerances and limit the horizontal and vertical emittance dilutions to 20% and 30% respectively.

5.2.2 Beam Requirements

In this section, we describe the injected and extracted beam requirements; these are summarized in Table 5-1. The damping rings are designed to produce beams with a relative energy spread of $\sigma_\delta \sim 1 \times 10^{-3}$ and a bunch length of $\sigma_z \approx 4$ mm. The beam emittances are damped to $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad and the spin is oriented vertically to preserve the polarization in the rings.

The compressor system has been designed to accept beams whose incoming phase varies due to the transient beam loading in the damping rings. The present design allows for phase errors of $\pm$20 ps while limiting the phase error of the extracted beams to less than $\pm 0.2^\circ$ of X-band. This will limit the resulting energy error at the end of the main linac to less than 0.1% and will limit the luminosity loss due to shifts of the collision point to less than 1%. It should be noted that, in the present damping ring design, the phase variation is less than ± 12 ps (Chapter 4).

The compressor system will operate with a train of 90 bunches having bunch charges of 0.68×10^{10} particles for NLC-Ia and 1.31×10^{10} particles for NLC-IIc; these values are 5% higher than the design bunch charges at the IP to provide overhead in the bunch charge. The higher charge for NLC-IIc coincides with longer bunches in the main linacs ($\sigma_z = 150\ \mu$m) while the lower charge in NLC-Ia corresponds to a shorter bunch length (100 μm).

In all cases, the expected beam loss through the compressor system is much less than 1%. To provide an additional margin, the compressor system has been designed to accept beams with an initial bunch length of 5 mm. Because

Inj. Energy	1.8–2.2 GeV
Inj. Bunch charge/Beam current	$< 1.31 \times 10^{10} / < 1.5$ A
Inj. Phase error Δz	< 6 mm (4 mm expected)
Inj. Bunch length σ_z	< 5 mm (4 mm expected)
Inj. Energy error $\Delta E/E$	$< 0.01\%$
Inj. Energy spread σ_δ	0.1%
Inj. Emittances $\gamma\epsilon_x$, $\gamma\epsilon_y$	$3 \times 10^{-6}, 3 \times 10^{-8}$
Ext. Energy	10 GeV
Charge loss	$< 1\%$
Ext. Phase error Δz	< 15 μm
Ext. Bunch length σ_z	$100 - 150$ μm
Ext. Energy error $\Delta E/E$	$< 3\%$
Ext. Energy spread σ_δ	$< 1.5\%$
Ext. Emittances $\gamma\epsilon_x$, $\gamma\epsilon_y$	$3.6 \times 10^{-6}, 4 \times 10^{-8}$

Table 5-1. *Input and extracted beam requirements for compressor system.*

there are no quadrupoles in regions of large dispersion, this has minimal impact on the design requiring only slightly larger horizontal apertures in the bending magnets and vacuum system. All apertures have been specified in excess of $\pm 5\sigma_{x,y}$ plus an allowance for the injected phase error of ± 6 mm and an additional allowance for steering errors of 2 mm; the allowance of 2 mm for steering errors is based on present SLC operation.

Finally, the bunch compressor system has been designed to have minimal emittance dilutions. In particular, the system needs to limit the horizontal emittance dilution to 20% and the vertical dilution to 30% while providing a stable platform for the tuning of the downstream systems. Wherever possible we have chosen technologies that, while more expensive or less efficient, will ease the tolerances. We believe the added cost and inefficiency is acceptable for these relatively small sections.

5.2.3 Spin Rotator Options

The spin rotator system must orient the spin so that the polarization is longitudinal at the IP. Because there is net bending in the collider to provide for the multiple IPs with a 20-mr crossing angle and because the beam energy at the IPs will be varied, it is necessary to provide full control over the orientation of the spin vector. To control the orientation of the spin vector, one can either use solenoids or a snake or serpent design that consists of interleaved horizontal and vertical bending magnets.

The choice of a rotator system is determined by emittance dilution constraints and limitations on the momentum compaction (R_{56}). In the NLC, it would be difficult to preserve the vertical emittance of $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad through a "half-serpent" system [Fieguth 1987] which necessarily includes vertical bending. To limit the vertical emittance dilution due to synchrotron radiation to 2%, one can show that the length of the vertical dipole magnets is related to the beam energy as [Emma 1994].

$$L[\text{m}] > 190 E[\text{GeV}] \quad . \tag{5.1}$$

Even at the minimum beam energy of 2 GeV, the bending magnets must become excessively long. For this reason, a solenoid-based rotator system was chosen. It is located immediately after damping ring extraction where the beam

energy is low and the beam energy spread is small. The betatron coupling introduced by the solenoids is corrected by using pairs of solenoids separated by an optical transformation so that the coupling is canceled.

5.2.4 Bunch Compressor Options

For the NLC design, we are currently planning to use a two-stage compressor design which is outlined in Ref. [Raubenheimer 1993]. The first stage, located after the spin rotator, compresses the bunch from the damping rings to $\sigma_z = 500\ \mu$m for injection into an S-band prelinac. There, the beam is accelerated from the damping ring energy of roughly 2 GeV to an energy of 10 GeV where it is injected into the second stage of the compressor. This second stage compresses the bunch to the final bunch length of $\sigma_z = 100\ \mu$m–$150\ \mu$m, before injection into the X-band main linac. The first compressor stage rotates the phase space by roughly $\pi/2$ and the second stage rotates the longitudinal phase space by 2π so that the beam phase in the X-band main linac is independent of the phase in the damping rings.

While the linear design of such a two-stage bunch compressor system is conceptually straightforward, the design is complicated by the large longitudinal nonlinearities introduced in both the bending sections and the rf sections. The solution we have adopted is to use the nonlinearity of an additional sinusoidal rf to cancel the nonlinearity introduced in the bending sections, but this is difficult and is potentially dangerous; systems where one large effect cancels another large effect are inherently extremely sensitive. We have investigated some options of decreasing the optical longitudinal nonlinearities by adding sextupoles but these made the transverse phase space unacceptable. These issues will be discussed further in Section 5.4.

Alternate Approaches

Unfortunately, the second-stage bunch compressor is complicated because of the required 2π rotation in longitudinal phase space. One possible alternate approach would be to use a $\pi/2$ compressor for the second stage *if the S-band linac can be used to compensate the bunch-to-bunch energy deviations*. This would have two advantages: first, it simplifies the design of the second stage compressor; the required R_{56} (the R_{56} is the (5,6) element in the linear transfer matrix) in the second compressor would be 0.05 m which would be easily achieved with a simple FODO arc. Second, such a system might allow one to tune the bunch-to-bunch energy deviation of the beam after the X-band linac by making small phase changes in the second compressor. The allowable phase changes must be small compared to the final beta function of $100\ \mu$m, *i.e.*, $\Delta\phi \ll 1^\circ$ X-band.

It is also possible to consider designing a single-stage bunch compressor system at the exit of the damping rings, but the large energy spreads and large space charge forces at the exit of the compressor are thought to make preservation of the transverse emittances difficult; a systematic study of the required tolerances has not been performed. In addition, the nonlinearity in the longitudinal phase space would seem to complicate if not preclude this option (see the discussion in the next section).

Alternatively, one could consider accelerating the uncompressed beam from the damping rings in a low-frequency structure and then use a single-stage compressor. This option would reduce the relative energy spread and the space charge forces, but, because of the nonlinearity of both the rf and the T_{566}, it is more difficult to perform the 90° rotation. This system does not appear to have significant advantages over the present design and thus we will not consider it further.

Single-Stage *vs.* **Multistage**

At first glance, a single-stage compressor has several advantages compared with the two-stage solution. It is simpler, contains less elements, and, perhaps most importantly, avoids the significant wakefield effects in the prelinac, thereby improving the longitudinal dynamics at high intensity. It has also been argued that a single-stage compressor may ease the compensation of multibunch effects [Oide 1994].

The main disadvantages of a single-stage compressor are the increase of the final energy spread to 5%, which enhances the harmful effects of the T_{566} and U_{5666} and the much larger sensitivity to incoming phase variations; the T_{566} and U_{5666} are the TRANSPORT notation for the higher-order dependences of the path length on the energy deviation: z-δ^2 and z-δ^3. The nonlinearity of the rf alone is already significant. Assuming an L-band rf system, an initial phase error $z_0 \approx \sigma_{z0} \approx 5$ mm would change the final phase at the exit of the compressor by $\Delta z_f \approx k^2 z_0^3/6 \approx 20\,\mu$m, where, as before, $k \equiv 2\pi/\lambda$. Such a change is too large to be tolerable, considering that the beam in the main linac will be off crest by about 10–15° and that the energy change at the end of the linac has to be small.

As a concrete example, we have performed simulations of the longitudinal dynamics for a single compressor consisting of 750-MV L-band rf and a wiggler (or chicane) with $R_{56} \approx 9$ cm. In the ideal case of perfectly compensated U_{5666} and T_{566}, an initial phase error of ± 6 mm increases the final bunch length to 155 μm (from 100 μm for no initial phase offset). If the T_{566} is included, without compensation, the final rms bunch length for zero initial phase error is as large as 58 mm. Applying a compensating S-band rf, as discussed for the double-stage compressor, the minimum final bunch length as a function of compensating voltage V_c was found to be 400 μm, for $V_c \approx 195$ MV and no initial phase error. Such a single-stage compressor does not fulfill our specifications.

To make a single-stage compressor a viable and attractive option, the T_{566}-matrix-element has to be smaller than about a third of the R_{56}-matrix-element (for a generic wiggler or chicane, $|T_{566}|$ is 1.5 times R_{56}). If the T_{566}-matrix-element is of this size, the residual nonlinearity can be compensated by a decelerating rf as in the double-stage scheme, without introducing other intolerable aberrations. A small T_{566}-matrix-element may be realized by manipulating the second-order dispersion via quadrupoles at dispersive locations and/or via additional sextupoles. A further complication arises, however, in the transverse phase space, where dilutions of the transverse emittance have to remain tolerable for an rms energy spread as large as 5%. These two requirements, small T_{566} and tolerable transverse emittance dilution, appear to be in conflict with each other, as changes to the second-order dispersion also generate higher-order dispersion and chromaticity, which affect the transverse emittances.

For these reasons, attempts to design a single-stage compressor with satisfactory performance in all phase-space dimensions have not yet been successful.

5.2.5 RF Frequency Choices

The choice of rf frequency depends upon the bunch length and the expected phase variations from the damping rings, the sensitivity to multibunch beam loading, the alignment tolerances, and the length and cost of the rf system. The tolerances due to all of these effects are significantly eased by using lower rf frequency with lightly beam-loaded structures. The disadvantage of lower-frequency rf is the system cost and poorer efficiency.

For example, in the low-energy compressor, one can use 139 MV of L-band rf, which requires roughly 10 m of accelerator structure, or 70 MV of S-band rf which could be provided with roughly 3 m of accelerator structure. But, the multibunch loading is more difficult to compensate in the S-band system and would require using the ΔF compensation technique. Therefore, we would need three different structures with slightly different rf frequencies and three klystrons to power the structures. Furthermore, for the same transverse emittance dilution, the transverse alignment tolerances on the S-band structure are more than five times tighter than on the L-band structures. Since

the bunches are long in the bunch compressors the alignment tolerances are tight anyway and the factor of five is significant. For these reasons and because the additional cost is not a significant limitation in these short rf sections, the present design is based on the lower-frequency rf systems. In particular, we use L-band rf (1.4 GHz) in the low-energy compressors and S-band rf (2.8 GHz) in the prelinacs and the high-energy compressors.

In all cases, we will use Damped-Detuned Structures similar to those developed for the X-band main linacs. These minimize the long-range transverse wakefields as well as provide a straightforward method of measuring the induced dipole modes which are used to align the structures to the beam. The structures, power sources, rf pulse compressors, and details of the beam-loading compensation are all described in Chapter 6.

5.2.6 Tuning and Diagnostics

Although the tolerances on components in the bunch compressor systems are not nearly as tight as in the main linacs or the final foci, we have chosen to adopt the same methods of beam-based alignment and tuning. In particular, to ease the alignment procedures, all of the quadrupoles will be mounted on remote magnet movers and will be powered by individual power supplies. In addition, each quadrupole will contain a BPM with a resolution of $1\ \mu$m. Similarly, all of the accelerator structures will be instrumented with an rf BPM to measure the induced dipole modes and each girder will be remotely movable; depending on location, the girders will support either one or two structures. All of these components: the $1\ \mu$m resolution BPMs, the magnet and structure movers, and the structure BPMs, are described in Chapter 7 which discusses the design of the main X-band linacs.

Finally, there are three emittance diagnostic stations located through each compressor system: one after the low-energy compressor, one after the prelinac, and one after the high-energy bunch compressor. Each diagnostic station will contain four laser wire scanners so that emittance measurements can be made without disrupting the collider operation. Lastly, there will be a pulsed beam dump after each diagnostic station so that the various subsystems can be tuned before sending the beams further downstream.

5.3 System Design

In this section, we will describe the present design of the bunch compressor system. We start with a description of the overall layout and then describe the various subsystems in sequence. A discussion of the longitudinal and transverse dynamics in the compressor system can be found in Sections 5.4 and 5.5, respectively. Finally, the system tolerances are discussed in Section 5.6.

5.3.1 Layout

The bunch compressor system starts at the end of the extraction lines from the main damping rings where the beam energy is between 1.8 and 2.2 GeV. It includes the matching and spin rotation sections, the first bunch compressor, and subsequent diagnostics. It also includes the S-band prelinac that accelerates the beam from roughly 2 GeV to 10 GeV. Finally, it includes the second bunch compression system which contains a 180° arc, a compressor rf section, a chicane, and the matching and diagnostics beyond the end of the chicane for injection into the main linac. The compressor system also includes three pulsed beam dumps to allow tuning of the various subsystems before taking the beam further down the collider.

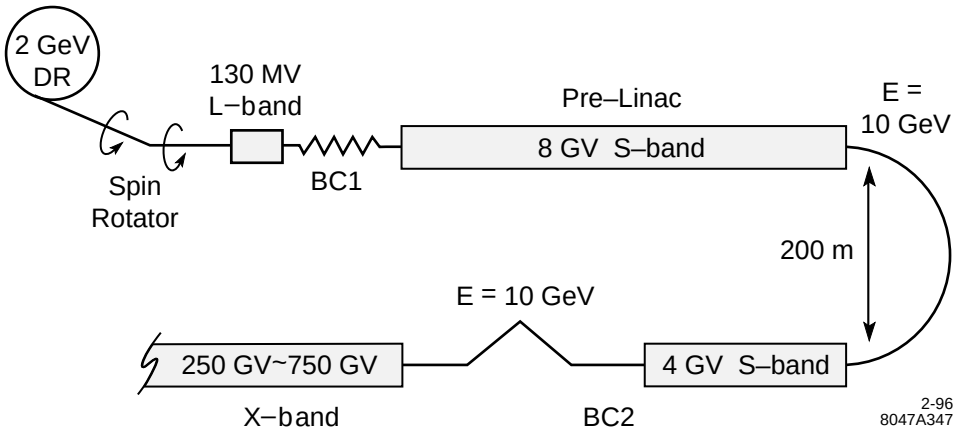


Figure 5-1. *Schematic of the NLC bunch compressor system.*

The length of the spin rotator and first bunch compressor is roughly 170 m while the prelinac is roughly 450-m long. Finally, the 180° arc is roughly 330 m and the second compressor rf and chicane are 350-m long. Thus, the overall system is 1.3 km in length. A schematic view is given in Figure 5-1.

5.3.2 Ring Extraction

The beams are extracted from the main damping rings using an achromatic kicker system and achromatic extraction septum. These are described in Chapter 4 and the component layout is illustrated in Figure 4-14. The bunch compressor beam line begins after the compensation kicker. It starts with six quadrupoles to match the lattice functions into the spin rotator system. A single laser wire scanner is included in this section to provide a continuous diagnostic on the ring performance and quad-scan emittance measurements when necessary. This wire will be used to verify the beta match and minimize any residual dispersion leaking from the damping ring. Finally, there will be a fast-feedback system to control the extraction positions and angles from the damping rings. Similar feedback systems are used in many places through the SLC and are described in Appendix D.

5.3.3 Spin Rotator

The spin rotator system is located immediately after the damping rings. It consists of four paired superconducting solenoids and an intermediate short arc that bends the beam by roughly 20°. The system is designed to operate over the energy range of 1.8 GeV–2.2 GeV. At the nominal energy of 2 GeV, it allows complete control over the spin orientation; at other energies, some of the tuning range is lost but it should still be sufficient. The betatron coupling introduced by the solenoids is canceled by using pairs of solenoids separated by an I transform in X and a $-I$ transform in Y. Although such a cancelation system is sensitive to errors, tuning simulations show that it is straightforward to correct the system; the tuning studies are described in Section 5.5. In this section, we describe the rotator optical design and performance.

Cross-plane Coupling Compensation in a Solenoid Rotator

Solenoid magnets can be implemented to rotate the electron spin about the longitudinal axis by ψ; however, they radially focus the beam and unfortunately introduce a roll about the beam axis [Fieguth 1989]. The damped, flat beam is rolled through an angle ψ_b, which is roughly one-half that of the spin precession.

$$\psi_s = \left[1 - \left(\frac{g-2}{2}\right)\right] \frac{B_z L_s}{(B_0\rho)} \approx \frac{B_z L_s}{(B_0\rho)} = 2\psi_b \tag{5.2}$$

In Eq. 5.2, L_s is the effective solenoid length, B_z is the longitudinal magnetic field and $(B_0\rho)$ is the usual magnetic rigidity. For electrons, $(g-2)/2 \approx 1.16 \times 10^{-3}$, is the anomalous magnetic moment and can be ignored. The large $x-y$ coupling introduced when passing the solenoid can potentially destroy the vertical emittance. It can be shown that the vertical projected emittance at the end of a solenoid of strength ψ_s is

$$\epsilon_y^2 = \epsilon_{x_0}^2 S^4 + \epsilon_{y_0}^2 C^4 + \epsilon_{x_0}\epsilon_{y_0} C^2 S^2 \left(\beta_x\gamma_y - 2\alpha_x\alpha_y + \beta_y\gamma_x\right) \quad , \tag{5.3}$$

where $C \equiv \cos(\psi_s/2)$, $S \equiv \sin(\psi_s/2)$, and $\beta_{x,y}, \alpha_{x,y}, \gamma_{x,y}, \epsilon_{x_0,y_0}$ are the beam Twiss parameters and initial emittances at the solenoid entrance. For $\psi_s = \pi/2$, equal Twiss parameters in the two planes and a horizontal-to-vertical emittance ratio of 100 as in the NLC, the vertical emittance after the solenoid increases as

$$\frac{\epsilon_y}{\epsilon_{y_0}} = \frac{\epsilon_{x_0}}{\epsilon_{y_0}} S^2 + C^2 \approx 50 \quad . \tag{5.4}$$

A coupling compensation system needs to be included which is able to reliably correct this large emittance increase. A system of skew quads can be imagined but has the undesirable character of requiring new skew quad settings for each new solenoid setting. A more robust correction scheme can be designed by splitting the solenoid in half and introducing a canceling symmetry between the two halves [Spence 1991]. The first half-solenoid rotates the beam about the longitudinal axis by $\psi_s/2$ while the spin rotates by ψ_s. If this is followed by a transfer matrix which is $+I$ in the x-plane and $-I$ in the y-plane, the beam is reflected about the y-axis where another half-solenoid of equal strength rotates the beam to its flat state again and a net spin precession of $2\psi_s$ is accomplished. If $k \equiv \psi_s/2L_s = B_z/2(B_0\rho)$, the transfer matrix of the solenoid [Brown 1980] is

$$R_s = \begin{pmatrix} C^2 & SC/k & SC & S^2/k \\ -kSC & C^2 & -kS^2 & SC \\ SC & -S^2/k & C^2 & SC/k \\ kS^2 & -SC & -kSC & C^2 \end{pmatrix} \quad . \tag{5.5}$$

Inserting the reflector beam line between the two solenoids produces a system transfer matrix of

$$R_s \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} R_s = \begin{pmatrix} \cos\psi_s & k^{-1}\sin\psi_s & 0 & 0 \\ k\sin\psi_s & \cos\psi_s & 0 & 0 \\ 0 & 0 & \cos\psi_s & k^{-1}\sin\psi_s \\ 0 & 0 & k\sin\psi_s & \cos\psi_s \end{pmatrix} \quad . \tag{5.6}$$

As long as the solenoids have equal strength ($\pm$ ~1% in this case), all $x-y$ coupling is canceled independent of solenoid settings; the solenoid pairs will be powered in series to insure comparable strengths. Finally, the focusing dependence can be compensated with matching quadrupoles.

Location of the Rotator

Solenoid rotator systems are limited to low-energy applications due to the necessary scaling of solenoid strength with energy. For the NLC, this may seem like a major disadvantage since the second bunch compressor at 10 GeV

includes a 180° arc and the beam delivery section (Collimation–Final Focus) includes a 10-mr arc at 500 GeV to achieve the 20-mr crossing angle. These bends will rotate a non-vertically oriented spin vector many times, potentially depolarizing the beam due to the incoming energy spread. However, a calculation of this depolarization shows it to be small. If the bunch compressor arc at 10 GeV bends the beam by π, the spin will be rotated by $\psi_s = \alpha\gamma_0(1+\delta)\pi$ (11.4 turns) while the beam delivery arc at 500 GeV will rotate the spin vector by 1.8 turns. The change in the spin magnitude will be $P(\delta)/P_0 = \cos(\alpha\gamma_0\delta\pi)$ and the mean polarization over a beam with a Gaussian energy spread σ_δ is

$$\bar{P}/P_0 = \frac{1}{\sqrt{2\pi}}\sigma_\delta \int_{-\infty}^{\infty} e^{-\delta^2/2\sigma_\delta^2}\cos(\alpha\gamma_0\delta\pi)\,d\delta = e^{-(\alpha\gamma_0\pi\sigma_\delta)^2/2} \quad . \tag{5.7}$$

For an energy spread of $\sigma_\delta = 0.25\%$ in the 10-GeV 180° arc, the relative depolarization is 1.6%; it is roughly five times smaller through the beam delivery system. Therefore, the solenoid-based rotator system can be placed at 2 GeV immediately after damping ring extraction with little spin diffusion. Note that the net spin rotation in the low-energy compressor wiggler and the high-energy compressor chicane are zero since there is no net bending and thus these sections cause no depolarization.

Rotator Flexibility

A fully flexible rotator can be made by placing a short horizontal bending section between two split-solenoid segments as previously described. If the four solenoids are each capable of providing a maximum ±45° of spin rotation around the longitudinal axis and the bend section provides 90° around the vertical axis, the system will provide arbitrary control of the IP spin orientation as long as the spin sign is reversible by some means prior to damping. The net spin rotation through the system is symbolized in Eq. 5.8.

$$\Omega_{\rm tot} = \Omega_{sol/34}\Omega_{bend}\Omega_{sol/12} = \begin{pmatrix} c_3 & -s_3 & 0 \\ s_3 & c_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} c_2 & 0 & s_2 \\ 0 & 1 & 0 \\ -s_2 & 0 & c_2 \end{pmatrix}\begin{pmatrix} c_1 & -s_1 & 0 \\ s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.8}$$

Here $c_i \equiv \cos(\psi_i)$, $s_i \equiv \sin(\psi_i)$, where ψ_i is the spin rotation of the ith section and $i = 1,2,3$ indicates the first solenoid pair, the bend section, and the second solenoid pair respectively. Since the input spin vector from damping ring extraction will be vertical, the spin vector after the full system is

$$\bar{s} = \Omega_{\rm tot}\begin{pmatrix} 0 \\ \pm 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \mp\big(\sin(\psi_3)\cos(\psi_1) + \cos(\psi_2)\cos(\psi_3)\sin(\psi_1)\big) \\ \pm\big(\cos(\psi_3)\cos(\psi_1) - \cos(\psi_2)\sin(\psi_3)\sin(\psi_1)\big) \\ \pm\sin(\psi_2)\sin(\psi_1) \end{pmatrix} \quad . \tag{5.9}$$

If the solenoid fields are reversible, then any arbitrary spin orientation is achievable at the nominal beam energy of 2 GeV where $\sin(\psi_2) = 1$; there is a slight loss in the tunable orientation when the damping ring energy is raised or lowered but this is not expected to be significant.

Rotator Beam Line Optics

The reflector beam line (between solenoids) is built with four FODO cells each with 90° betatron phase advance in x and 45° in y. Three cells of 120°, 60° are possible; however, the chromaticity of the 120° cells is larger than desirable. The short bend section (mini-arc) is a simple missing magnet scheme containing six horizontal bend magnets and four 90° x and y FODO cells. The net horizontal bend angle at 2 GeV is $\pi/2\alpha\gamma = 19.83°$. The peak horizontal dispersion is 260 mm and the total R_{56} is 40 mm.

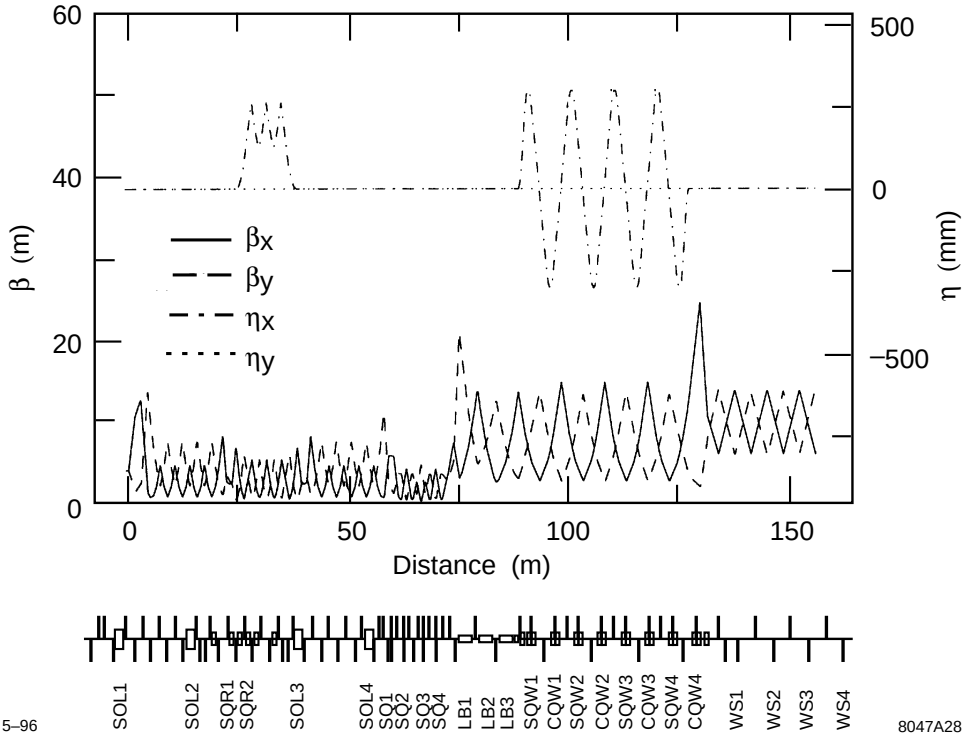


Figure 5-2. *Dispersion and beta functions for the entire low-energy bunch compressor and spin rotator system. Tuning elements and major component locations are indicated. Notation is explained in Section 5.5.*

The solenoids are 1.50 m in length with a maximum field strength of ± 38.5 kGauss; this will rotate the spin by 45° at a beam energy of 2.2 GeV. There is also a four-quadrupole beta matching section between the first solenoid pair segment and the mini-arc, and another matching section between the mini-arc and the second solenoid pair segment. These matching sections are used to maintain the periodic beta functions in the mini-arc through all possible solenoid settings. This is necessary since the solenoids radially focus the beam. The reflector sections provide coupling cancelation while the matching sections provide constant beta functions at the output of the entire system. Figure 5-2 shows the lattice functions for the combined spin rotator and first bunch compressor while Figure 5-3 shows the beta functions of the spin rotator system with (a) solenoids off and (b) solenoids on. With the solenoids at full strength, the x and y beta functions between them are strongly coupled creating very large vertical beta functions due to the locally rolled beam. The dispersion functions are shown in Figure 5-3c.

Chromaticity of the Rotator System

It is important to cancel the large solenoid coupling over a range of particle energies which is comparable to the beam energy spread. The reflector beam line transfer matrix will break down at some level for off-energy particles resulting in energy dependent-coupling. By locating the solenoid rotator system immediately after damping ring extraction and before the first bunch compressor rf section, the energy spread in the solenoids is small ($\sim$0.1% rms). This bandpass is easily achieved in the worst case (all four solenoids at maximum strength) by using FODO cells with $\leq$90° phase advance. Figure 5-4 shows the mono-energetic system bandpass with all four solenoids at full strength, ψ_s(pair) = $\pi/2$. Figure 5-5 shows the emittance as the Gaussian rms energy spread is increased. The vertical emittance increase at 0.1%-rms Gaussian energy spread is $<$0.5%.

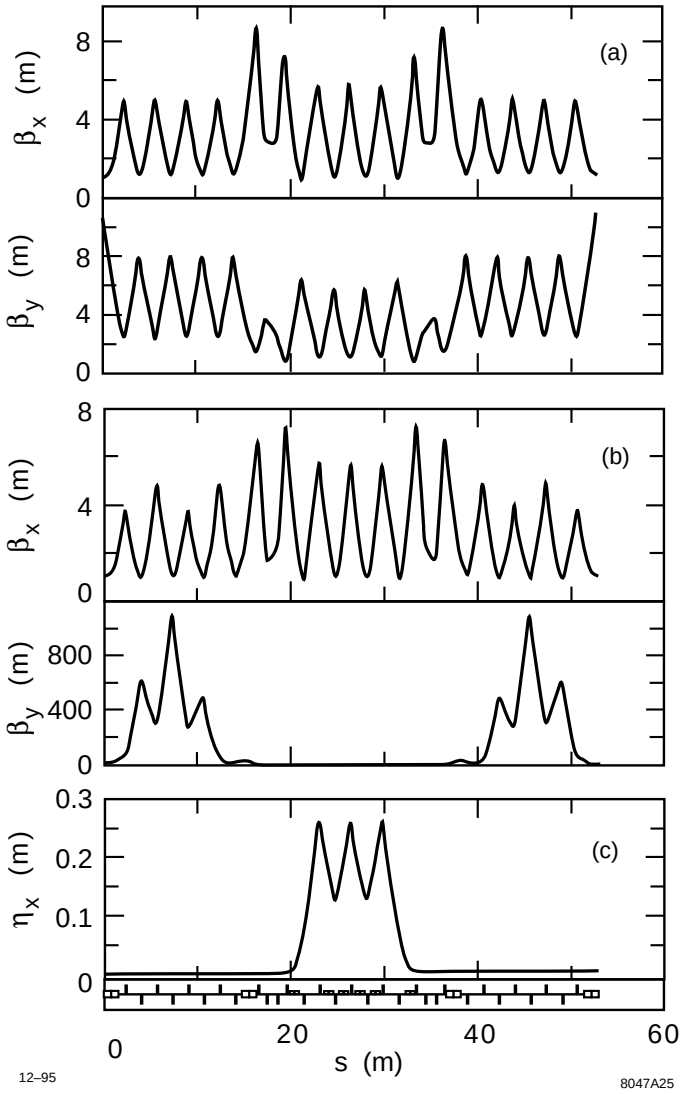


Figure 5-3. *(a) Horizontal and vertical beta functions for spin rotator system with all four solenoids switched off. (b) Beta functions with all solenoids at maximum field, ϕ_s(pair) = $\pi/2$. (c) Horizontal dispersion function for spin rotator system (same for solenoids off or on); the vertical dispersion is zero.*

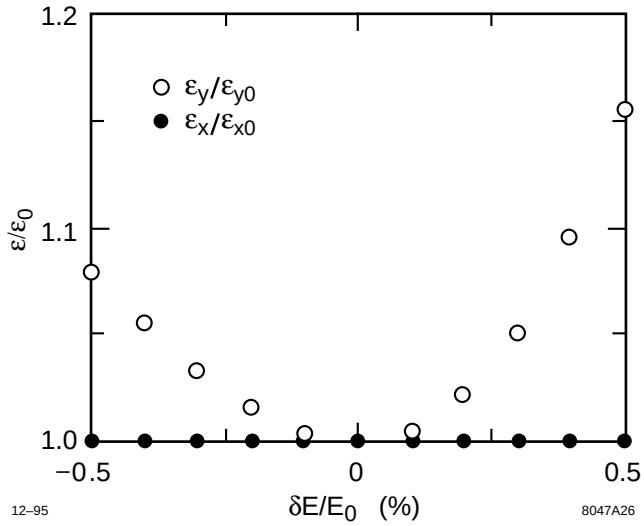


Figure 5-4. *Energy bandpass of the complete solenoid-based rotator system for a mono-energetic beam. The emittance values used are* $\gamma\epsilon_{x_0} = 2.3 \times 10^{-6}$ *m,* $\gamma\epsilon_{y_0} = 2.3 \times 10^{-8}$ *m.*

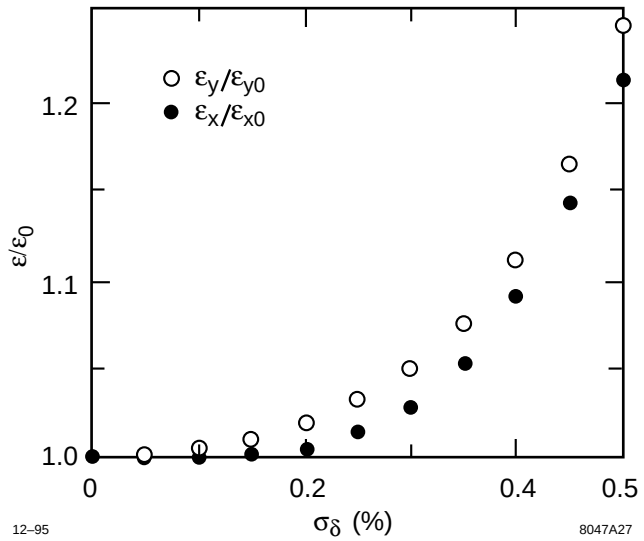


Figure 5-5. *Gaussian rms energy spread dependence of relative chromatic emittance growth.*

Longitudinal Phase Space and Synchrotron Radiation

The rotator system has very little impact on the performance of the first bunch compressor. The longitudinal transfer matrix of the first bunch compressor, not including the spin rotator, is (see Section 5.3.4)

$$\mathbf{R}_{\rm bc1} = \begin{pmatrix} 1+fR_{56} & R_{56} \\ f & 1 \end{pmatrix} \quad . \tag{5.10}$$

To minimize the final bunch length after compression, the rf parameter, f, is chosen so that $1 + fR_{56} = 0$. In this case, adding the spin rotator system with $\alpha \equiv R_{56}$(rot) changes only the R_{66} element of the the total transfer matrix.

$$\mathbf{R}_{\rm bc1}\mathbf{R}_{\rm rot} = \begin{pmatrix} 0 & R_{56} \\ f & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & R_{56} \\ f & 1+\alpha f \end{pmatrix} \tag{5.11}$$

The bunch length after the first compression in unchanged by the rotator and the energy spread after compression is insignificantly smaller ($f = 2\,\mathrm{m}^{-1}, \alpha = -0.04\,\mathrm{m}$).

$$\sigma_{zf} = \sigma_{\delta i} R_{56} \quad , \quad \sigma_{\delta f} = \sqrt{\sigma_{zi}^2 f^2 + \sigma_{\delta i}^2 (1+\alpha f)^2} \tag{5.12}$$

Feedback, Tuning, and Diagnostics Sections

The tightest tolerances on this system are in the construction of the reflector beam line. Without any tuning correction, one of the eight quadrupoles in the reflector would have to be built with a gradient which meets the absolute specification to an accuracy of ∼0.2% to limit the vertical emittance increase to 2%; most other quadrupoles are significantly looser in this tolerance. In any case, a small tunable coupling correction section is already needed to minimize damping ring extraction coupling and this same correction section can be used to cancel any residual coupling due to reflector imperfections.

The tuning of the spin rotator optics, *i.e.,* the betatron match and coupling, will be done with the bunch compressor rf off and use an emittance diagnostic section located at the end of the first bunch compressor–this is described in Section 5.3.4; this procedure is similar to that used in the SLC bunch compressors. The tunability of this system has been studied in some detail using the *Final Focus Flight Simulator* program [Woodley 1994]. To facilitate the tuning, four skew quadrupoles are placed at the end of the rotator system in locations with appropriate phase advances so they are, at least initially, orthonormal to each other [Emma 1995]. In addition, two more skew quadrupoles are placed in the 20° arc to correct any residual vertical dispersion. This system can then be easily compensated for random quadrupole gradient errors of 1% rms and is discussed further in Section 5.5.

Because of the small incoming energy spread, the alignment tolerances are relatively loose, $\sim 50\ \mu$m. Regardless, the alignment of the solenoids and quadrupoles will be adjusted using remote magnet movers similar to those developed for the FFTB. All of the quadrupoles have internally-mounted BPMs and independent power supplies to facilitate the alignment; the tolerances are discussed further in Section 5.6.

Finally, as described earlier, there will be a fast-feedback system in the ring extraction line to control the extraction positions and angles from the damping rings. Because the alignment and vibration tolerances are relatively loose, no other feedback systems are expected in this region.

Energy	2 GeV
σ_z	5 mm $\rightarrow$ 500 μm
σ_ϵ	0.1% $\rightarrow$ 1.0%
V_{rf}	139 MV
f_{rf}	1.4 GHz
L_{rf}	10 m
$\phi_{\rm rf}$	$-101°$
R_{56}	0.5 m
$\Delta\epsilon_{SR}/\epsilon$	$\sim$1.3%
Bend Hor. Aperture	4 cm
Quad. Aperture	1 cm
Length	100 m

Table 5-2. *Parameters for 2-GeV compressor.*

5.3.4 Low-Energy Compressor

The low-energy bunch compressors follow the spin rotators. Assuming an rms energy spread of $\sigma_\delta = 1 \times 10^{-3}$, they compress the damping ring beam to a bunch length of $500\ \mu$m. The final bunch length is roughly independent of the incoming bunch length and the system has been designed to accept an initial bunch length of 5 mm; this is 25% larger than expected from the damping ring. The bunch compressors consist of a 139-MV L-band (1.4 GHz) rf section followed by a long period wiggler which generates the R_{56} needed for the bunch compression; parameters are listed in Table 5-2.

The compressors are designed to rotate the longitudinal phase space by roughly $\pi/2$. This prevents synchronous phase errors, originating due to the transient beam loading in the damping rings, from becoming phase errors, and thereby extraction energy errors, in the linacs. Of course, the phase tolerance is then reflected on the phase stability of the compressor rf system—one reason for choosing the L-band rf frequency is that it is much less sensitive to the multibunch loading which manifests itself as a phase error.

A "$\pi/2$" compressor can be constructed from an rf section introducing a $\delta - z$ correlation and a bending system where the path length is energy-dependent. The linear equations determining the parameters are simple. The transfer matrix is:

$$\mathbf{R} = \begin{pmatrix} 1 + fR_{56} & R_{56} \\ f & 1 \end{pmatrix} \quad , \tag{5.13}$$

where

$$f \equiv \frac{d\delta}{dz} = -\frac{Vk\sin\psi}{E_o + V\cos\psi} \quad . \tag{5.14}$$

Here, R_{56} is the R_{56} of the bending system, V is the rf voltage, k is the rf wave number, and ψ is the rf phase where $0°$ corresponds to running on the rf crest. Finally, the compressed phase, energy, bunch length, etc., are given by:

$$z_f = z_i(1 + fR_{56}) + \delta_i R_{56} \tag{5.15}$$

$$\delta_f = z_i f + \delta_i \tag{5.16}$$

$$\sigma_{zf}^2 = \sigma_{zi}(1 + fR_{56})^2 + 2\langle\delta_i z_i\rangle R_{56}(1 + fR_{56}) + \sigma_{\delta i}^2 R_{56}^2 \tag{5.17}$$

$$\sigma_{\delta f}^2 = \sigma_{zi}^2 f^2 + 2\langle\delta_i z_i\rangle f + \sigma_{\delta i}^2 \tag{5.18}$$

$$\langle\delta_f z_f\rangle = \sigma_{zi}^2 f(1 + fR_{56}) + \langle\delta_i z_i\rangle(1 + 2fR_{56}) + \sigma_{\delta i}^2 R_{56} \tag{5.19}$$

For full compression, the R_{56} is determined by the initial energy spread and the desired bunch length. The rf parameters are chosen so that $f = -1/R_{56}$. In such a design, the only free parameters are the rf frequency and the rf phase—all other parameters are determined by the input beam. Finally, note that such a compressor does not really rotate the longitudinal phase space by $\pi/2$, but the final phase is independent of the initial phase as desired.

Wiggler

The bending system used to generate the R_{56} for the low-energy compressor is constructed as a very long period wiggler. Quadrupoles are placed between the dipoles in locations where the dispersion passes through zero. In this manner, they can constrain the beta functions but they will not generate any second-order dispersion and do not need large apertures. This eliminates the need for strong sextupole magnets which would have tight vertical alignment tolerances due to the coupling they introduce—sextupole correctors and other multipole fields will still be required for tuning purposes. Four wiggler cells are used with a 90° phase advance per cell to reduce the sensitivity to systematic errors in the bending magnets. The optical functions for the entire low-energy compressor, including spin rotator, rf section, and diagnostics section are plotted in Figure 5-2.

The bending magnets and quadrupoles for the low-energy compressor are straightforward. The required apertures are listed in Table 5-2. The main bending magnets are 1.2 m in length with a magnetic field of 14 kGauss; they will be powered as a string by a single power supply to reduce the regulation tolerance on the supply. The quadrupoles are 15 cm in length and have pole-tip fields that are less than 7 kGauss. They will be powered by independent power supplies and mounted on remote magnet movers to facilitate the beam based tuning.

RF Section

The rf section consists of two 5-m L-band rf structures that provide a total rf voltage of 139 MV and operate at a phase of $-101°$ which is slightly past the nominal $-90°$; this is to compensate the T_{566} generated in the wiggler. These structures are located in FODO cells just prior to the beginning of the wiggler. The multibunch beam loading compensation is accomplished using the ΔT method (early injection) and is discussed in Chapter 6. To facilitate the alignment, both structures will be instrumented with rf BPMs and will be mounted on remote movers.

Feedback, Tuning, and Diagnostic Sections

Because of the small beam emittances, the alignment and field tolerances in the compressor would be tight. Fortunately, it is straightforward to use simple tuning techniques to ease the tolerances substantially. The quadrupoles will be powered with independent power supplies and will be mounted on remote movers to facilitate the alignment and the optical tuning will be performed using an emittance diagnostic section consisting of four laser wires located after the compressor wiggler. This is followed by a beam dump that will allow full operation of the damping rings and compressor systems for tuning and commissioning purposes.

Four weak skew quadrupoles and four weak normal quadrupoles will be placed in the dispersive regions of the wiggler to facilitate the tuning of the horizontal and vertical dispersion. In addition, four skew sextupoles will be added to control the vertical second order dispersion. All of the tuning elements and procedures are described in Section 5.5.

The rf voltage and phase of the L-band cavity can be calibrated by varying the rf phase and measuring the beam centroid offsets in the wiggler section. It is more difficult to tune the beam loading compensation in the rf structures. This will be done using a beam phase monitor located at the end of the bunch compressor with a phase resolution of

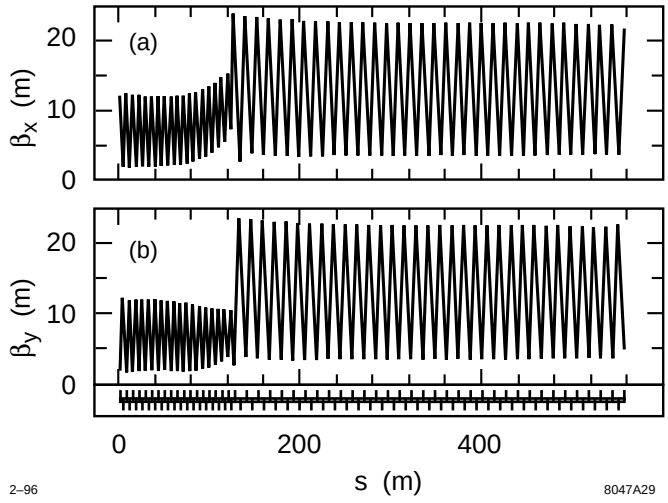


Figure 5-6. *Optical functions in S-band prelinac.*

0.1° S-band. Finally, there will be a phase feedback system that will measure the bunch train centroid through the wiggler and adjust the damping ring synchronous phase; this is discussed in Chapter 4.

5.3.5 Prelinac

Optical Design

The optical design of the prelinac consists of simple FODO cells with an initial quadrupole spacing of 3.1 m to accommodate the 3-m S-band structures. As the beam energy is increased, the quadrupole spacing is increased to a 6-m spacing, accommodating two accelerator structures between quadrupoles. The matching across this transition has been tuned to have a bandwidth in excess of 20%. Finally, note that the horizontal and vertical phase advances are separated by roughly 5%. This reduces the sensitivity to systematic skew quadrupole fields and the nonlinear betatron coupling due to the beam space charge field and trapped ions [Raubenheimer 1992]. The lattice functions are plotted in Figure 5-6.

The quadrupoles are 10 cm in length and have a 1-cm radius with pole-tip fields less than 7.6 kGauss. As is standard practice in the NLC, all of the linac quadrupoles will have internal BPMs and will be mounted on remote magnet movers. In addition, all of the quadrupoles are powered by independent power supplies. This will allow the electrical centers of the BPMs to be accurately determined with respect to the magnetic centers of the quadrupoles and then the quadrupoles can be aligned by adjusting the movers.

RF System

The S-band prelinac consists of 3-m DDS S-band accelerating structures—the structures, klystrons, and rf pulse compressors are described in greater detail in Chapter 6. The linac consists of 33 rf modules with four accelerator structures per module. Each module is powered with two 5045 klystrons and a SLED pulse compressor to produce a net acceleration of 254 MeV. Thus the linac can provide 8.4 GeV of acceleration and would normally operate with one

rf module in standby. The multibunch beam loading is compensated using the ΔT method (early injection) so that the bunch-to-bunch energy deviations are corrected locally in each structure; this minimizes the emittance growth due to dispersive and chromatic effects. The expected bunch-to-bunch energy variation is less than 10^{-4}; this is discussed in Chapter 6.

Finally, all of the accelerator structures will have rf BPMs and will also be mounted on girders that can be remotely aligned; there is one structure support girder between quadrupoles and thus at the beginning of the prelinac there is only one structure per girder while at the end of the linac there are two structures per girder. This will allow the structures to be aligned to the beam trajectory. The alignment techniques are described in Section 5.5.

Feedback, Tuning, and Diagnostic Sections

As noted earlier, there is a diagnostic section, consisting of four laser wire scanners, after the low-energy bunch compressor at the beginning of the prelinac. This is followed by a pulsed dump which can absorb a full bunch train at the nominal repetition rate so that the damping rings and first bunch compressors can be tuned before the beams are sent through the prelinacs.

A second diagnostic section, again consisting of four laser wire scanners, is located at the end of the prelinac. To save space, this section is integrated into the linac lattice. It is followed by another pulsed beam dump to facilitate the tuning of the prelinac. As described, the transverse alignment of the quadrupoles and accelerator will be performed using remote magnet movers; details are described in Section 5.5.

The rf phases and the beam loading compensation will be tuned using the spectrometer in the beam dump or using the 180° arc as a spectrometer. The BPMs in the arc have a resolution of 1 μm and will be able to resolve relative energy variations of roughly 10^{-5}; this is more than sufficient to optimize the klystron phases and the beam loading compensation.

Finally, one fast-feedback system will be used to control the transverse position and angle of the bunch train as it is injected into the prelinac and another will be located at the transition region where the quadrupole spacing is increased from 3 to 6 m. In addition, an energy feedback will be used to stabilize the beam energy; this will adjust the phases of the two klystrons feeding one of the rf modules to vary the beam voltage.

5.3.6 High-Energy Compressor

The high-energy bunch compressors follow the prelinacs. Assuming an rms bunch length of $\sigma_z = 500\,\mu$m, they compress the beam to a bunch length between 100 and $150\,\mu$m depending on the NLC parameters. The bunch compressors consist of a 180° arc which is followed by a 4-GeV S-band (2.8 GHz) rf section and a chicane. Parameters of the high-energy bunch compressors are listed in Table 5-3 for four different NLC scenarios.

As described, the second compressor is designed to rotate the longitudinal phase space by 2π; it is a telescope in longitudinal phase space. This requires two rf sections and two bending regions. We can use the prelinac as the first rf region. Then, a 180° arc will be used for the first bending region, followed by an rf section and then a magnetic chicane as the second bending region. The linear equations that determine the compressor are:

$$\mathbf{R} = \begin{pmatrix} 1 + f_2 R_{56\,2} + f_1(f_2 R_{56\,1} R_{56\,2} + R_{56\,1} + R_{56\,2}) & R_{56\,1} + R_{56\,2} + f_2 R_{56\,1} R_{56\,2} \\ f_1 + f_2 + f_1 f_2 R_{56\,1} & --- \end{pmatrix}, \tag{5.20}$$

	NLC-I a	NLC-I c	NLC-II a	NLC-II c
Energy	10 GeV			
Initial σ_z	500 μm			
Final σ_z	100 μm	150 μm	125 μm	150 μm
Initial σ_ϵ	0.25 %			
Final σ_ϵ	1.5 %	1.2 %	1.3 %	1.4 %
180° arc R_{56}	0.211	0.147	0.184	0.162
Arc Length	330 m			
$\Delta\epsilon_{SR}/\epsilon$ (arc)	2.2%	1.5%	2.0%	1.6 %
$V_{\rm rf}$ (GV)	3.85	3.50	3.73	3.61
$f_{\rm rf}$	2.8 GHz			
$L_{\rm rf}$	250 m			
$V_{{\rm rf}\ c}$ (MV)	274	226	256	240
$f_{{\rm rf}\ c}$	11.4 GHz			
$L_{{\rm rf}\ c}$	7.6 m			
Chicane R_{56}	36 mm			
Chicane Length	100 m			
$\Delta\epsilon_{\rm SR}/\epsilon$ (chicane)	$\sim 0.05\%$			
Bend Hor. Aperture	10 cm			

Table 5-3. *Parameters for 10-GeV compressor.*

where $R_{56\,1}$ and $R_{56\,2}$ are the R_{56}s of the first and second bending regions and f_1 and f_2 are the rf parameters for the first and second rf regions. Finally, the (2,2) matrix element can be found from the other (the determinant equals 1) but it is not relevant for our study.

To make the final phase independent of the incoming energy and only sensitive to the incoming phase, we require that the (1,2) matrix element equal zero. In addition, to compress the bunch by a factor m, we require that the first matrix element be equal to $\pm 1/$ m. Finally, to perform a full $n\pi$ rotation, we require that the (2,1) element also be zero, but this is not actually very important for us since the incoming phase errors should be small. Thus, no restriction needs to be placed on the rf of the prelinac and the two constraints on the compressor design are:

$$1 + f_2 R_{56\,2} = \pm 1/m \qquad\qquad R_{56\,1} = \mp m R_{56\,2} \tag{5.21}$$

where the upper sign corresponds to a 2π rotation and the lower signs are for a π rotation of the phase space.

Because the NLC collider geometry suggests using a 180° arc and a chicane which have opposite values of R_{56}, we have designed the compressor to perform a 2π rotation. Now, given a value of the R_{56} in the chicane and a chosen rf frequency (S-band), the R_{56} of the arc and the parameters of the rf section between the arc and the chicane are determined. To reduce nonlinearities of the longitudinal phase space, the R_{56} of the chicane should be as small as possible. However, a smaller R_{56} increases the required rf voltage, so a compromise had to be found. According to longitudinal single-bunch simulations, an R_{56} of 100 mm is too large and impairs the performance of the compressor. In the present design, therefore, the R_{56} of the chicane is chosen as 36 mm, which implies a 200-m-long S-band rf section.

Parameter	Symbol	Unit	$\psi_x = 90°$	$\psi_x = 108°$	$\psi_x = 135°$
Momentum compaction	R_{56}	m	-0.286	-0.208	-0.150
Synch. rad. emit. growth	$\Delta\epsilon_{\rm SR}/\epsilon_0$	%	3.7	2.4	1.6
Max. hor. beta function	$\beta_x^{\rm max}$	m	7.27	7.67	10.35
Max. hor. dispersion function	$\eta_x^{\rm max}$	m	0.133	0.102	0.078
F-quad pole-tip magnetic field	B_{QF}	kGauss	$+7.15$	$+8.08$	$+9.26$
D-quad pole-tip magnetic field	B_{QD}	kGauss	-7.15	-7.32	-7.50
F-sext. pole-tip magnetic field	B_{SF}	kGauss	$+1.18$	$+1.74$	$+2.63$
D-quad pole-tip magnetic field	B_{QD}	kGauss	-2.06	-2.98	-4.43
# FODO cells	$N_{\rm cell}$	—	68		
Energy	E	GeV	10		
Total arc length	$L_{\rm tot}$	m	295		
Total bend angle	θ	°	180		
Mag. length of quadrupoles	L_Q	m	0.32		
Mag. length of dipoles	L_B	m	1.25		
Mag. length of sextupoles	L_S	m	0.10		
Drift between every magnet	$L_{\rm BPM}$	m	0.3		
Vert. betatron phase adv. per cell	ψ_y	°	90		
Quadrupole magnet pole radius	r_Q	mm	10		
Sextupole magnet pole radius	r_S	mm	10		
Bend angle of dipole	θ_B	°	1.304		
Bend radius of dipole	ρ	m	54.9		
Spin advance per cell	$\Phi_{\rm Spin}$	°	59.2		

Table 5-4. *Parameters for 180° arc of high-energy compressor for several momentum compaction factors adjusted by varying the x-betatron phase advance per cell, ψ_x.*

Arc

The bending system used to generate the R_{56} for the 180° arc is constructed from a series of FODO cells. Although one could reduce the emittance growth through this section using combined-function bending magnets, we have chosen separated-function magnets so that one can use beam-based alignment techniques to align the quadrupole magnets.

We have studied several versions of the arc, with different numbers of cells. In all cases, the R_{56} is easily adjusted with the horizontal phase advance and the choice of the number of cells depends upon the tuning range desired. The present arc design consists of 68 FODO cells, and has an average radius of about 100 m. Variation of the horizontal phase advance per cell from 90° to 135° changes the R_{56} from 29 cm to 15 cm, which results in an increase of the final bunch length from 100 μm to 150 μm. This variability covers the entire NLC operating plane. While the horizontal phase advance is varied, the vertical phase advance per cell is held constant at 90°. Figure 5-7 illustrates the tuning range of the R_{56} in the arc and the corresponding emittance growth due to synchrotron radiation. In addition, the beta functions and dispersion functions are plotted in Figures 5-8 and 5-9. Finally, the arc parameters are summarized in Table 5-4.

The arcs are constructed using bending magnets that are 1.25 m in length and have a maximum field of 6 kGauss and quadrupole magnets that are 32-cm long and have a 1-cm radius with pole-tip fields of 8 kGauss. To relax the stability and voltage requirements on the bending magnet power supplies, 18 bends are powered as a string by a single power supply. As stated, all of the quadrupoles have independent power supplies to facilitate the beam-based alignment.

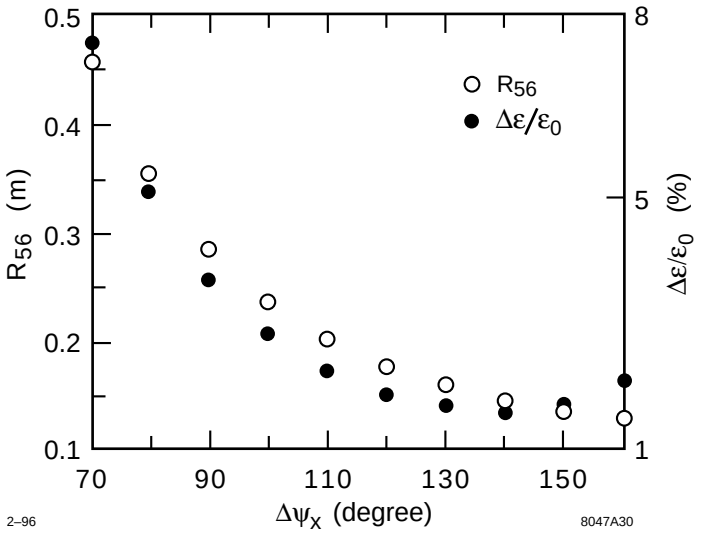


Figure 5-7. *R_{56} (solid) and $\Delta\epsilon_x$ (dashes) for 68-cell arc.*

Sextupole compensation is also included in the optical design to desensitize the arc to the generation of dispersion due to incoming betatron oscillations. This also desensitizes the arc to the generation of second order dispersion due to a dispersion oscillation through the entire arc. Without sextupole compensation a vertical betatron oscillation (jitter) which propagates through the entire arc will generate vertical dispersion due to the chromaticity of the quadrupoles. For an initial oscillation amplitude of $\eta_\sigma = y_0/\sqrt{\beta_y \varepsilon_{y0}}$ in the first QD ($y_0' = 0$) which oscillates through an arc with $N(\gg 1)$ FODO cells of phase advance per cell μ and rms energy spread σ_δ, the vertical emittance dilution due to dispersion is approximately (thin lens QDs only)

$$\frac{\Delta\varepsilon}{\varepsilon_0} \approx \frac{1}{2}\left(n_\sigma N \sigma_\delta \frac{1+\sin\mu/2}{\cos\mu/2}\right)^2 \quad . \tag{5.22}$$

For the high-energy bunch compressor (BC2) arc parameters without sextupoles, an initial vertical oscillation of $n_\sigma = 1$ ($y_0 = 3.3\,\mu$m) would induce a 10% emittance dilution due to vertical dispersion. Likewise, a small vertical dispersion error of 1 mm (at QDs) through the entire arc would generate second-order dispersion which dilutes the emittance by 10%. The inclusion of sextupoles opens these oscillation tolerances by more than a factor of 10 and actually allows full recovery of the emittance using the simple linear tuning schemes discussed in Section 5.5.3, even for extreme vertical emittance dilution cases of $\Delta\varepsilon_y/\varepsilon_{y0} \sim 30$. The sextupoles are 10-cm long with 1-cm pole radius and 1.7 and -3.0 kGauss pole-tip fields for the `focusing' (SF) and `defocusing' (SD) sextupoles, respectively. Their alignment tolerances are similar to the quadrupole alignment tolerances of roughly $60 \sim 100\,\mu$m and are discussed in Section 5.5.3.

Figure 5-10 is a plot of the emittance dilution, without synchrotron radiation effects, at the end of the 68-cell arc as a function of the incoming energy spread for a horizontal phase advance of 108°. For the nominal operation parameters, the emittance growth is less than 0.5%.

RF Section

The high-energy bunch compressor rf sections consist of roughly 4 GeV of S-band linac. The linacs are similar to the S-band prelinacs, consisting of 3-m DDS S-band accelerator structures. They differ from the prelinacs in

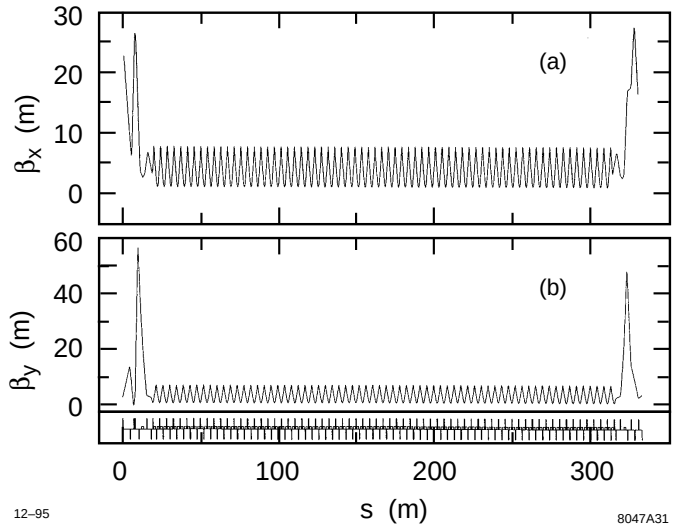


Figure 5-8. *Beta functions for 68-cell arc.*

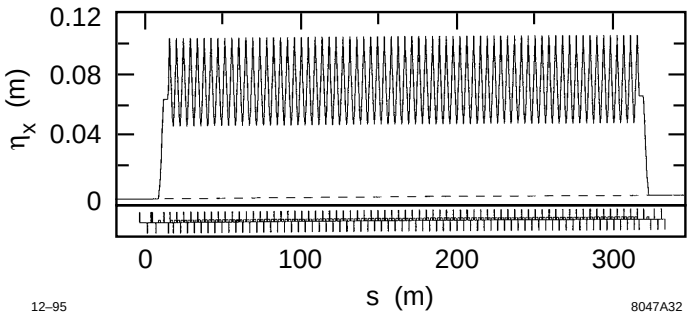


Figure 5-9. *Dispersion function for 68-cell arc.*

that the multibunch beam loading is compensated using a combination of the ΔT method (early injection) and the ΔF method; this is necessary to attain the required compensation. The linacs consist of 16 rf modules with a net acceleration of 253 MeV per module. Thus, they can provide 4 GeV of acceleration and would normally operate with at least one rf module in standby; see the parameters in Table 5-3. Each module has five accelerator structures, four tuned to the nominal frequency and one structure that is either 1.1 MHz above the nominal 2.856 GHz or 1.1 MHz below 2.856 GHz; the sign of the frequency compensation alternates between modules. The four normal structures in each module are powered with two 5045 klystrons and a SLED pulse compressor while the off-frequency structure is powered with a single 42-MW klystron. The structures, klystrons, rf pulse compressors, and beam loading compensation are described in greater detail in Chapter 6.

Because the bunch length is long in these rf sections, the transverse wakefields are significant. Furthermore, because the chicane will rotate the longitudinal phase space, the transverse emittance dilutions will not be correctable after the beam passes the chicane. To reduce the emittance dilution, this rf section has strong focusing with a 3.1-m quadrupole spacing. To attain the required fields, the quadrupoles are 15 cm in length with a 1-cm aperture and maximum pole-tip fields of 7 kGauss.

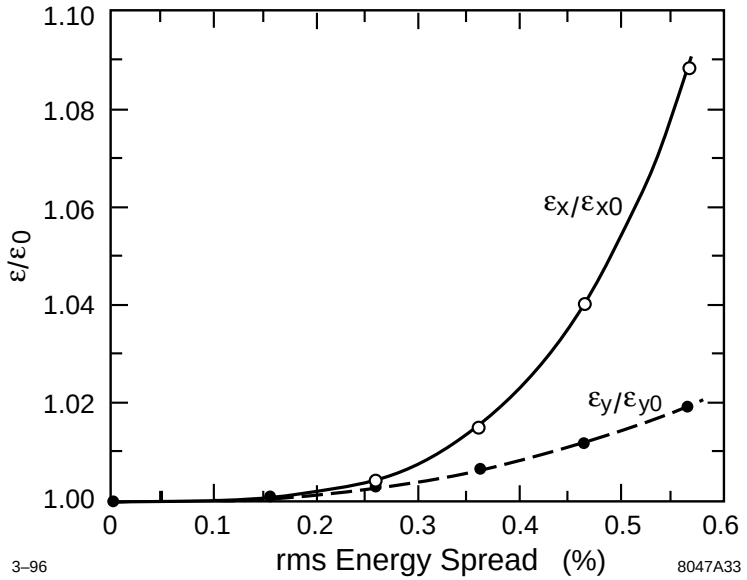


Figure 5-10. *Chromatic emittance growth for nominal operation, ($R_{56} = 210$ cm).*

Finally, the high-energy compressor also needs a higher frequency rf to compensate the T_{566} from the chicane. The present design requires roughly 275 MV of X-band rf provided in three 1.8-m structures.

Chicane

The chicane is constructed from sixteen 2.5-m bending magnets having a 1.1-kGauss field; the beta functions and the dispersion function are plotted in Figures 5-11 and 5-12. The lengths and spacing are chosen so as to achieve the R_{56} of 3.6 cm in a reasonable length, without generating substantial emittance dilution due to synchrotron radiation; all of these are subjective constraints and the parameters might be optimized in another manner.

One potential issue is the peak value of the dispersion which is roughly 65 cm. This means that the horizontal good field aperture of the bending magnets must be large ($\sim$ 10 cm full width) to provide an aperture in excess of ± 5 sigma and without tuning correction, the alignment and field tolerances will be severe. The field tolerances can be attained by using large aperture magnets and careful construction techniques and, fortunately, the tuning of the errors is straightforward as described in Section 5.5.

Feedback, Tuning, and Diagnostic Sections

A diagnostic section consisting of four laser wires is located at the end of the chicane so that the chromatic properties of the chicane can be tuned properly for injection into the main X-band linacs. Presently, it is thought that the 180° arc and the rf section can be tuned by turning the compressor rf off to minimize the beam energy spread and using the diagnostic section at the end of the chicane. If necessary, an additional diagnostic section can be integrated into the end of the high-energy compressor rf section before the chicane. Finally, there are pulsed beam dumps just before the beginning of the main linacs that allow the whole injector complex to be tuned before accelerating the beams to high energy; the tuning elements and procedures are described in Section 5.5.3.

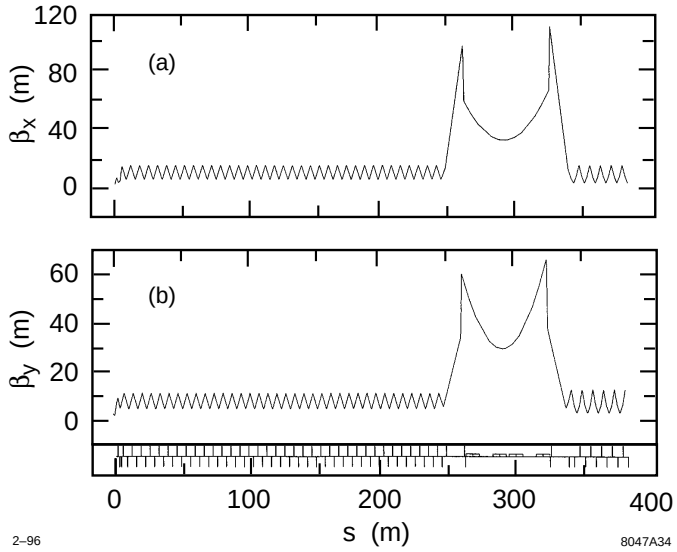


Figure 5-11. *Beta functions for chicane.*

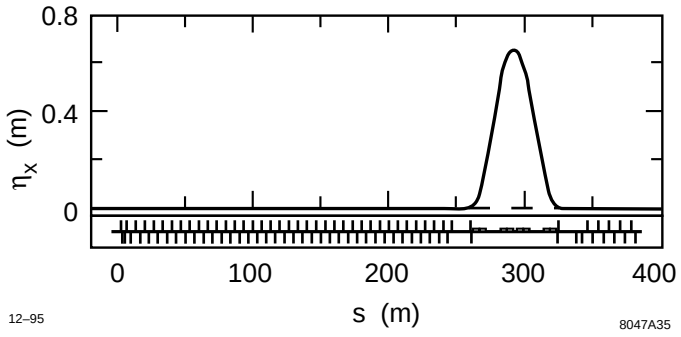


Figure 5-12. *Dispersion function for chicane.*

5.4 Longitudinal Dynamics

The primary purpose of the bunch compressor system is to rotate the longitudinal phase space from the damping rings by $\pi/2$ while compressing the bunch length. In this process, nonlinearities distort the phase space, changing the bunch shape and the sensitivity to incoming phase errors from the damping rings. In this section, we discuss the sources of nonlinearity and their compensation and we present simulation results illustrating the longitudinal beam distributions at the IP.

Finally, it should be noted that most of these simulations were performed before the structure parameters were finalized. Thus, the wakefields that were used in the simulations are stronger than those expected and only the ΔF technique was used for the beam loading compensation. Both of these differences will cause the simulations to *overestimate* the sensitivities and errors. Since this document is primarily a feasibility study and the feasibility is clearly demonstrated, we have not yet re-calculated the effects.

5.4.1 Optical and RF Nonlinearities

Having been given all parameters as described and having generated the desired linear transfer matrix for the longitudinal phase space, the tracking simulations still show a fairly large variation of rms energy spread, and, in particular, of the mean energy at the end of the main linac as a function of initial phase error. For instance, a 6-mm initial phase error causes a change of the final beam energy by about 0.2%–0.3%, which is too large to be tolerable. Furthermore, the mean energy and energy spread show a very asymmetric dependence on positive and negative phase errors.

An important nonlinear aberration of the system is due to the quadratic dependence of the final longitudinal phase on the incoming energy (the T_{566} transfer-matrix element in TRANSPORT notation [Brown 1977]) for the different subsystems. It is straightforward to see (compare Section 5.4.2) that for wiggler and chicane we have

$$T_{566}^{w,c} = -\frac{3}{2} R_{56}^{w,c} \tag{5.23}$$

while for the present lattice of the arc we find

$$T_{566}^{a} \approx 1.9 \cdot R_{56}^{a} \quad . \tag{5.24}$$

It is the nonzero $T_{566}^{w,c}$ coefficient of wiggler and chicane which causes the intolerably large variation of the final energy at the end of the main linac as a function of initial phase error. (The T_{566}^{a} of the arc is not significant.)

There are two effects that are important. First, initial phase errors give rise to energy errors δ_1 after the first compressor rf, which in turn, due to the T_{566}^{w} of the wiggler, cause a phase offset in the prelinac. The result is an additional energy change $\Delta\delta_{pl}$ which may either add to or cancel the energy error δ_1, dependent on the sign of the offset. The second compressor further enhances the energy error and, due to the T_{566}^{c} of the chicane, generates a significant phase error in the main linac. A second effect is that, if the bunch is longitudinally off-set, the T_{566} can either increase or counteract the nonlinearity of the rf, giving rise to an asymmetry in bunch length versus phase error, for each compressor stage separately.

The simulation results show that it is desirable to compensate the sensitivity to initial phase errors caused by the $T_{566}^{w,c}$ of wiggler and chicane. A compensation can be performed, for instance, with an additional, decelerating rf [Raubenheimer 1994b], whose purpose is to cancel the quadratic dependence of the final phase z_f on the longitudinal position (the phase error) at the compressor rf. Consider the main compressor rf

$$V \cos(\psi + kz) \tag{5.25}$$

and a compensating rf

$$V_c \cos(\psi_c + k_c z) \tag{5.26}$$

with V_c small compared with the beam energy E, $V_c \ll E$, followed by a chicane or a wiggler. The phase z_f at the exit of this system depends on the initial phase z_{rf} at the rf as

$$\begin{aligned} z_f &= (1 - R_{56}^{w,c}(f + f_c)) z_{rf} + T_{566}^{w,c}(f + f_c)^2 z_{rf}^2 \\ &\quad + \frac{1}{2} R_{56}^{w,c}(fk \cot\psi + f_c k_c \cot\psi_c) z_{rf}^2 \end{aligned} \tag{5.27}$$

where we have omitted a constant shift as well as higher-order terms, $f_c \equiv V_c k_c \sin\psi_c/(E + V\cos\psi)$ for the compensating rf, and $f \equiv Vk \sin\psi/(E + V\cos\psi)$ is the corresponding value for the main compressor rf; $k \equiv 2\pi/\lambda$ is the wave number. The sum of the quadratic terms in Eq. 5.27 is zero if we choose

$$T_{566}^{w,c}(f + f_c)^2 = \frac{1}{2} R_{56}^{w,c}(fk \cot\psi + f_c k_c \cot\psi_c) \tag{5.28}$$

which, for the special case $\psi = -\pi/2$ and $\psi_c = \pi$, simplifies to

$$V_c = -2\frac{T_{566}^{w,c}}{R_{56}^{w,c}}\frac{k^2V^2}{k_c^2E} = 3\frac{\lambda_c^2V^2}{\lambda^2E} \tag{5.29}$$

where we have used Eq. 5.23. Alternately, assuming no additional compensating rf system and only changing the rf phase, we find:

$$\cot(\psi) = 3\frac{f}{k} \quad . \tag{5.30}$$

In the first compressor stage, we can simply change the L-band rf phase from $-90°$ to $-101°$ to perform the compensation. Unfortunately, in the second compressor this is more difficult; the S-band rf would have to be run roughly $-50°$ from the zero crossing to provide sufficient curvature. Instead, we plan to use 275 MV of X-band rf operating at a phase of $-180°$ to perform the compensation.

With the compensation, the residual longitudinal aberrations are now third order and higher. An initial phase error or multibunch beam loading may cause an energy offset δ of the entire bunch at the entrance to the high-energy compressor (BC2). In addition, the longitudinal wakefields in the prelinac induce a (mainly) quadratic δ-z-correlation, where z is the longitudinal position of a single particle with respect to the bunch center. Due to the R_{56}^a of the arc, this energy error translates into a position error at the two rf systems of BC2, which is transformed once more into energy by the rf, and back into longitudinal phase by the T_{566}^c of the chicane. The final phase z_f of a single particle at the exit of BC2 is

$$z_f \approx az - bz^4 - c\delta^2z^2 - d\delta z^2 - e\delta z^4 + \ldots \tag{5.31}$$

where $a \approx 1/5$–$1/7$ is the desired linear compression ratio, and the four nonlinear terms on the right-hand side are about the same size (10–20 μm), for typical values $\delta \approx \pm 0.004$ and $z \approx 500\,\mu$m. These four terms are negligibly small if the prelinac wakefields are absent.

5.4.2 The T_{566}-Matrix Element

Let the trajectory of an arbitrary particle be described by a general dispersion function η as $x = \eta\delta$, where δ is the relative momentum deviation. The dispersion function may be expanded as a power series in δ:

$$\eta\delta = \eta_0\delta + \eta_1\delta^2 + \ldots \tag{5.32}$$

For a system of bending magnets and quadrupoles with bending radius ρ and gradient k, the functions η_0 and η_1 fulfill the differential equations [Delahaye 1985]

$$\eta_0'' + \left(\frac{1}{\rho^2} + k\right)\eta_0 = \frac{1}{\rho} \tag{5.33}$$

$$\eta_1'' + \left(\frac{1}{\rho^2} + k\right)\eta_1 \approx -\frac{1}{\rho} + \left(k + \frac{2}{\rho^2}\right)\eta_0 \tag{5.34}$$

which, for $\eta_0 \ll \rho$, may be further approximated by

$$\eta_0'' + k\eta_0 \approx \frac{1}{\rho} \tag{5.35}$$

$$\eta_1'' + k\eta_1 \approx -\frac{1}{\rho} + k\eta_0 \tag{5.36}$$

The R_{56}- and T_{566}-matrix-elements may be expressed in terms of $\eta_{0,1}$ [Delahaye 1985]

$$R_{56} = -\int \frac{\eta_0}{\rho} ds \tag{5.37}$$

$$T_{566} \approx -\int \left(\frac{\eta_1}{\rho} + \frac{1}{2}\eta_0'^2 \right) ds \tag{5.38}$$

In case of wiggler and chicane $k\eta_{0,1} = 0$ and, therefore, $\eta_1 = -\eta_0$. Since $\eta_0'' \approx 1/\rho$, the second term in the expression for T_{566} can be integrated by parts, with the final result

$$T_{566} \approx -\frac{3}{2} R_{56} \qquad \text{(for wiggler and chicane)} \quad . \tag{5.39}$$

To reduce the T_{566}-matrix-element, the value of the second-order dispersion needs to be reduced. Ideally, one would like to have $T_{566} = 0$, which requires a sign reversal of η_1 such as to cancel the contribution from $\eta_0'^2$. This can be accomplished by adding quadrupoles at locations with nonzero η_0, as evident from Eq. 5.36, or by adding sextupoles (not included in the above equations). Unfortunately, such changes to η_1 may adversely affect the transverse emittances.

5.4.3 Single Bunch Longitudinal Wakefields

Longitudinal wakefields in the S-band accelerator seriously modify the longitudinal phase space. For our calculations we have assumed the SLAC wakefield. The actual wakefield for the NLC will be roughly $\sqrt{2}$ smaller because of the increased iris aperture. Thus, we have over-estimated the tolerances and sensitivities.

5.4.4 Multibunch Wakefields

Thus far, the simulations have included only the fundamental longitudinal mode for the different S-band, L-band, and X-band rf sections. Parameters were provided by R. Miller [Miller 1995] and are listed in Table 5-6; these parameters differ slightly from the structure design parameters described in Chapter 6.

5.4.5 Bunch Shaping

At this time, we have not investigated the option of deliberately shaping the longitudinal profile of the bunch to reduce the energy spread induced by the longitudinal wakefields and/or the transverse emittance dilution. However, a two-stage bunch-length collimation section is part of the bunch-compressor and linac design; it is discussed in Chapter 9. This system, by itself, already reduces the beam energy spread, and it would be an integral part of any bunch-shaping strategy. More studies on this topic are left for the future.

5.4.6 Longitudinal Phase Space at the End of the Linac

To determine the optimum parameters of the system, we have considered four NLC scenarios, corresponding to different bunch lengths and energies (see Table 5-5). The parameters chosen enclose the NLC operating plane as described in Chapter 1.

	NLC-I		NLC-II	
Cms energy	500 GeV	500 GeV	1 TeV	1 TeV
Bunch length	100 μm	150 μm	125 μm	150 μm
Single-bunch gradient	33.25 MV/m	29.13 MV/m	60.19 MV/m	53.97 MV/m
$L_{\rm acc}$	8130 m	8130 m	8900 m	8900 m
N/bunch (10^{10})	0.65	0.85	0.95	1.25
$E_{\rm max}$ (GeV)	267	233	534	473
Average $\phi_{\rm rf}$ ($^\circ$)[a]	-14.4	-15.1	-7.2	-11.0

[a] with respect to beam center, for 0.8% FWHM energy spread.

Table 5-5. *Four different NLC scenarios considered in the bunch-compressor simulations.*

Computer simulations of the longitudinal single-bunch dynamics have been performed with the code LITRACK [Bane]. In these simulations, a distribution of particles as extracted from the damping rings is tracked through the different compressor subsystems and the main X-band linac. The parameters of the bunch compressor are optimized such that the final average energy and the energy spread at the end of the main linac are insensitive (*i.e.*, vary by less than 0.1%–0.2%) to initial phase errors up to 20° S-band, or ± 6 mm.

In the latest simulation studies, the phase of the main X-band linac was adjusted to obtain a final full-width, half-maximum (FWHM) energy spread of about 0.8%. This energy spread is slightly smaller than the energy bandwidth of the final focus (see Chapter 11) and it might be reduced further by tuning the rf properties in the X-band linac. Figures 5-13 through 5-16 present tracking results for two different bunch lengths in either NLC-I and NLC-II. Shown is the longitudinal distribution at the exit of the bunch compressor and, for five different initial phases, at the end of the main linac. Figure 5-17 presents the dependence of the average energy on the initial phase error, for the same four NLC scenarios. Figures 5-18 and 5-19 show equivalent pictures for the rms and FWHM energy spread, respectively. In all cases, the beams are acceptable for the final focus.

5.4.7 Multibunch Dynamics

The real purpose of the bunch compressors is not to reduce the length of a single bunch, but the length of each bunch in the train of 90 bunches. This task is complicated by the longitudinal long-range wakefields and the pertinent multibunch beam loading in the various accelerator sections of the compressor.

In this section, we will evaluate the ΔF compensation technique for the multibunch beam loading which adds two additional rf systems with slightly different frequencies to each acceleration section [Kikuchi 1992]. These rf systems generate an additional voltage which increases almost linearly along the bunch train, and which can be adjusted to cancel the linear part of the multibunch beam loading. It should be noted that the present rf system design actually uses a combination of ΔT and ΔF techniques that is much more effective than just the ΔF technique described here.

A simulation study has been performed to test this multibunch energy compensation scheme and the interplay of longitudinal single- and multibunch dynamics in the bunch compressor. For the purpose of this study, the long-range wakefields have been confined to the fundamental mode. The wakefield of a structure is then characterized by three numbers: the loss factor k, the mode frequency f, and the quality factor Q. In terms of these quantities, the wakefield at a distance z reads

$$W(z) = 2k\cos\left(\frac{2\pi f z}{c}\right) e^{-\frac{2\pi f z}{2Q}} \tag{5.40}$$

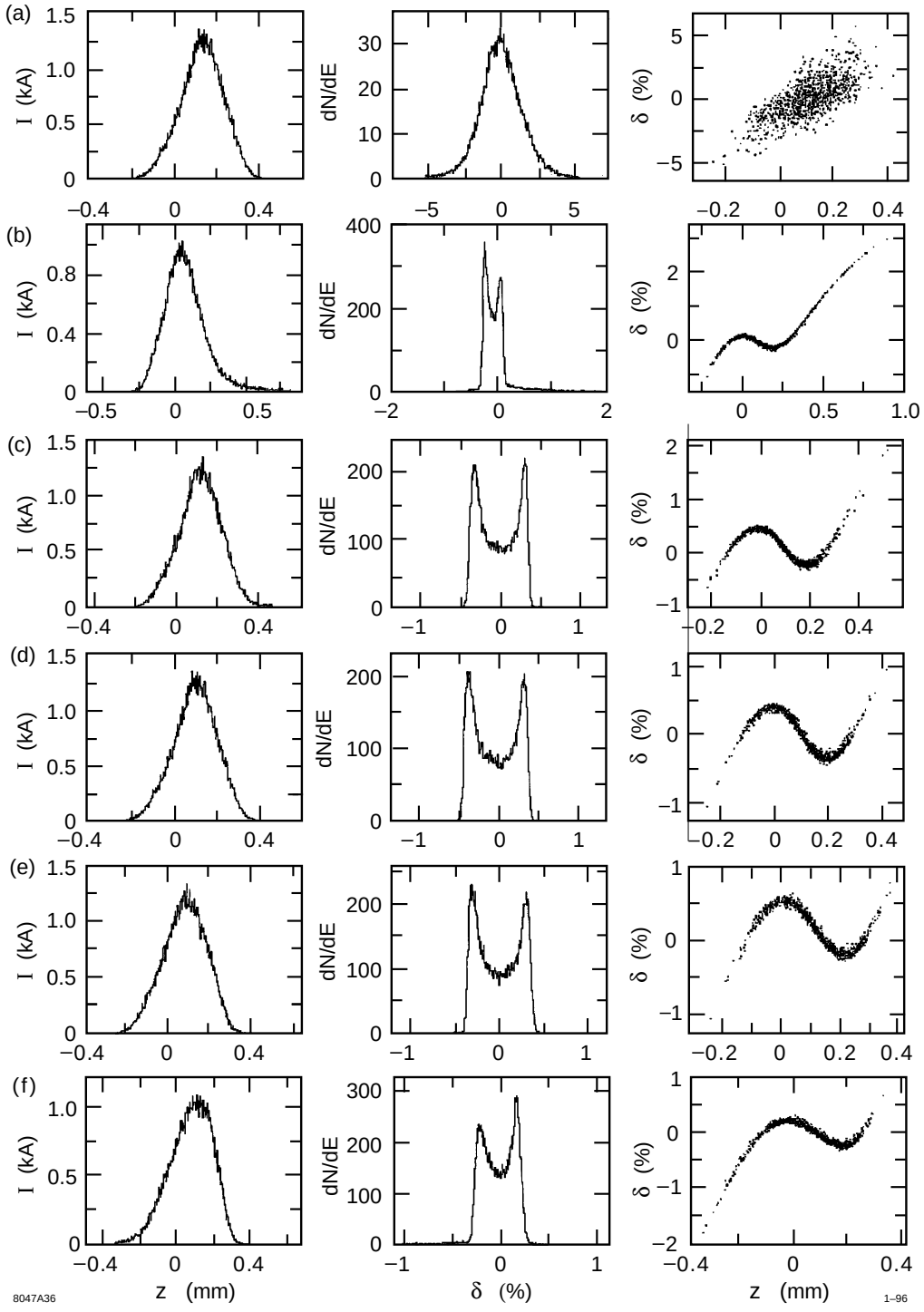


Figure 5-13. *Longitudinal phase space distributions for NLC-Ia, $N = 6.5 \times 10^9$ and a final bunch length of 100 μm: a) after the bunch compressor; and at the end of the main linac for an initial phase error of b) −12, c) −6, d) 0, e) 6, f) 12 mm.*

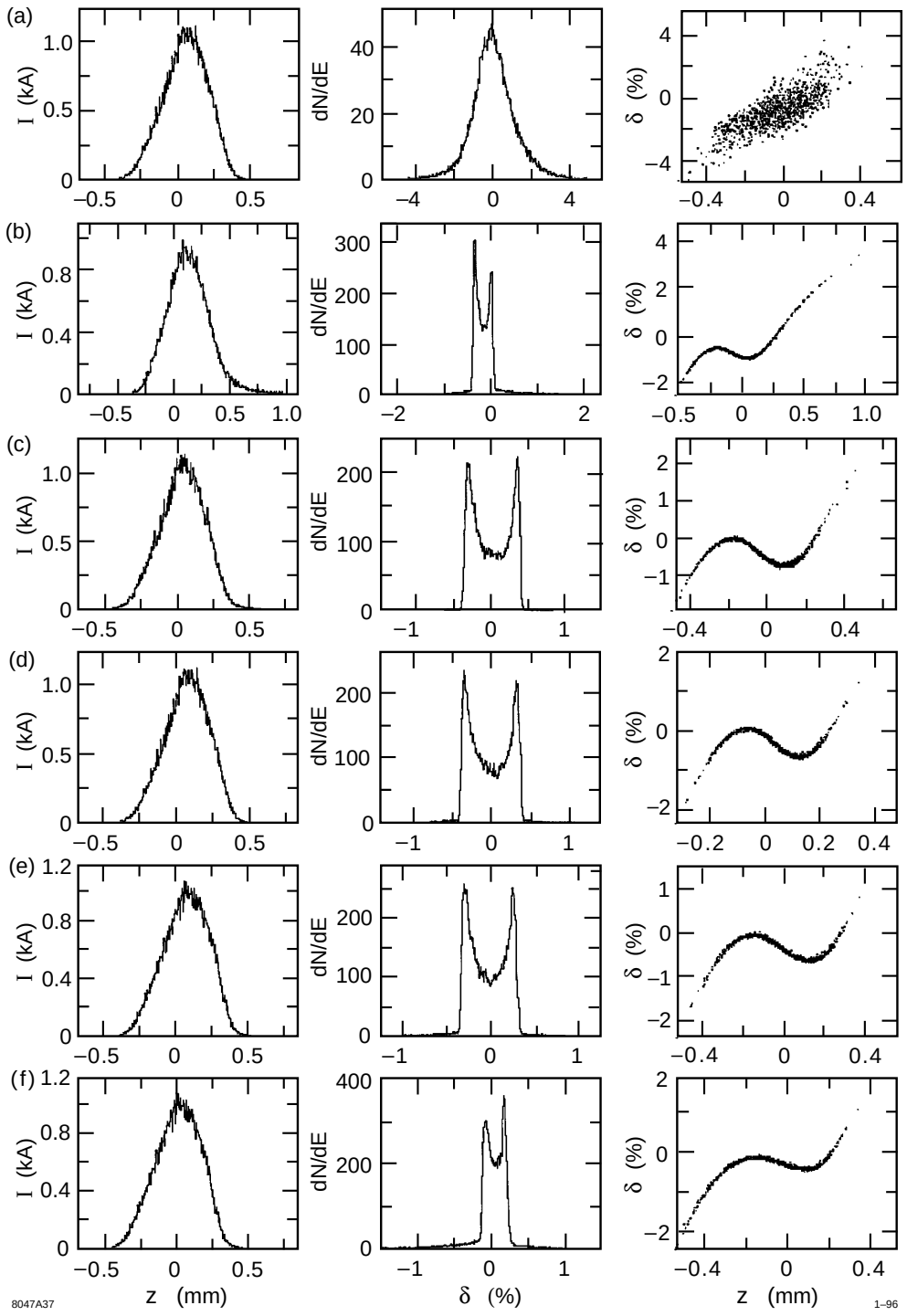


Figure 5-14. *Longitudinal phase space distributions for NLC-I c, $N = 8.5 \times 10^9$ and a final bunch length of 150 μm: a) after the bunch compressor; and at the end of main linac for an initial phase error of b) – 12, c) – 6, d) 0, e) 6, f) 12 mm.*

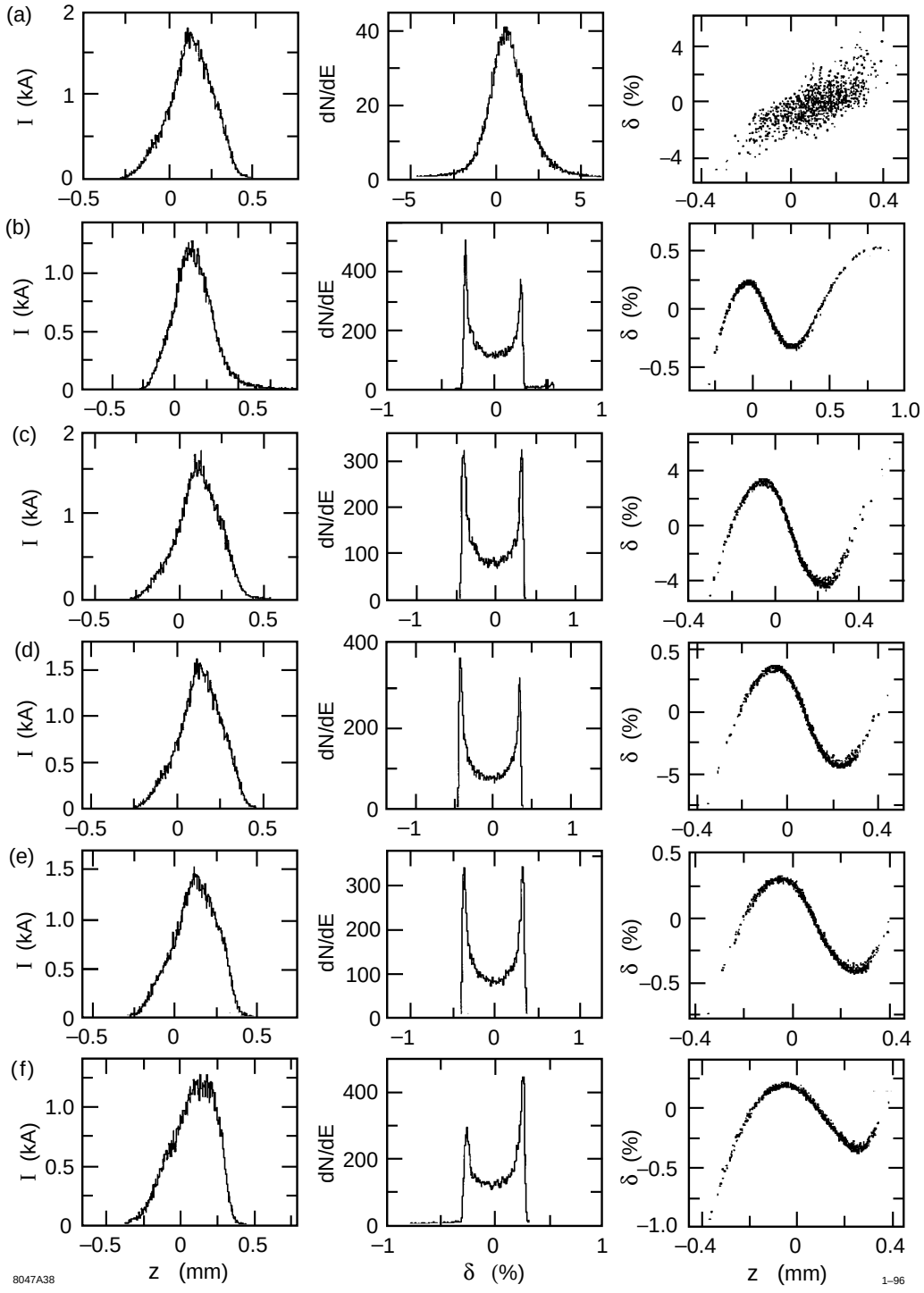


Figure 5-15. *Longitudinal phase space distributions for NLC-IIa, $N = 9.5 \times 10^9$ and a final bunch length of 125 μm: a) after the bunch compressor; and at the end of the main linac for an initial phase error of b) −12, c) −6, d) 0, e) 6, f) 12 mm.*

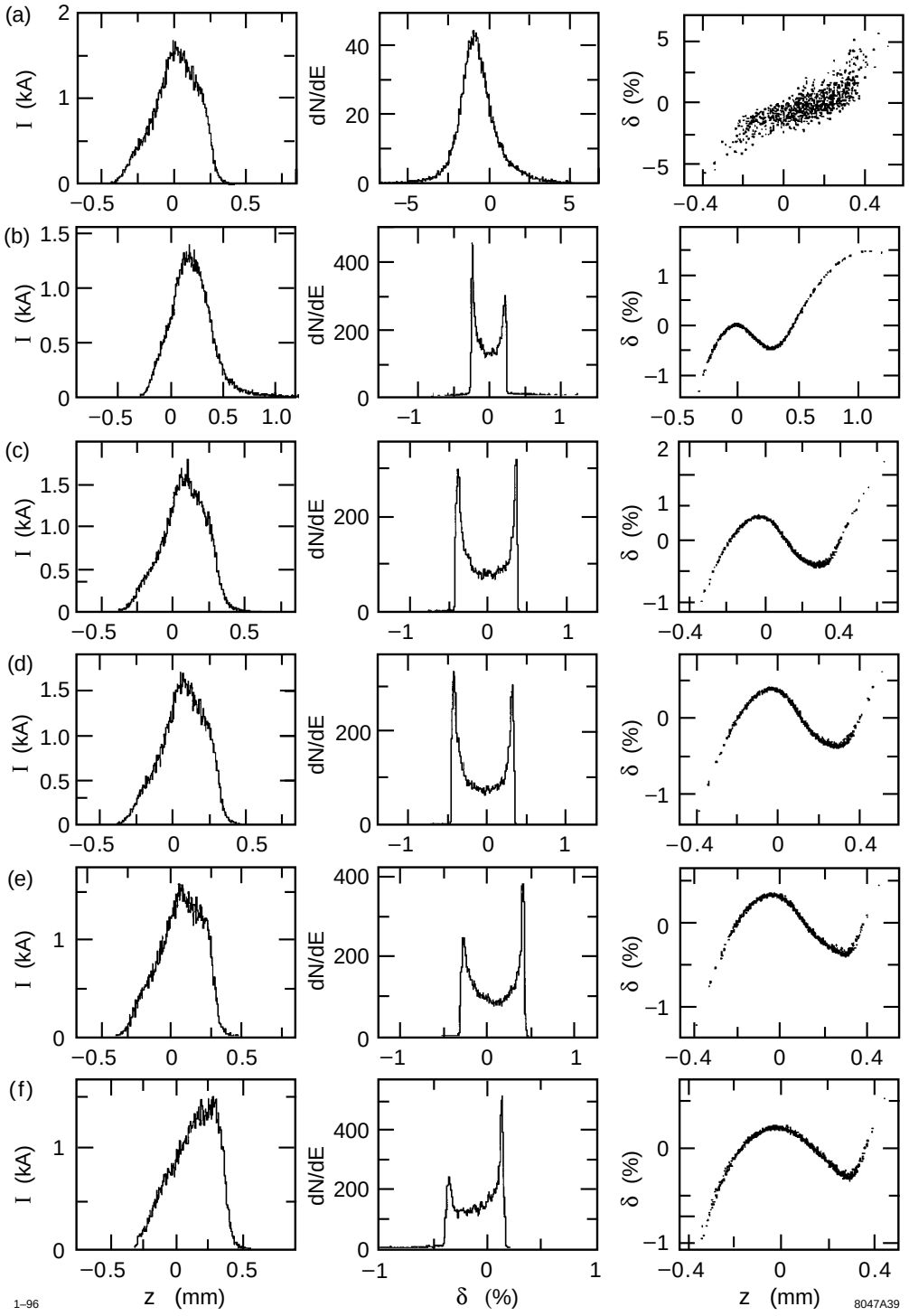


Figure 5-16. *Longitudinal phase space distributions for NLC-IIc, $N = 1.25 \times 10^{10}$ and a final bunch length of 150 μm: a) after the bunch compressor; and at the end of the the main linac for an initial phase error of b) −12, c) −6, d) 0, e) 6, f) 12 mm.*

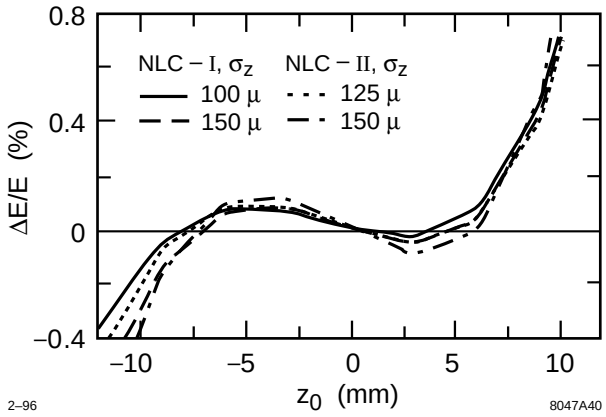


Figure 5-17. *Variation of average energy with initial phase error, for two different bunch lengths in the NLC-I and NLC-II.*

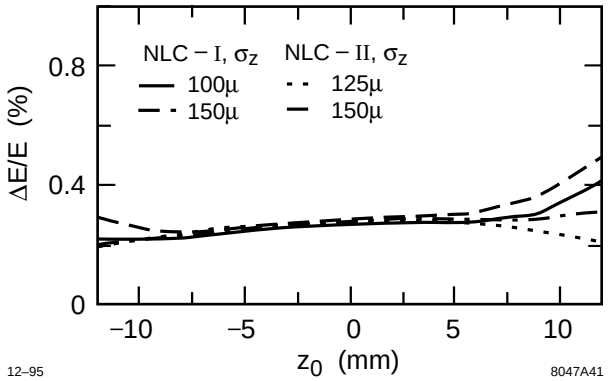


Figure 5-18. *Variation of rms energy spread with initial phase error, for two different bunch lengths in the NLC-I and NLC-II.*

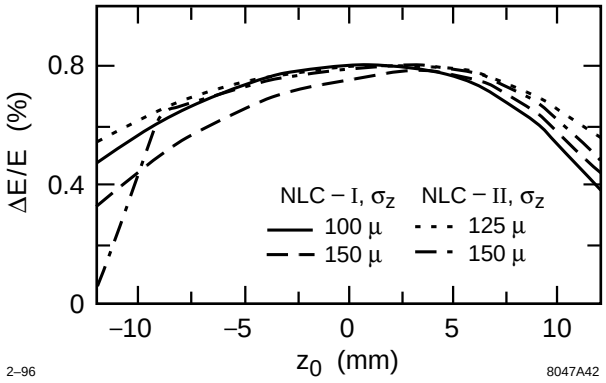


Figure 5-19. *Variation of FWHM energy spread with initial phase error, for two different bunch lengths in the NLC-I and NLC-II.*

Parameter	L-Band	S-Band	X-Band	Comment
f (GHz)	1.428	2.856	11.424	frequency
$2\,k$ (V/C/ m)	1.8×10^{13}	3.7×10^{13}	2.0×10^{14}	loss factor $\times$ 2
v_g/c	0.012	0.012	0.08	group velocity
Q	18, 000	13, 000	7000	quality factor
l (m)	6	3	1.8	structure length

Table 5-6. *Wakefield and structure parameters for the different accelerating sections in the bunch compressor.*

where c denotes the velocity of light. Also taken into account in the simulation is the group velocity of the fundamental mode, which introduces a small nonlinear component to the beam loading and makes perfect beam-loading compensation impossible—at least for the compensation scheme presently chosen. Wakefield parameters for the different structures were provided by R. Miller [Miller 1995]. They are compiled in Table 5-6.

The short-range wakefields used in the multibunch simulations are the same as those which were employed in the single-bunch case: The short-range wakefield for the X-band structure was derived by K. Bane [Bane 1995]. The wakefields for other frequencies were estimated from an approximative formula in Ref. [Palmer 1990].

The multi-bunch energy compensation scheme of choice was proposed by Kikuchi [Kikuchi 1992]. If two rf structures are driven at frequencies $f_0 \pm \Delta f$, the effective total voltage for the nth bunch is

$$\Delta V_n = V_c \sin\left(\frac{2\pi(f_0+\Delta f)z_n}{c}\right) - V_c \sin\left(\frac{2\pi(f_0-\Delta f)z_n}{c}\right) \tag{5.41}$$

$$\approx 4\pi V_c \frac{\Delta f}{c} \cos\left(\frac{2\pi f_0 z_n}{c}\right) \tag{5.42}$$

where z_n denotes the longitudinal position of the bunch, and $2\pi\ \Delta f z_n/c \ll 1$ was assumed. The required compensation voltage V_c is [Kikuchi 1992]

$$V_c = qc\frac{k_m L_m + 2k_c L_c}{b\ 2\pi\Delta f} \tag{5.43}$$

where $(k_m L_m)$ and $(2k_c L_c)$ are the beam-loading voltage in the main rf structure and compensating cavities, respectively, and b is the bunch spacing.

For all subsystems a compensation frequency in the S-band region was chosen, detuned from the main S-Band frequency by $\Delta f \approx \pm 1$ MHz. Several considerations determined the optimum choice of Δf. In general, a larger Δf reduces the required compensation voltage and the additional beam loading. The compensation, however, becomes less linear for larger beam frequency. It seems best to choose the frequency difference so as to partially cancel the nonlinear component of the beam loading. In the simulation, this is done empirically by minimizing the rms energy variation (or phase variation) as a function of Δf.

The initial setup of the multibunch compensating rf was performed by considering only one macroparticle per bunch in order to increase the computational speed. Subsequently, the main rf-voltages and phases were re-optimized for the single-bunch dynamics. This is necessary because of additional short-range wakefields in the compensation structures. Finally, a train of 90 bunches, 50 macroparticles each, was tracked through the entire system, and through the X-Band main linac. No long-range wakefields in the X-Band linac were considered, assuming that they are perfectly compensated by rf-pulse shaping.

The simulation includes an initial linear phase variation of ± 3 mm along the bunch train, as caused by beam loading in the damping rings. There is no need for a particular rf system to compensate this initial phase variation, since the compressor was designed to handle single-bunch phase errors up to ± 6 mm or larger.

Table 5-7 summarizes the compressor parameters used in the simulation. Note that the active length of the prelinac is about 320 m, and thus about 30% shorter than in the present ZDR design. The 320-m length assumes an unrealistic gradient of 30 MV/ m (or up to 36 MV/ m for the compensating rf) but the performance is not expected to be much different for a greater length.

Results of the simulation study are summarized in Figures 5-20 and 5-21 and in Table 5-8. Note that the bunch-to-bunch variation of energy, position, and bunch length, apparent in the two figures, arises from the finite number of macroparticles per bunch (50), and is much larger than the expected actual bunch-to-bunch variation. On the other hand, the slow change of energy, position, and bunch length over several bunches represents the effect of the long-range wakefields, and is the quantity of interest here. Table 5-8 lists the longitudinal bunch-to-bunch phase variation, the bunch-to-bunch energy spread, the intrabunch energy spread, and the rms bunch length at the end of the main linac, for different NLC scenarios. The longitudinal bunch position varies by 30–40 μm about the average value. The bunch length fluctuates by ± 10 μm. Changes of phase and bunch length give rise to a bunch-to-bunch rms energy variation of roughly 0.1–0.2%, and a total variation of 0.8% for the NLC-I and 0.6% for the NLC-II. The rms intrabunch energy spread is 0.3–0.4% in all cases.

We conclude that the ΔF multibunch energy compensation is straightforward and, in the simulation, its performance is adequate and satisfactory. The primary disadvantage of the ΔF technique is that the compensation is not local and thus the dispersive and chromatic emittance dilutions are larger. In the NLC injectors, we have adopted a combination of the ΔT and ΔF techniques that should perform far better than the ΔF scheme reported here–this is described in Chapter 6.

Parameter	Collider Version NLC-Ia	NLC-Ic	NLC-IIa	NLC-IIc	Comment
V_{LB} (MV) ϕ_{LB} (°) L (m)	136.0 −89.88	136.5 −89.88	136.0 −89.92 8.5	136.0 −89.88	Main L-band RF 1st stage $f = 1.428$ GHz
V_{SB} (MV) ϕ_{SB} (°) L (m)			6.9 −180.0 0.4		Comp. S-band RF 1st Stage $f = 2.856$ GHz
V_{SB} (MV) Δf (kHz) ϕ_{SB} (°) L (m)	±26.7	±34.4	±38.4 ± 856.8 −90. 2×3.0	±50.5	Multibunch Comp. S-band RF 1st Stage $f_c = f_0 \pm \Delta f$, $f_0 = 2.856$ GHz
V_{SB} (MV) ϕ_{SB} (°) L (m)	8050 −4.0	8049 −3.0	8054 −3.0 270	8068 −3.0	Prelinac S-Band RF $f = 2.856$ GHz
V_{SB} (MV) Δf (kHz) ϕ_{SB} (°) L (m)	±805	±1037	±1159 ± 856.8 −90. 2×27 / 2×39	±1525	Multibunch Comp. S-band RF Prelinac $f_c = f_0 \pm \Delta f$, $f_0 = 2.856$ GHz
V_{SB} (MV) ϕ_{SB} (°) L (m)	3850 −89.6	3500 −88.5	3730 −89.7 130	3610 −89.7	Main S-band RF 2nd Stage $f = 2.856$ GHz
V_{XB} (MV) ϕ_{XB} (°) L (m)	274	276	256 −183.0 8	240	Comp. X-band RF 2nd Stage $f = 11.424$ GHz
V_{SB} (MV) Δf (kHz) ϕ_{SB} (°) L (m)	±338	±440	±492 ± 1142.4 −90. 2×15.0	±647	Multibunch Comp. S-band RF 2nd stage $f_c = f_0 \pm \Delta f$, $f_0 = 2.856$ GHz
V_{XB} (GV) ϕ_{XB} (°) L (m)	270.32 −15.7 8130	236.83 −16.1 8130	535.69 −8.3 8900	480.33 −12.0 8900	Main Linac X-band RF $f = 11.424$ GHz

Table 5-7. *Rf parameters of compressor subsystems, as used in the multibunch simulation study.*

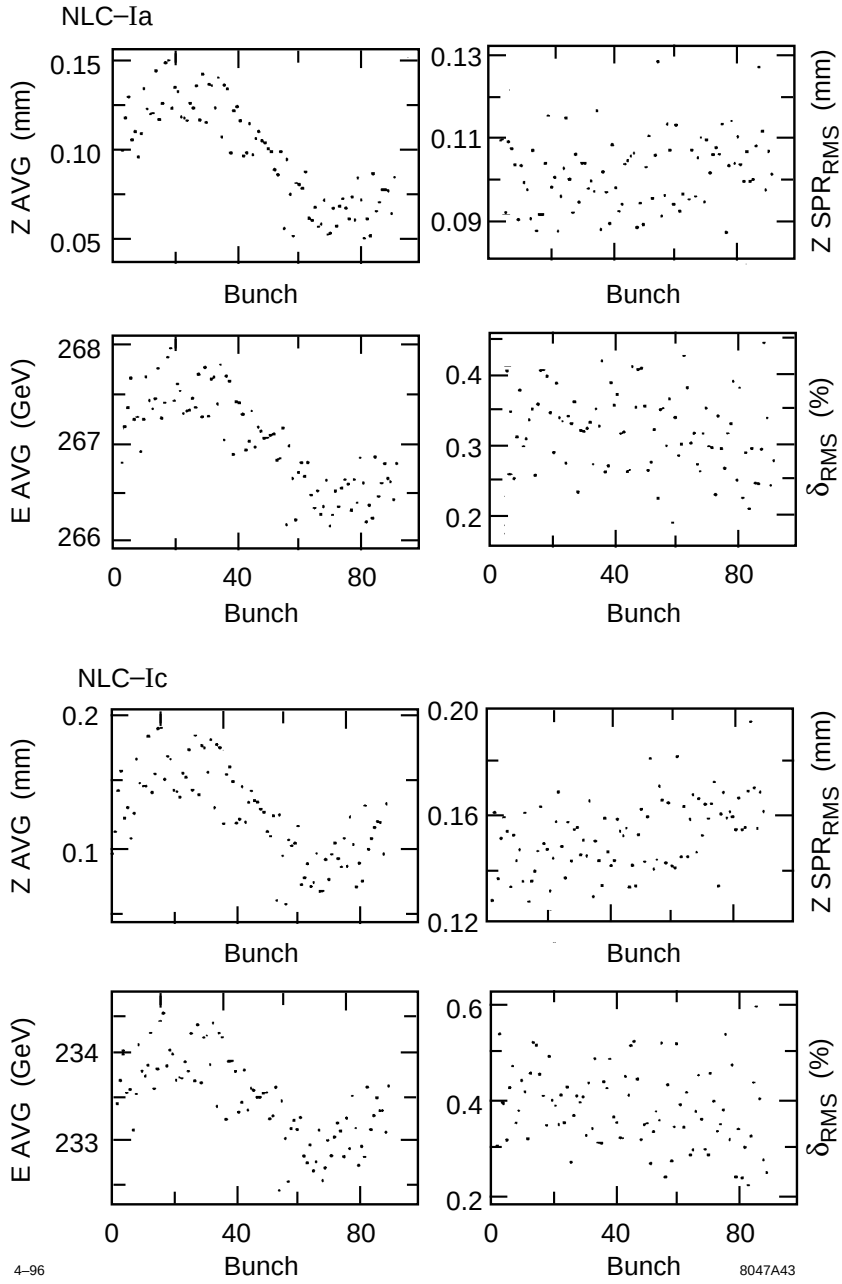


Figure 5-20. *Bunch centroid position, rms bunch length, average bunch energy, and rms energy spread, as a function of bunch number in a train of 90 bunches, for two different versions of NLC-I.*

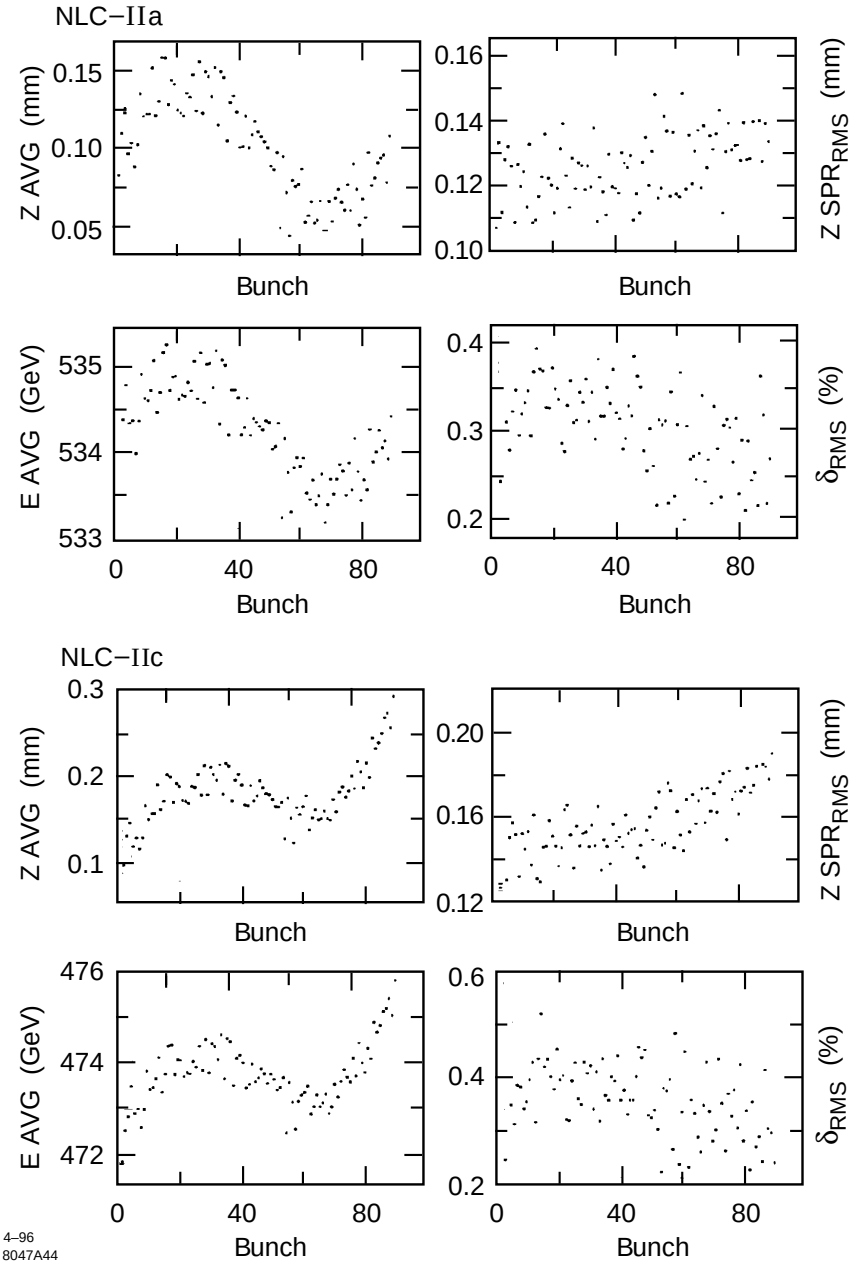


Figure 5-21. *Bunch centroid position, rms bunch length, average bunch energy, and rms energy spread, as a function of bunch number in a train of 90 bunches, for two different versions of NLC-II.*

Parameter	Collider Version				Comment
	NLC-Ia	NLC-Ic	NLC-IIa	NLC-IIc	
$z_{b,\ rms}$ (μm)	28	33	31	36	Rms phase variation
$\sigma_{z,\ ave}$ (μm)	100 ± 9	153 ± 13	125 ± 10	156 ± 15	Average bunch length
$\delta_{b,\ rms}$ (%)	0.18	0.20	0.09	0.14	Rms energy variation
$\sigma_{\delta,\ ave}$ (%)	0.31	0.38	0.31	0.36	Intrabunch rms energy spread
E_b (GeV)	266.9	233.46	534.2	473.7	Average energy

Table 5-8. *Inter- and intrabunch energy spread, rms bunch length, and longitudinal phase variation at the end of the X-band main linac, for a train of 90 bunches in two different versions of the NLC-I and NLC-II. Numbers were obtained by a macroparticle simulation, which includes the effect of short- and long-range wakefields, multibunch energy compensation, and an initial phase variation of 6 mm along the train.*

5.5 Transverse Dynamics

In this section, we will examine the effects that can dilute the transverse emittances. The primary source of dilution are optical aberrations in the spin rotator and compressors and single bunch wakefields in the linacs. We will describe these effects and the respective tuning compensation systems in three regions which are organized according to the respective diagnostic sections. Thus, we will describe the tuning in the spin rotator and low-energy bunch compressor, in the prelinac, and in the high-energy bunch compressor. Next, we will discuss the effect of the multibunch transverse wakefields and then finally we will describe the effects of: space charge fields and coherent radiation, ion trapping, and synchrotron radiation. All of these later effects impose weak constraints on the system design.

5.5.1 Spin Rotator and Low-Energy Bunch Compressor

Given the small vertical emittance at damping ring extraction ($\gamma\epsilon_y = 3 \times 10^{-8}$ m), it is crucial that tuning elements and algorithms be provided which will allow full correction of the inevitable fabrication and alignment errors of the first bunch compressor (BC1). This is especially necessary when considering the large energy spread incurred after the BC1 rf structures ($\sim$1% rms) as well as the large cross-plane coupling generated and canceled again between spin rotator solenoids. The following section describes dispersion correction elements, coupling correctors, beam matching, and diagnostics necessary for preserving the damped transverse emittance. In addition, the results from a full tuning simulation are presented.

Tuning Elements and Aberrations

Dispersion Correction Elements. The large energy spread of the bunch compressor sets very tight tolerances on residual dispersion, especially in the vertical plane. The relative emittance dilution due to residual spatial, η, and angular, $\eta\prime$, dispersion is expressed in Eq. 5.44 including subsequent chromatic filamentation.

$$\frac{\delta\epsilon}{\epsilon_0} = \frac{\eta^2 + (\eta\alpha + \eta\prime\beta)^2}{2\epsilon_0\beta}\sigma_\delta^2 \quad . \tag{5.44}$$

The tolerable vertical spatial dispersion at a point $\beta = 10\,\mathrm{m}, \alpha = 0$, is 175 μm (2% emittance dilution at $\gamma\epsilon_y = 3 \times 10^{-8}$ m and 1% rms energy spread). To correct at this level, four small skew quadrupoles are placed in the wiggler section of the BC1 at points of large horizontal dispersion. These are tiny correction elements (10-cm long, 30-mm pole-tip radius, $\pm$1 kGauss pole-tip field). They are paired at $-I$ (transfer matrix) intervals so that coupling is not generated and the pairs are separated by 90°. In this way, orthogonal η and $\eta\prime$ knobs are available. A similar system of four normal quadrupoles is used for horizontal dispersion correction. Figure 5-22 shows the placement of these skew and normal elements (SQW1,...4, CQW1,...4), as well as other elements discussed below.

In addition to dispersion correction after the BC1 rf structures, it is useful to include two more skew quadrupoles (SQR1, SQR2) in the mini-arc section of the spin rotator to provide correction of residual vertical dispersion generated before the rf sections. The net emittance dilution is a linear combination of the two residual dispersion functions weighted by their respective energy spreads (1% post-compression and 0.1% pre-compression). Although post-compression dispersion is much more intolerable, any significant pre-compression dispersion will not be correctable by the wiggler section skew quadrupoles. The two skew elements in the mini-arc are separated by 90° to provide spatial and angular pre-compression vertical dispersion correction. The ratio $(\eta_x\sigma_\delta)^2/\beta_x\epsilon_x >> 1$ so that paired coupling cancelation is not necessary.

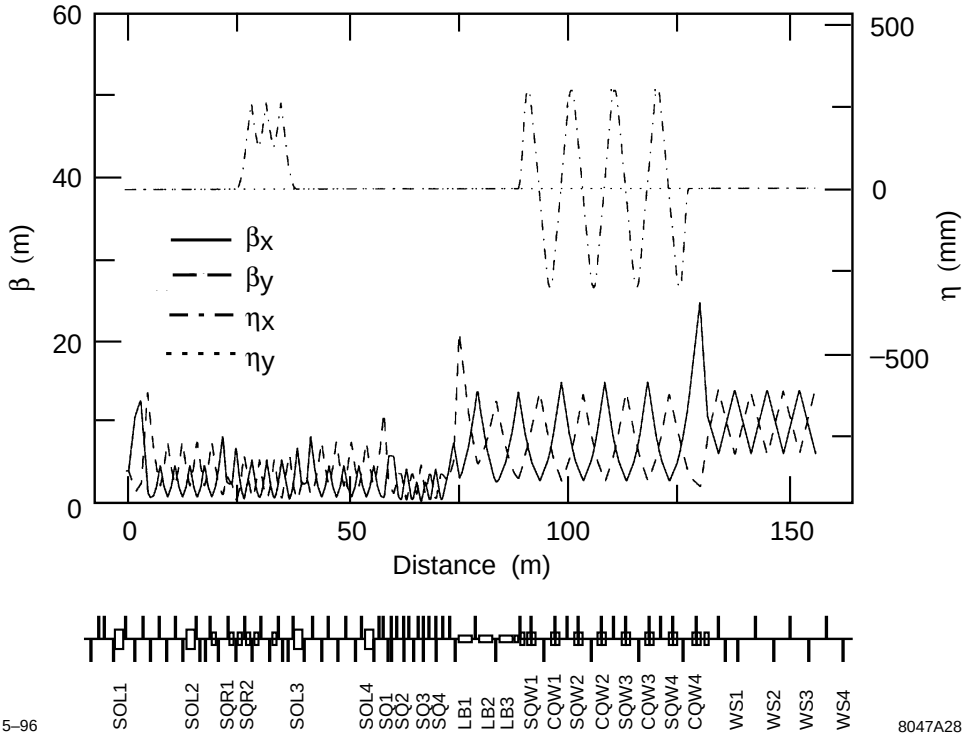


Figure 5-22. *Dispersion and beta functions for the entire low-energy bunch compressor and spin rotator system. Tuning elements and major component locations are indicated.*

Coupling Correction. The spin rotator solenoids (SOL1,...4) are nominally separated by optics which will cancel their cross-plane coupling. However, imperfections in the cancelation optics will result in emittance dilution. Furthermore it may be necessary to correct coupling generated within the damping rings. Therefore, a skew correction section is included following the spin rotator. This section contains four skew quadrupoles (SQ1,...4) in a non-dispersive section which is placed at ideal phase advance points so that they are initially orthonormal. The vertical emittance is minimized with each of these four skew elements. At the end of this process, and allowing for some reiteration, the coupling is completely corrected.

Beam Matching. Several provisions exist in the BC1 to compensate for matching errors. There are four variable strength quadrupoles which match the extracted beam to the spin rotator optics. There are also matching quadrupoles within the spin rotator to compensate for the focusing effect of the solenoids. Finally there are several matching quads between the skew correction section and the wiggler and also after the wiggler section. The BC1 rf may be switched off in order to distinguish more clearly between dispersion errors and matching errors.

Second-Order Corrections. Although it is not necessary to include local sextupole chromatic corrections for the design optics, simulations have shown (Section 5.5.1, Tuning Simulations) that for large vertical dispersion corrections $(\epsilon_y/\epsilon_{y0} >> 10)$ some second-order dispersion arises. If this situation develops, it is a simple matter to include four skew sextupoles adjacent to the wiggler skew quadrupoles to provide T_{366} and T_{466} knobs which are paired so that vertical chromaticity is not generated.

RMS Quadrupole Strength Errors ($\Delta k/k$)	1.0%
RMS Dipole Magnet Roll Errors (θ)	0.25°
Initial Cross-Plane Coupling (DR output)	$\epsilon_y/\epsilon_{y0} = 2.2$
All Four Solenoids at Full Field	35 kGauss

Table 5-9. *Errors and conditions for bunch compressor and spin rotator tuning simulations.*

Phase Space Diagnostics. To measure the BC1 extracted beam, four wire scanners (WS1,...4) are placed at ideal phase advance intervals (45°) at the end of the wiggler. The corrections described above are applied by minimizing the measured emittance. It is not necessary to actually measure the cross-plane coupling. It is fully corrected by minimizing the projected vertical emittance. The matched beam size is the same at each wire ($\sim$10 μm vertically by $\sim$65 μm horizontally) and therefore the matched condition is easily recognizable. Wire scanner roll tolerances are readily achievable ($\sim 1°$). In addition, a wire scanner placed at a high-dispersion point ($\sim$220 mm) in the spin rotator mini-arc will be useful to measure the extracted damping ring energy spread ($\sim$0.1%). The ratio of the dispersive beam size to the betatron beam size is $\sim$4:1 here. Finally a wire scanner at a peak dispersion point ($\sim$240 mm) in the wiggler will be useful to measure the damping ring extracted bunch length (similar size ratio $\sim$4:1).

Tuning Simulations

Tuning simulations have been run for the many corrections described above using the *Final Focus Flight Simulator*. Table 5-9 summarizes the set of errors and conditions introduced to the bunch compressor and spin rotator system for these simulations.

These errors produce extreme conditions that result in a vertical emittance which is a factor of $>$200 times larger than the initial intrinsic emittance $\gamma\epsilon_{y0} = 2.3 \times 10^{-8}$ m. In this case, the vertical beam profiles measured on the wire scanners will increase from the nominal $\sim$10 μm to $>$130 μm.

The simulations are run to test the tuning algorithms, especially in this extreme case. In the tuning process, a realistic measurement resolution is not attempted. The emittance measurements are infinitely precise and the initial diluted vertical emittance is $\gamma\epsilon_{y0} = 540 \times 10^{-8}$ m. Although it may not be the optimal procedure, the following describes the tuning procedure used:

Step 1. With L-band rf switched on, scan both vertical dispersion knobs to minimize projected vertical emittance (some iteration useful at this extreme stage).

Step 2. With L-band rf switched on, scan both horizontal dispersion knobs to minimize projected horizontal emittance.

Step 3. With L-band rf switched off, scan the four coupling correction skew quadrupoles to minimize projected vertical emittance.

Step 4. With L-band rf switched off, match horizontal and vertical beta functions with damping ring extraction quadrupoles (not actually tested here).

Step 5. Repeat Steps 1 and 2 (most dilution caused by dispersion).

Step 6. With L-band rf switched on, scan both second-order vertical dispersion knobs to minimize projected vertical emittance (only necessary for these extreme conditions).

Step 7. With L-band rf switched off, scan pre-compression vertical dispersion knobs.

The simulation of this procedure with the conditions of Table 5-9 has proven very successful and results in a final emittance dilution of $<$2% in both planes. Quadrupole alignment errors were not included in the simulation; however,

this should simply add to vertical dispersion errors which are already extreme due to the large dipole magnet roll errors. It is estimated that such a procedure will probably require approximately 16 hours of real machine time.

5.5.2 Prelinac

The prelinac quadrupoles will be aligned using the standard quadrupole beam-based alignment procedure used in the main linacs and the beam-delivery sections. Each quadrupole will be powered with an individual power supply. Then, the power supply can be varied making it straightforward to determine the alignment of the quadrupole magnetic center with respect to the electrical center of the BPM captured in the magnet. Each quadrupole will be mounted on a remote magnet mover to facilitate the alignment. In addition, the accelerator structures will also be mounted on remote movers and will contain BPMs that will measure the induce dipole modes in the structures. This will allow the structure to be accurately aligned to the beam trajectory established by the quadrupoles. Finally, additional tuning, using oscillation bumps similar to those used in the SLC linac, could be performed using the emittance diagnostic station located at the end of the prelinac.

Simulations of the transverse emittance dilution were performed with the LInear Accelerator Research simulation program described in Chapter 7. In simulations, the residual vertical emittance dilution was less than 15% assuming errors due to 10 μm quad-BPM resolution, 40 μm structure BPM resolution, 200 μm initial quadrupole offsets, and 1 μm mover step size. The dilution is dominated by the dispersive dilution due to the large incoming energy spread (1%) of the beam. These are very loose tolerances. The expected errors are 2 μm quad-BPM misalignments, 20 μm structure BPM resolution, and 0.3 μm mover step size. These would yield a vertical emittance dilution that is less than 1%.

5.5.3 High-Energy Bunch Compressor

The high-energy bunch compressor can be separated into 3 stages: the 180° arc, the compressor rf, and the chicane. The quadrupoles in the arc and the compressor rf section will be aligned using the standard beam-based techniques that will also be used in the pre-linac, the main linac, and the beam-delivery section of the collider. Here, each quadrupole is powered with an individual power supply and the power supply is varied to determine the magnetic center of the quadrupole with respect to a BPM captured within the bore. The quadrupoles are then aligned with remote magnet movers; this algorithm is described in greater detail in Chapter 7. and the preceding section on the prelinac alignment. The arc was explicitly designed using separated function magnets to facilitate this form of alignment.

The 4-GeV compressor rf section has strong focusing to reduce the sensitivity to wakefields. To keep the vertical emittance dilution due to transverse wakefields to less than 1%, the accelerator structures must be aligned within 45 μm of the beam trajectory. This will be accomplished using structure BPMs ("S"-BPMs) to measure the induced dipole mode and the remote movers to align the structures to the beam path. Given this loose tolerance, no significant dilution is expected.

As in the low-energy bunch compressor, tuning elements must also be included in the high-energy compressor. The most likely mechanisms for emittance dilution are rolled dipoles and quadrupoles, misaligned quadrupoles and sextupoles, as well as quadrupole gradient errors, all of which generate mostly anomalous dispersion which phase mixes in the main linac (see Eq. 5.45). Cross-plane coupling and beta-mismatch errors are much less likely since the horizontal beam size in the arc and chicane is dominated by dispersion (*i.e.*, $\{\eta_x\sigma_\delta\}^2/\beta_x\epsilon_x \gg 1$) (see Table 5-4 and Figures 5-11 and 5-12). The next sections describe the tuning elements and algorithms for controlling anomalous dispersion in the arc and chicane; the tuning of the rf section was described earlier.

Name	Quantity	Roll (mrad)	$\Delta B/B_0$ (10^{-4})	b_1/b_0 (%)	b_2/b_0 (%)
BB	138	4.7	1.2	0.3	5

Table 5-10. *BC2 arc dipole magnet single element tolerances for 8% emittance growth each ($\gamma\varepsilon_{x0} = 3\times10^{-6}$ m, $\gamma\varepsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.25\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 4 mm. They generate 1st and 2nd order x-dispersion.*

Name	Quantity	Roll (mrad)	Δx offset (μm)	Δy offset (μm)	$\Delta x_{\rm rms}$ vibrate (μm)	$\Delta y_{\rm rms}$ vibrate (μm)	$\Delta B/B_0$ (%)	b_2/b_1 (%)
QF	69	1.12	920	230	2.3	0.58	0.9	10
QD	68	1.10	2930	100	7.3	0.25	4.8	120

Table 5-11. *BC2 arc quadrupole magnet single element tolerances for 8% emittance growth each ($\gamma\varepsilon_{x0} = 3\times10^{-6}$ m, $\gamma\varepsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.25\%$). Sextupole field component tolerances (b_2/b_1) are evaluated at a radius of 4 mm and are very loose for the QDs (120%).*

Name	Quantity	Roll (mrad)	Δx offset (μm)	Δy offset (μm)	$\Delta x_{\rm rms}$ vibrate (μm)	$\Delta y_{\rm rms}$ vibrate (μm)	$\Delta B/B_0$ (%)
SF	68	325	800	160	800	160	>50
SD	68	325	1740	86	1740	86	>50

Table 5-12. *BC2 arc sextupole magnet single element tolerances for 8% emittance growth each ($\gamma\varepsilon_{x0} = 3\times10^{-6}$ m, $\gamma\varepsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.25\%$). The field regulation and roll tolerances are not important (for aligned sextupoles) because the chromaticity of the arc is insignificant.*

180° Arc Section

In order to estimate the dispersion correction range needed we first examine the single element tolerances for the arc. These tolerances are then scaled to estimate the emittance dilution associated with an achievable set of alignment and quadrupole gradient tolerances. This estimated emittance dilution is used to choose the necessary correction range and the scheme is tested for induced non-linear aberrations. The single element tolerances for the 180° arc FODO magnets are listed below in Table 5-10 (dipoles), 5-11 (quadrupoles), and 5-12 (sextupoles). Each tolerance represents a 2% luminosity loss for that single element's effect on one beam (8% emittance increase for one beam in one plane). The effects of these errors increase the IP beam size except for dipole field regulation and quadrupole transverse vibration which steer the beams out of collision. In this case, the betatron phase to the IP is not calculated so phase averaging is applied.

The important vertical emittance dilution mechanisms include vertical misalignments of the quadrupoles and sextupoles and roll errors of quadrupoles and dipoles. Horizontal dilution mechanisms include horizontal misalignments and gradient errors of the QF quadrupoles. The magnetic multipole tolerances in the tables, when scaled down by $\sqrt{N}$, are easily achievable, as are the horizontal alignment and gradient errors of the QD quadrupoles and the roll and field errors of the sextupoles. Static field strength errors of the dipoles can be encompassed in the horizontal alignment of

Rms magnet roll angles	Rms vertical misalignments	Rms horizontal misalignments	Rms static gradient errors	$\Delta\varepsilon_y/\varepsilon_{y0}$	$\Delta\varepsilon_x/\varepsilon_{x0}$
0.5 mrad QF/QD/BB	100 μm QF/QD 60 μm SF/SD	200 μm QF/QD 200 μm SF/SD	0.5% QF/QD	—	—
68 QDs				1.1	$\sim$0
69 QFs				1.1	$\sim$0
138 BBs				0.1	$\sim$0
	68 QDs			5.4	0
	69 QFs			1.0	0
	68 SDs			2.6	$\sim$0
	68 SFs			0.8	$\sim$0
		68 SDs		0	0.1
		68 SFs		0	0.3
		68 QDs		0	$\sim$0
		68 QFs		0	0.2
			68 QDs	$\sim$0	0.1
			69 QFs	$\sim$0	1.7
			Total Expected Dilution *without* Tuning =	12.1	2.4

Table 5-13. *Chosen set of rms alignment tolerances and quadrupole static gradient tolerances for the BC2 arc and their impact on emittance dilution for each plane ($\gamma\varepsilon_{x0} = 3\times10^{-6}$ m, $\gamma\varepsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.25\%$). Vertical alignment of SDs and QDs is most critical.*

the quadrupoles. Regulation and vibration tolerances are not addressed by the tuning schemes discussed here. These non-static problems will need to be corrected by orbit feedback systems or, as in the case of dipole field regulation, all dipoles or groups of dipoles will be powered in series to significantly loosen these tolerances. An achievable set of alignment and quadrupole gradient tolerances is now chosen and the associated emittance dilution is tabulated for each mechanism in Table 5-13.

For these tolerances, the total emittance dilution factor for the vertical plane may be $\sim$12, whereas the horizontal may be $\sim$2. Dispersion correctors should handle at least this range and should also not affect beta functions or introduce betatron coupling. The quadrupole roll tolerances of 0.5 mr and sextupole vertical alignment tolerances of 60 μm were chosen so that cross-plane coupling will not be a problem and therefore no betatron-coupling correction is required (*i.e.*, for a monochromatic beam, 69 QF and 68 QD magnets with 0.5 mr rms roll errors plus 68 SF and 68 SD magnets with 60 μm rms vertical misalignments will produce a 2% luminosity loss for the one beam due to betatron coupling). The sextupoles may easily be beam-based aligned if vertical movers are included with ±1 mm range. If an SD's pole-tip field is temporarily raised to -4 kGauss (from -3 kGauss) and its vertical position is scanned in $\sim$5 steps over the range ±1 mm while the horizontal beam centroid offset is recorded on a $<$4-μm resolution BPM near the next QF, then each sextupole may be aligned to $<$60-μm resolution. This resolution may be improved by using more than one BPM. Note that energy jitter will vary the BPMs reading due to the $\sim$100-mm x-dispersion there. However, all arc x-BPMs may be used together to correct the data for this energy jitter down to $\delta E/E_0 = (4\,\mu\text{m}/\sqrt{68})/(100\,\text{mm}) \sim 5\times10^{-6}$. Figure 5-23 shows a simulated sextupole (SD) alignment scan using one 4-μm resolution x-BPM.

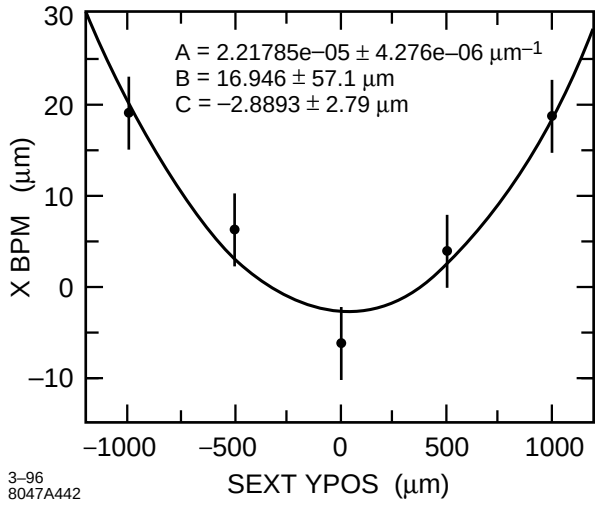


Figure 5-23. *Simulated sextupole (SD) alignment scan using one 4-μm-resolution x-BPM located at the next QF. The SD's pole-tip field is set to -4 kGauss (from -3 kGauss) for the alignment.*

Dispersion Correctors for the Arc. Vertical dispersion correction at the end of the arc can be provided by adding four small skew quadrupoles (of zero nominal field)—one per cell in the last four cells (cells 65–68). This scheme takes advantage of the 90° vertical phase advance per cell by pairing skew quadrupoles at $-I$ transfer matrix (2 cell) separation so that, for equal and opposite skew quadrupole settings, no betatron cross-plane coupling is generated. The two asymmetrically powered pairs of skew quadrupoles are shifted by one cell ($\Delta\psi_y = 90°$) with respect to each other so that both betatron phases are correctable (η_y and η'_y). Betatron coupling is also correctable, if necessary, if one skew quadrupole is also added to each of cells 61–64. In this way four symmetric pairs may be formed (61/63, 62/64, 65/67 and 66/68) which correct betatron coupling without introducing vertical dispersion so that, given the unequal x and y tunes of the arc, all four cross plane correlations are removable within skew quadrupole field limits (not possible for $\Delta\psi_x = 90°$/cell). The range of vertical dispersion correction for one pair of 10-cm-long, ± 1 kGauss pole-tip field, 30-mm pole-tip radius skew quadrupoles located 10-cm upstream of the QD at $\beta_y = 6.5$ m, $\eta_x = 50$ mm is $\Delta\varepsilon_y/\varepsilon_{y0} \sim 12$ at 0.25% rms energy spread for $\gamma\varepsilon_{y0} = 3 \times 10^{-8}$ m (or the peak vertical dispersion induced at the center of the next QD magnet is as much as ± 7 mm per pair). No significant coupling or beta beat is generated over this range. For betatron coupling, one of the four symmetric pairs of skew quadrupoles at $\pm$2-kGauss pole-tip field will correct a 50% vertical emittance dilution due to coupling. These four coupling correctors may be easily added if betatron coupling becomes a problem. Their inclusion will effectively loosen sextupole vertical alignment tolerances by nearly a factor of 2. Of course the skew quadrupoles may also be replaced, with some loss in correction speed, by QD roll-movers with $\pm$10-mr range or SD vertical movers of $\pm$1-mm range (SD vertical mover steering in the x-plane, is small and can be removed with a post-arc launch feedback system); the magnet movers planned exceed these ranges.

The horizontal dispersion may similarly be controlled by adding two pairs of small normal quadrupoles (of zero nominal field). Due to the nominal 108° horizontal phase advance per cell, these quad pairs must be spaced by 5 cells in order to provide a $-I$ separation. It is probably best not to nest the skew and normal correction quads. If cell 54 and 55 as well as cell 59 and 60 include a 10-cm-long, $\pm$2.3-kGauss, 30-mm radius quadrupole which is 10-cm upstream of the QF at $\beta_x = 6.4$ m, $\eta_x = 93$ mm, the emittance correction range per pair will be $\Delta\varepsilon_x/\varepsilon_{x0} = 2.4$ (or the additional horizontal dispersion induced at the center of a QF magnet is as much as ± 28 mm per pair). If cells 56, 57 and 58 also include a 10-cm-long normal quadrupole, then the horizontal dispersion correction scheme will also work for 90°/cell and 135°/cell horizontal phase advance configurations if the appropriate correction quadrupole (CQ) pairs are tuned for a given R_{56} configuration (*i.e.*, $\psi_x = 90°$/cell: CQ pairs 57/59 and 58/60, $\psi_x = 108°$/cell: CQ

Quad type	Quantity	Length (m)	Pole radius (mm)	Max. field (kGauss)	2% lum. step size (kGauss)	$\Delta\eta_{x,y}^{\max}$ at $\beta_{x,y}^{\max}$ (mm)	Max. $\varepsilon_y/\varepsilon_{y0}$
Skew	4–8	0.1	30	$\pm$1.0	$\pm$0.08	7	12
Normal	7	0.1	30	$\pm$2.3	$\pm$0.42	28	2.4

Table 5-14. *BC2 arc dispersion correction magnet specifications for 0.25% rms energy spread, $\gamma\varepsilon_{x0} = 3 \times 10^{-6}$ m and $\gamma\varepsilon_{y0} = 3 \times 10^{-8}$ m. There is one skew quadrupole (or SD vertical mover) in each of the last four (or eight) FODO cells 10-cm upstream of each QD and one normal quadrupole in each of cells 54 through 60, 10-cm upstream of each QF (or QF trim supplies).*

pairs 54/59 and 55/60, $\psi_x = 135°$/cell: CQ pairs 55/59 and 56/60; see Table 5-4). Since the vertical transfer matrix between paired normal quadrupoles (5 cells for 108°/cell) is not $-I$ there will be an insignificant perturbation to the vertical beta function which amounts to a $< 0.2\%$ vertical beta beat amplitude at full horizontal dispersion correction ($\pm$2.3-kGauss quadrupole fields). Of course these CQ quadrupoles may also be replaced by independent QF trim supplies of $\pm$5% range (SF horizontal movers would require a large $\pm$2.5-mm range and may steer too much). The dispersion correction specifications are summarized in Table 5-14.

This scheme of skew and normal quadrupole dispersion correction has been tested over various vertical emittance dilution mechanisms ranging to $\Delta\varepsilon_y/\varepsilon_{y0} \sim 30$. The dispersion was easily corrected with the skew/normal quadrupole pairs and then the arc was tracked with 3000 particles. The emittance dilution was corrected to within 4% without any coupling correction necessary. The remaining 4% dilution was due to betatron coupling which is also correctable as described above. There are also two quadrupoles placed in the dispersion suppressor section at the beginning and end of the arc which can be used to vary the horizontal dispersion function. However, these variable quadrupoles are included to properly match/suppress the arc dispersion function over the various R_{56} configurations (different horizontal phase advance per cell).

Dispersion Measurements. Dispersion correction for the arc can be accomplished in practice by correlating arc and post-arc BPM readings with prelinac energy variations and applying calculated corrections to the skew and normal quadrupole pairs, and/or by switching off the bunch compression rf between the arc and chicane (so that chicane dispersion errors are minimized), then minimizing the measured emittance on the post-chicane wire scanners using simple orthogonal combinations of the dispersion correction quadrupole pairs.

Beta Matching Corrections for the Arc. There are four matching quadrupoles placed at both entrance and exit of the arc which may be used for empirical beta matching into the post-chicane wire scanner section (with pre-chicane rf switched off to minimize dispersion effects). If the arc dispersion is first matched using the BPM correlation technique described above, then dispersion errors from the arcs should not bias the beta matching process significantly. The tolerance on the beta match into the arc is fairly loose. It can be shown that the synchrotron radiation (SR) emittance increase approximately scales with the amplitude of the incoming beta mismatch.

$$\Delta\varepsilon_{\rm SR} \approx B_{\rm mag}\Delta\varepsilon_{\rm SR-nom.} \tag{5.45}$$

Here $B_{\rm mag}(\geq 1)$ is the beta mismatch amplitude in the horizontal plane and $\Delta\varepsilon_{\rm SR-nom.}(\ll \varepsilon_{x0})$ is the nominal SR emittance increase for a matched incoming beam. A very large mismatch of $B_{\rm mag} = 2(\beta_x \approx 4\beta_{x0}, \alpha_x = \alpha_{x0} = 0)$ will increase a nominal 2.4% SR emittance increase to 4.8%. The vertical match has no such constraint.

Name	Quantity	Length (m)	Roll (mrad)	$\Delta B/B_0$ (10^{-5})	b_1/b_0 (10^{-4})	b_2/b_0 (10^{-3})
B1	2	10	0.14	2.1	33.5	377.0
B2	2	10	0.18	2.6	2.8	2.1

Table 5-15. *BC2 chicane dipole single element tolerances for 8% emittance growth each ($\gamma\varepsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\varepsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 1.5\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 100 mm. They generate 1st and 2nd order dispersion.*

Steering Correctors for the Arc. Horizontal and vertical dipole correctors at each QF and QD, respectively, will be required (depending on the availability of quadrupole movers) to initially steer the beam line and to use in fast feedback applications. Correctors with $\pm$0.4-kGauss fields and 10-cm length will be adequate to displace the beam nearly $\pm$1 mm at the next similar quadrupole. The horizontal correctors will then need to regulate at $\sim 5 \times 10^{-4}$ over the 100-ms (10 pulse) range while similar vertical correctors will need $\sim 4 \times 10^{-5}$ regulation ($\sim$0.5% luminosity loss due to all correctors in both arcs for both planes). Quadrupole movers will be used to realize the vertical alignment tolerances using beam based alignment techniques.

Chicane Section

As in the BC2 arc, the chicane presents mostly a dispersion correction problem. Since the chicane includes no quadrupoles at dispersion points, emittance dilution mechanisms will include dipole magnet roll errors and dipole field quality effects. Dipole roll errors should be correctable with vertical steering. Table 5-15 shows the approximate single element tolerances (2% luminosity reduction per error per beam) for the two chicane dipole locations; two B1 dipoles at chicane start/end and the two B2 dipoles at chicane center.

If these tolerances are not met, small correction quadrupoles and skew quadrupoles at the half-way point of the B1 dipoles (split at $\eta_x = 40$ mm, $\beta_{x,y} = 50$ m) can be added to control horizontal and vertical dispersion. These quadrupoles would need to be $\sim$10 cm in length, with $\sim$50-mm radius and $\pm$1-kGauss fields. Both dispersion phases can be handled by placing these correctors at the first and last B1 chicane dipole centers which are separated by $\sim \pi/2$ in phase. Betatron coupling and beta-mismatch effects will not be significantly generated due to the large dispersion here (*i.e.*, $\{\eta_x\sigma_\delta\}^2/\beta_x\varepsilon_x \sim 50$). If the sextupole field harmonic tolerances are not met, then small tuner normal sextupoles will also need to be included here to correct second order horizontal dispersion.

5.5.4 Space Charge and Coherent Radiation

The space-charge field of the beam is not insignificant when it exits the low-energy and the high-energy bunch compressors. This field can drive coupling resonances and increase the beam emittance or add to the beam halo. However, estimates of these effects indicate that they are small because the beams are rapidly accelerated and the optics has been chosen to avoid the most significant resonance, namely $2\psi_x \approx 2\psi_y$.

The space-charge fields in the bending magnets manifest themselves in a slightly different manner. There have been a few estimates of these effects which can be separated into a centrifugal space charge force (CSCF) and a coherent synchrotron radiation force (CSR). A calculation of the CSCF in bunched beams [Carlsten 1995] shows that this should not be a limitation in the NLC. In contrast, the CSR may be significant. Here, the longitudinal electric field of the coherent synchrotron radiation induces an energy variation along the bunch [Derbenev 1995]; the coherent radiation

field is similar to a longitudinal wakefield except that the field travels faster than the bunch and acts on leading particles rather than a normal wakefield which acts on trailing particles. In the wigglers and chicanes, this field will destroy the achromaticity of the compressors, thereby causing an increase in the horizontal emittance; the effect is not important in the 180° arcs where the dispersion is smaller and bends are shorter.

If we neglect the shielding effects of the vacuum chamber, the induced energy spread is predicted to be roughly [Derbenev 1995]

$$\sigma_\delta \approx 0.2 \frac{N r_0 L_d}{\gamma R^{2/3} \sigma_z^{4/3}} \quad , \tag{5.46}$$

where N is the bunch charge, L_d and R are the bending magnet length and bending radius, and r_0 and γ are the classical electron radius and the relativistic factor.

As stated, this induced energy spread will break the chromatic cancelation of the wiggler or chicane. The increase in the horizontal emittance is roughly [Raubenheimer 1995b]

$$\Delta\gamma\epsilon_x \approx 0.5\alpha^2 \sigma_x \frac{I}{17kA} \left(\frac{R}{\sigma_z}\right)^{1/3} \tag{5.47}$$

where α is the bend angle and I is the bunch current. In both the NLC low- and high-energy compressors, this predicts roughly 5% horizontal emittance growth.

Finally, the long-wavelength radiation is suppressed by boundary conditions imposed by the vacuum chamber. The point where shielding becomes important can be estimated as:

$$\sigma_z \gtrsim \sqrt{\frac{h^2 w}{\pi^2 R}} \tag{5.48}$$

where h and w are the full height and width of the chamber and R is the bending radius. In the low- and high-energy bunch compressors, we find that shielding is important for bunch lengths greater than 500 μm and 150 μm, respectively. These values are close to the fully-compressed bunch lengths and thus the chamber shielding may reduce the predicted emittance dilution.

At this time, this emittance dilution has not been observed experimentally and thus there is some uncertainty in the calculations. Regardless, it does not appear that the coherent radiation will be a significant limitation to the NLC compressors. It should be noted that this might be one reason not to consider a single stage compressor as the predicted dilution becomes much more important.

5.5.5 Ion Effects

At the start of the prelinac, all ions created by the beam are stably trapped within a single bunch train, except for hydrogen ions which are over-focused between bunches. An upper limit on the critical ion mass is obtained for the maximum current of 90 bunches with 1.5×10^{10} particles each. In this case the critical ion mass is six proton masses at the start and 29 at the end of the prelinac.

The ions will, first, increase the rate of filamentation, due to the additional focusing experienced by the trailing bunches, and, second, cause a nonlinear octupole-like coupling of horizontal and vertical betatron motion [Raubenheimer 1992]. At a pressure of 10^{-8} Torr, the focusing along the bunch train varies by $\Delta Q_y \approx 10^{-4}$ m^{-1} assuming 90 bunches of 1.5×10^{10} particles each. To attain an additional filamentation due to ions which is comparable to that caused by the intrinsic energy spread, a pressure of less than 10^{-9} Torr has to be achieved. Finally, a separation of the horizontal and vertical phase advances by about 5% effectively eliminates the effect of nonlinear coupling.

As in the damping rings and main linacs, the ionization of the residual gas may give rise to a transverse multibunch instability of the electron bunch train in the prelinac. At a pressure of 10^{-8} Torr, the estimated growth time τ_{asym} of this instability is about 90 ns [Raubenheimer 1995a], assuming 90 bunches of 1.5×10^{10} particles each and ignoring decoherence of the ion oscillations. The ion-driven instability causes an increase of the Fourier component at the ion-oscillation frequency of the vertical bunch positions. Traversing the 500-m-long prelinac at a pressure of 10^{-8} Torr amplifies an initial bunch-to-bunch displacement at this Fourier component, characterized by an action value J_{y0}, by a factor 3×10^4. As an example, if the original perturbation is due to the finite number of electrons in the train, the initial action value is about $J_{y0} \approx \sigma_y^2/(2\beta_y N_b n_b) \approx 3 \times 10^{-24}$ m-rad, which yields a negligible emittance growth of $\Delta(\gamma\epsilon_y) \approx 2 \times 10^{-15}$ m-rad.

However, assuming that the initial perturbation is due to a realistic bunch-to-bunch orbit jitter of about 1% σ_y, the emittance blow-up can be as large as 500%. Fortunately, ion decoherence, due to the variation of the vertical ion oscillation frequency with the horizontal position, increases the rise time of the instability by about a factor of two or three, thereby relieving the pressure tolerance by the same factor [Stupakov 1995].

Taking the decoherence into account, and assuming a partial pressure of 5×10^{-9} Torr of CO, the emittance increase resulting from an initial 1% σ_y harmonic bunch-to-bunch offset at the ion frequency would be about 10%. Because the main constituent of the vacuum is hydrogen, which is not trapped in the beam, this is a relatively loose vacuum requirement.

Until further studies and experiments improve our understanding of the ion-driven instability and possible cures, it is assumed that the required pressure in the prelinac is below 10^{-8} Torr of which more than 70% would be H_2.

5.5.6 Synchrotron Radiation

Incoherent synchrotron radiation in bending and quadrupole magnets may increase the energy spread and, more importantly, dilute the horizontal and (if linear coupling or vertical dispersion is nonzero) the vertical emittance. The total emittance growth in the bunch compressor should be smaller than 5% horizontally, and less than 1–2% vertically. The latter number translates into a tolerance for the tuning procedures and the diagnostics. In addition, the induced energy-spread should be small compared with the original energy spread of the bunch.

The effects of synchrotron radiation in the longitudinal and the horizontal phase space are summarized in Table 5-16 for wiggler, arc, and chicane. In all cases, the number of radiated photons per electron, N_γ, is much larger than one so that the resulting net effect can be described by a Gaussian distribution. The critical energy $E_c \equiv 3\hbar c\gamma^3/(2\rho)$ is only a few keV. Therefore, the radiating particles stay inside the energy acceptance of the system. The average energy loss, ΔE_{rad}, and the additional rms momentum spread due to synchrotron radiation, $\Delta\delta_{\rm rms}$, are both small. The change of the horizontal emittance can be estimated from [Raubenheimer 1994a]

$$\Delta(\gamma\epsilon_x) \approx 4 \cdot 10^{-8} \left(\frac{E}{\text{GeV}}\right)^6 I_5 \tag{5.49}$$

where I_5 denotes the fifth synchrotron radiation integral [Helm 1973]. A mismatch between the synchrotron radiation effects and the original beam distribution can lead to an emittance growth which is larger than predicted by Eq. 5.49 [Spence 1993]. This difference could be especially important for the chicane, since the phase advance over the four bending magnets is only about 90°. To estimate the actual emittance growth, a simulation study has been performed using the computer code MAD [Grote 1990]. The simulations show an emittance growth of approximately $\Delta(\gamma\epsilon_x) \approx 2 \times 10^{-9}$ m-rad in the chicane, which is consistent with the simple estimate of Eq. 5.49, and indicates that the effect of the radiation-lattice mismatch is small. The total increase of the horizontal normalized emittance in the bunch compressor is then about 2.6%. We conclude that the effects of incoherent synchrotron radiation appear insignificant.

	Wiggler	Arc	Chicane
N_γ	99	662	28
E_c (keV)	3.5/4.6	25	7.4
ΔE_{rad} (keV)	134	5300	66
$\Delta\delta_{\rm rms}$ (%)	$1.3 \cdot 10^{-3}$	$4 \cdot 10^{-3}$	$8 \cdot 10^{-5}$
$\Delta(\gamma\epsilon_x)$ (10^{-6} m-rad)	0.039	0.038	0.0015
$\Delta(\epsilon_x)/(\epsilon_{x\,0})$ (%)	1.3	1.3	0.05

Table 5-16. *Effects of synchrotron radiation in wiggler, arc, and chicane.*

5.6 Tolerances

5.6.1 Alignment and Field Tolerances

These are discussed in the text. In general, all transverse alignment and magnet tolerances in the bunch compression system are relatively loose. The only field tolerances that are tight are those in the spin rotator system and the bending magnet multipoles in the chicane. Both of these tolerances are eased using simple tuning techniques similar to those routinely used in the SLC bunch compressors. All of the magnet alignment tolerances are substantially larger than the expected resolution from the beam-based alignment; the alignment techniques are discussed in greater detail in Chapter 7.

5.6.2 Tolerances on RF Phase, Energy, and Intensity

Some of the tightest tolerances in the bunch compressor system are those on the (relative) rf phase stability. The tolerance on the phase difference of electron and positron compressor rf is set by the loss of luminosity when the beams do not collide at the minimum waist position. A collision which is longitudinally offset by $\Delta s = 0.2\beta_y^\star$ leads to a 2% loss of luminosity, which translates into a maximum tolerable rf phase difference $\Delta\phi_{SB,e^+e^-}$ or $\Delta\phi_{LB,e^+e^-}$ between the S-band and L-band rf systems, respectively, on the electron and positron side. This tolerance can be as tight as 0.17° (Table 5-17) and may be difficult to meet because the compressors are located more than 20-km apart.

Even tighter tolerances on relative-phase stability have to be achieved, on either side of the IP, between different compressor and linac rf systems. If the beam is injected into the main linac at a wrong phase, the final beam energy changes. This energy change sets a tolerance on the relative rf phase stability of linac and compressor rf, and on the allowed bunch-to-bunch energy variation due to multibunch beam loading or intensity variation in the various compressor subsystems.

In case the beam is nominally off-crest by an X-band phase angle ϕ_{XB}, the maximum allowed phase error $\Delta\phi_{XB}$ (in degree), which causes a 0.1% average energy change, is given by

$$\Delta\phi_{XB} \leq \frac{10^{-3}\,360}{2\pi\tan\phi_{XB}} \tag{5.50}$$

		NLC-Ia	NLC-Ic	NLC-IIa	NLC-IIc
$E_{\max}$	(GeV)	267	233	534	473
$N_{\rm part}$	(10^{10})	0.65	0.85	0.95	1.25
Final σ_z	(μm)	100	150	125	150
σ_y^*	(nm)	4.52	7.00	3.57	5.35
θ_y^*	(μrad)	36.17	35.02	28.59	26.75
$\Delta\phi_{SB\ e^+e^-}$	(°)	0.17	0.27	0.17	0.27
$\Delta\phi_{LB\ e^+e^-}$	(°)	0.50	0.56	0.44	0.62
$\Delta z_{e^+e^-}$	(μm)	50	80	50	80
ϕ_{XB}	(°)	-14.4	-15.1	-7.2	-11.0
ϕ_{PL}	(°)	-4	-5	-5	-8
$\Delta\phi_{XB}$	(°)	0.22	0.21	0.45	0.30
$\Delta\phi_{SB}$	(°)	0.055	0.053	0.112	0.073
$\Delta\phi_{LB}$	(°)	0.16	0.11	0.29	0.16
$\Delta\phi_{PL}$	(°)	0.32	0.22	0.57	0.33
Δz_{XB}	(μm)	16	15	33	21
Δz_{SB}	(μm)	16	15	33	21
Δz_{LB}	(μm)	95	63	167	96
Δz_{PL}	(μm)	95	63	167	96
$\left.\frac{\Delta E}{E}\right\|_{XB}$	(10^{-3})	1.0	1.0	1.0	1.0
$\left.\frac{\Delta E}{E}\right\|_{SB}$	(10^{-3})	0.45	0.43	0.91	0.59
$\left.\frac{\Delta E}{E}\right\|_{LB}$	(10^{-3})	0.19	0.13	0.33	0.19
$\left.\frac{\Delta E}{E}\right\|_{PL}$	(10^{-3})	4.0	4.7	6.1	5.3
$\frac{\Delta N}{N}_{XB}$	(%)	6.4	4.5	8.1	5.7
$\frac{\Delta N}{N}_{SB}$	(%)	30	21	40	20
$\frac{\Delta N}{N}_{LB}$	(%)	290	148	351	154
$\frac{\Delta N}{N}_{PL}$	(%)	116	105	121	80

Table 5-17. *Tolerances on rf phase, single-bunch energy, and intensity in the various compressor subsystems for four different NLC scenarios, corresponding to either 2% loss of luminosity or to a 0.1% change of average bunch energy, at the end of the main linac. Subindices XB, SB, LB and PL refer to main X-band linac, S-band rf of the second compressor, L-band rf of the first compressor, and the prelinac (S-band rf), respectively.*

and the corresponding longitudinal position error (meter) in main linac and second compressor is

$$\Delta z_{XB} \leq \frac{10^{-3}\,\lambda_{XB}}{2\pi \tan\phi_{XB}} \tag{5.51}$$

where λ_{XB} denotes the wavelength of the X-Band rf (26 mm). The maximum phase error $\Delta\phi_{SB}$ for the S-band rf of the second compressor is equal to

$$\Delta\phi_{SB}\,(^\circ) \leq \frac{\Delta z_{XB}\,360}{\lambda_{SB}}, \tag{5.52}$$

and the maximum single-bunch energy variation due to multi-bunch beam loading in the S-Band rf section is

$$\left.\frac{\Delta E}{E}\right|_{SB} \leq \frac{\Delta z_{XB}}{R^c_{56}} \tag{5.53}$$

The tolerances for the first (L-Band) compressor are obtained in a similar way, taking account of the compression ratio $-R^c_{56}/R^a_{56}$ of the second stage. We thus have

$$\Delta z_{LB} \leq -\frac{\Delta z_{XB}\,R^a_{56}}{R^c_{56}}\quad , \tag{5.54}$$

$$\Delta\phi_{LB}\,(^\circ) \leq \frac{\Delta z_{LB}\,360}{\lambda_{LB}}\quad , \tag{5.55}$$

$$\left.\frac{\Delta E}{E}\right|_{LB} \leq \frac{\Delta z_{LB}}{R^w_{56}}\quad . \tag{5.56}$$

There is also a tolerance on the bunch-to-bunch energy variation in the prelinac. A relative energy error δ_{PL} at the end of the prelinac causes a change of the longitudinal position Δz_{XB} of the bunch center at the exit of the second compressor, which in lowest order is given by

$$\Delta z_{XB} \approx \frac{1}{2} f_{2,c} k_{2,c} {R^a_{56}}^2 R^c_{56} \delta^2_{PL} + T^c_{566}(-f_{2,c}R^a_{56}+1)^2\delta^2_{PL} + \mathcal{O}(\delta^3_{PL}) \tag{5.57}$$

from which the maximum allowed energy variation $\Delta E/E|_{PL} \equiv \delta_{PL,\,max}$ in the prelinac can be estimated:

$$\left.\frac{\Delta E}{E}\right|_{PL} \leq \frac{3}{4}\left[\frac{\Delta z_{XB}}{\frac{1}{2} f_{2,c} k_{2,c} {R^a_{56}}^2 R^c_{56} + T^c_{566}(-f_{2,c}R^a_{56}+1)^2}\right]^{\frac{1}{2}} \tag{5.58}$$

where the factor $3/4$ in front is an empirical correction due to finite bunch length and higher-order contributions.

A phase error in the prelinac causes both an energy error at the exit of the prelinac and a phase error in the main linac. Both these effects lead to a change of the final energy, but the latter effect is ten times as severe as the former. For this reason, the energy error at the end of the prelinac may be ignored, and the tolerance on the longitudinal position in the prelinac is the same as that for the L-Band compressor, while the phase tolerance (in degree S-Band) is simply twice that of the L-Band rf.

A limit on single-bunch intensity variation arises from the above tolerances on energy and from the single-bunch beam loading. Denoting the relative energy change per 10^{10} particles due to longitudinal wakefields by δ_{10}, the intensity tolerance for an rf section may be written as

$$\frac{\Delta N}{N} \leq \frac{\Delta E}{E}\cdot\frac{10^{10}}{\delta_{10}\,N}, \tag{5.59}$$

where $\Delta E/E$ designates the allowed energy variation.

		NLC-Ia	NLC-Ic	NLC-IIa	NLC-IIc
$\Delta\phi_{c,\,BC2}$	(°)	0.76	0.56	1.06	0.53
$\Delta\phi_{c,\,BC1}$	(°)	0.81	0.42	0.99	0.44
$\Delta\phi_{c,\,PL}$	(°)	0.17	0.20	0.26	0.23
$\Delta f\|_{c,\,BC2}$	(kHz)	17	12	23	12
$\Delta f\|_{c,\,BC1}$	(kHz)	18	9	22	10
$\Delta f\|_{c,\,PL}$	(kHz)	3.7	4.5	5.7	5.0

Table 5-18. *Tolerances on phase and frequency stability for the additional rf systems which are employed to compensate multibunch beam loading, considering four different NLC scenarios. Tolerances correspond to a 0.1% change of average bunch energy, at the end of the main linac. Subindices $BC2$, $BC1$, and PL refer to the second compressor, first compressor, and prelinac, respectively. The subindex `c' indicates the compensating nature of the rf.*

Finally, if the ΔF beam loading compensation is used in either of the two compressor stages or the prelinac, then there are two additional rf systems which are used for multibunch beam loading compensation. The tolerance on the phase stability for these rf systems follows directly from the tolerable energy error ΔE in the pertinent compressor subsystem:

$$\Delta\phi_{mbc} \leq \Delta E / V_{mbc} \tag{5.60}$$

where V_{mbc} is the amplitude of the compensating rf voltage. The limit on the energy deviation, ΔE, is taken from Eqs. 5.53, 5.56 or 5.58, respectively. Similarly, the allowed frequency error for the compensating rf is determined by the corresponding phase (or energy) error at the end of the bunch train, *i.e.*,

$$\frac{\Delta f}{f} \leq \frac{c\,\Delta E}{2\pi V_{mbc} l_{\rm train} f} \quad . \tag{5.61}$$

Table 5-17 compiles tolerances on main rf and T_{566}-compensating rf for all compressor subsystems and four different NLC scenarios. Some tolerances shown in the table are easy to achieve, for instance those on the intensity. Other tolerances may need more attention. For example, the phase-stability tolerance for the second compressor is less than a tenth of a degree S-Band.

Table 5-18 lists tolerances on phase and frequency stability for the three double rf systems which compensate for multibunch beam loading in the different compressor subsystems. In this case, the tolerances are tightest for the prelinac, where a phase stability of about 0.2° S-band needs to be achieved. By comparison, the phase tolerances for the compensating rf of the two compressor stages are at least 0.4° S-band. They are, thus, looser than those of the corresponding main rf systems.

5.7 Conclusions and Comments

A design of a two-stage bunch compressor has been described that meets all specifications, as far as single- and multibunch dynamics is concerned. The bunch compressor has been designed in view of NLC-I and NLC-II, *i.e.,* for a c.m. energy of 0.5 and 1 TeV, and may have to be reevaluated once a specific operating range for NLC-III, with 1.5 TeV-c.m. energy, is determined.

Possible topics of further study in the bunch-compressor design are the following: First, if attempts are successful to correct the synchronous phase variation in the damping rings, the bunch compressor may be simplified and shortened. Second, a final value has to be chosen for the $R_{56}^{a,c}$ matrix elements of arc and chicane. A reduction of the present values would reduce the (significant) length of the rf system in the second compressor, at the expense of increased sensitivity to the incoming beam conditions. In particular, the choice of $R_{56}^{a,c}$ depends, of course, on the assumed initial phase variation. Third, the possibility of and schemes for bunch shaping need to be explored. Also, the design of two 90° compressor stages as an alternative to the present "90°/360°" scheme could be investigated. Next, one or both compressor stages may be redesigned so as to reduce the T_{566}-matrix elements. This would improve the performance in the longitudinal phase space, and could also reduce the length of the bunch compressor.

As a final point, we note that a single-stage compressor remains the most attractive alternative to the present design. A single stage would be shorter, easier to tune, and would improve the longitudinal dynamics at high intensity. The main disadvantage is that the nonlinearities in the longitudinal phase space are more severe, because of the much larger energy spread. Consequently, for a single stage, the T_{566}-matrix element has to be reduced to about a third of the R_{56}-matrix element. The residual nonlinearity can then be compensated by a decelerating rf as in the double-stage scheme. A small T_{566}-matrix element may be realized by placing quadrupoles at dispersive locations and/or by adding sextupoles. In reducing the T_{566}, care has to be taken, since the emittance dilution in the transverse phase space has to remain acceptable for an rms energy spread as large as 5%. So far, attempts to design a single-stage compressor with satisfactory performance in all phase-space dimensions have not been successful. Nevertheless, this option should be pursued further.

References

[Bane] The program LITRACK for simulating longitudinal wakefield effects in linear accelerators has been written by K. Bane.

[Bane 1995] K. Bane, "The Short-Range Wake Field for the NLC X-Band Structure", NLC–Note–9 (1995).

[Brown 1977] K. Brown, F. Rothacker, D. Carey, C. Iselin, "TRANSPORT – A Computer Program for Designing Charged Particle Beam Transport Systems", SLAC–91 (1977).

[Brown 1980] K.L. Brown *et al.*, CERN 80-04, (March 1980).

[Burke 1995] D. Burke, "Proposal for New NLC Operating Plane", SLAC memorandum, unpublished (1995).

[Carlsten 1995] B.E. Carlsten and T.O. Raubenheimer, "Emittance Growth of Bunched Beams in Bends," *Phys. Rev. E* **51**:1453 (1995).

[Delahaye 1985] J.P. Delahaye and J. Jäger, "Variation of the Dispersion Function, Momentum Compaction Factor and Damping Partition Numbers with Particle Energy Deviation", *Particle Accelerators* **18**, 183 (1986).

[Derbenev 1995] Y.S. Derbenev, J. Rossbach, E.L. Saldin, V.D. Shiltsev, "Microbunch Radiative Tail-Head Interaction", (DESY) TESLA–FEL 95–05 (1995).

[Emma 1994] P. Emma, "A Spin Rotator System for the NLC", NLC–Note–7 (1994).

[Emma 1995] P. Emma, "A Skew Correction and Diagnostic Section for Linear Colliders", in preparation (1995).

[Fieguth 1987] T. Fieguth, "Snakes, Rotators, Serpents and the Octahedral Group", SLAC–PUB–4195 (January, 1987).

[Fieguth 1989] T.H. Fieguth, "The Optical Design of the Spin Manipulation System for the SLAC Linear Collider", *Proc. of the 1989 Part. Acc. Conf.*, Chicago, IL (1989).

[Grote 1990] H. Grote and F. C. Iselin, *The MAD Program, Version 8.1, User's Reference Manual*, CERN/SL/**90-13** (AP) (1990).

[Helm 1973] R.H. Helm, M.J. Lee and P.L. Morton, M. Sands, "Evaluation of Synchrotron Radiation Integrals", *Proc. of the 1973 Part. Acc. Conf.*, San Francisco, CA; *IEEE Trans. Nucl. Sci.* NS–20, **3**, 900 (1973).

[Kikuchi 1992] M. Kikuchi, "Multi-Bunch Energy Compensation in Bunch Compressor of Linear Colliders", *Proc. of the 1992 High Energy Acc. Conf.*, Hamburg, Germany, 864 (1992).

[Miller 1995] R. Miller, private communication (1995).

[Oide 1994] K. Oide, discussed at XLC miniworkshop, SLAC, Dec. 1994 (1994).

[Palmer 1990] R. Palmer, "Prospects for High-Energy e^+e^- Linear Colliders", *Ann. Rev. Nucl. Part. Sci.* **40**, 529–592 (1990).

[Raubenheimer 1992] T.O. Raubenheimer and P. Chen, "Ions in the Linacs of Future Linear Colliders", *Proc. of LINAC 92*, Ottawa, Canada (1992).

[Raubenheimer 1993] T.O. Raubenheimer, P. Emma, S. Kheifets, “Chicane and Wiggler Based Bunch Compressors for Future Linear Colliders”, *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC (1993).

[Raubenheimer 1994a] T.O. Raubenheimer, “Bunch Compressor Parameters”, NLC–NOTE–2 (1994).

[Raubenheimer 1994b] T.0. Raubenheimer, private communication (1994).

[Raubenheimer 1995a] T.O. Raubenheimer and F. Zimmermann, “A Fast Beam-Ion Instability: Linear Theory and Simulations”, *Phys. Rev. E* **52**, 5487 (1995).

[Raubenheimer 1995b] T.O. Raubenheimer, “Summary of the Linac Working Group”, Proc. of the MicroBunches Workshop, Brookhaven, NY (1995).

[Rogers 1991] R. Rogers and S. Kheifets, “Misalignment Study of the NLC Bunch Compressor”, Proc. of the 1991 IEEE Part. Acc. Conf., San Francisco, CA (1991).

[Spence 1991] W. Spence, private communication (1991).

[Spence 1993] W. Spence, “Phase Space and Synchrotron Radiation in High Energy Electron Transport Lines”, *Proc. of the Emittance 93 Workshop*, KEK **93-13** 362 (1993).

[Stupakov 1995] G. Stupakov, T.O. Raubenheimer and F. Zimmermann, “A Fast Beam-Ion Instability: Effect of Ion Decoherence”, *Phys. Rev. E* **52**, 5499 (1996).

[Woodley 1994] M. Woodley, private communication (1994).

[Zimmermann 1994] F. Zimmermann, “Longitudinal Single-Bunch Dynamics and Synchrotron Radiations Effects in the Bunch Compressor”, NLC–Note–3 (October 1994).

Contributors

- Paul Emma
- Zenghai Li
- Tor Raubenheimer
- Kathy Thompson
- Frank Zimmermann

6

Low-Frequency Linacs and Compressors

Contents

6.1 Introduction

The NLC complex has seven S-band linacs and four L-band linacs. The main L-band linac accelerates the positron beam from the positron target to 2 GeV where it enters the damping ring. In addition there are one 80-MeV L-band linac associated with the positron energy compressor at the output of the positron booster linac and two 134-MeV L-band linacs associated with the electron and positron first stage compressors at the output of the two damping rings. These linacs are run in phase quadrature with the beam to introduce the linear correlation between time and energy within each bunch which is required for compression. The seven S-band linacs are:

1) The Positron Drive Linac which accelerates the electrons which strike the positron target.
2) The Electron Booster Linac which accelerates the electron beam to 2 GeV where it enters the Electron Damping Ring.
3) 42-MeV electron energy compressor at the output of the electron booster linac.

4,5) The Electron and Positron Preaccelerators which accelerate the beams from the damping rings to 10 GeV where the second bunch compression occurs.

6,7) 3.85-GeV S-band linacs to produce the linear energy correlation required for each of the second-stage compressors.

These linacs share enough common characteristics and design problems that it is appropriate to discuss the rf design of these linacs together in this chapter. The principal design issues confronting the rf design of these linacs are beam loading and dipole wakefields. Except for the first 250-MeV portion of the Positron L-band linac which is inside a 0.5-T solenoid, all of these linacs must be capable of accelerating about 2-A beams with about 100-ns macropulse length. The part of the Positron L-band linac inside the solenoid needs to accelerate a total effective current of about 14 A since both electrons and positrons get captured onto their respective accelerating crests, $\lambda/2$ apart. Because the electrons are produced by ionization as well as by pair production, there are about 65% more electrons than positrons. Furthermore, the solenoid transports many particles which are outside the phase and energy range acceptable to the pre-damping ring. These excess particles will be scraped off in the energy analyzing slits after the bend immediately after the solenoid, but in the two capture sections within the solenoid they create extremely heavy beam loading.

6.2 Beam-Loading Compensation

All of these linacs are fairly heavily loaded by a beam with a pulse length much shorter than the reasonable filling times of the structures, which are in turn shorter than the ringing time, $2Q/\omega$, for the structure. In this situation, the energy of the beam will drop approximately linearly with time during the pulse as a result of beam loading. For this case there are two natural choices for beam-loading compensation: 1) ΔT (or Early Injection), *i.e.*, injecting the beam into the structure before the structure is full; 2) ΔF, *i.e.*, having one or more accelerator structures running at a frequency 1 to 2 MHz above or below the nominal frequency and roughly in phase quadrature from the accelerating phase. Thus the beginning of the pulse can be decelerated by the off-frequency section(s), while the end of the pulse is accelerated. The advantage of ΔT compensation is that the compensation occurs in every accelerator section, so that the energy spectrum can be good throughout the linac, thus minimizing emittance growth from dispersion and chromatic effects.

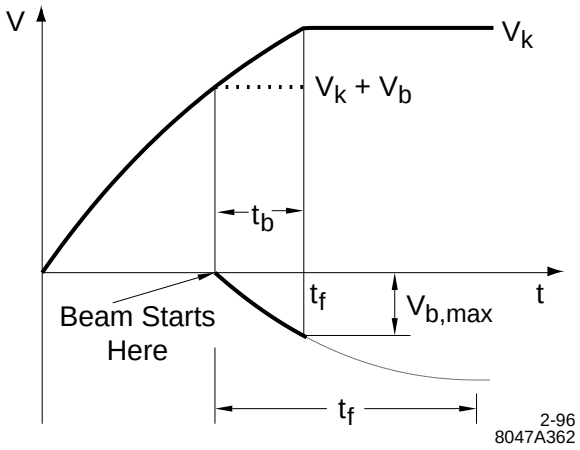


Figure 6-1. *ΔT beam-loading compensation for rectangular rf input pulse.*

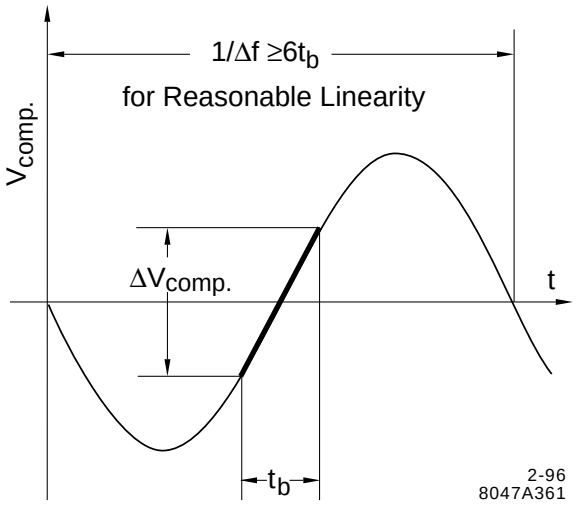


Figure 6-2. *Energy gain, $V_k(t)$, in a ΔF compensation section.*

The way ΔT compensation works is shown in Figure 6-1 in which the voltage $V_k(t)$ produced by a step-function rf pulse is plotted as a function of time for a traveling-wave linac section. Also plotted is the beam-induced voltage $V_b(t)$. The resultant sum of $V_k(t)$ and $V_b(t)$ is plotted for the case where the beam is turned on before the linac structure is full. To illustrate the principle of ΔF compensation, the voltage gain $V_{\rm comp}(t)$ of a beam with an F_0 bunch structure in a ΔF-compensation section operating at a frequency $F_0 \pm \Delta$F is plotted in Figure 6-2. The beam is bunched with a bunch-repetition frequency of 714 MHz, a subharmonic of the accelerator frequency F_0. In an accelerator section powered by rf at a frequency $F_0 \pm \Delta F$, the bunches see a field which appears to vary with the difference frequency ΔF as shown in Figure 6-2. If the beam pulse length satisfies the relation $t_b \leq \frac{1}{6\Delta F}$ and is phased as shown, the energy gain will vary quite linearly with time.

It is instructive to calculate the power required for compensation by early injection for a rectangular rf pulse (non-sledded) in the approximation that $t_b << t_f << 2Q/\omega$, where t_b is the beam pulse length and t_f is the structure

filling time.

$$\frac{dV_b}{dt} = -\frac{i}{2}\frac{\omega r}{Q}L \tag{6.1}$$

$$\frac{dV_k}{dt} = \left(\frac{P v_g \omega r}{Q}\right)^{1/2} \tag{6.2}$$

$$V_k^2 = L^2 \frac{P}{v_g}\frac{\omega r}{Q} \tag{6.3}$$

where V_b= beam-induced voltage
i = beam current
L = accelerator length
V_k= voltage produced by klystron power
P = klystron power
v_g= average group velocity in the structure

Setting $\frac{dV_k}{dt} = -\frac{dV_b}{dt}$ and multiplying by Eq. 6.3 we find that

$$P = \frac{iV_k}{2} \tag{6.4}$$

This remarkable result says that for compensation by early injection the klystron power has to be approximately equal to half the power absorbed by the beam *and is independent of the shunt impedance of the structure and the length of the linac* under the inequalities stated above. For compensation by early injection with the approximation used here the filling time must satisfy the relation:

$$\frac{t_b}{t_f} = \frac{V_b}{V_k} \tag{6.5}$$

With ΔF compensation one is free to pick the filling time. In the approximation that $t_b << t_f$, and that r/Q is independent of group velocity, the beam-induced voltage is independent of filling time (see Eq. 6.1). For filling times short compared with $2Q/\omega$ the peak power required varies inversely with filling time. One can also reduce the power required by decreasing the gradient. For a given required voltage, the power varies linearly with gradient, however the beam-induced voltage varies linearly with length or inversely with gradient. Increasing the beam-induced voltage tightens the tolerance on the compensation and increases the fraction of the overall length and the fraction of the total rf power which must be dedicated to the off-frequency compensating sections. One finds that for either ΔT or ΔF compensation there is an optimum product Lr/Q which minimizes the power required for accelerating the beam to a given voltage and compensating for beam loading. The optimum is very broad and changing the product by a factor of two in either direction only raises the power required by about 15%. We have chosen an unloaded gradient of 25 to 30 MeV/m for the S-band accelerators to minimize dark current problems. This determines the length of the accelerator required for a given beam energy. For a fixed length the beam loading for a short pulse varies linearly with r/Q, so we pick a value about half of the optimum which increases the power required about the 15% mentioned above the minimum. With these choices the power and lengths required for ΔT compensation are somewhat greater (about 10%) than for ΔF. Our present feeling is that the ΔF system is probably easier to operate. Depending on the percentage beam loading and the energy spread after compensation, a total phase shift of 1/2 to one radian is likely to be acceptable. Thus an off-frequency compensating section can correct for a beam-loading voltage V_b between 1/2 and one times the peak voltage of the section.

The principal disadvantage of the ΔF compensation, is that the beam energy spread reaches half of the compensation of a single off-frequency section. The compensation section then overcorrects by a factor of two which reverses

Frequency	2856 MHz
Structure type	DDS disk-loaded[a]
Structure length	3 m
# sections per module	4
Filling time	371 ns
Group velocity	0.05 c to 0.0125 c
# klystrons per module	2
Klystron power	65 MW
Rf pulse length (flattop)	4 μs
Pulse compression	SLED I

[a] Damped Detuned Structure

Table 6-1. *S-band linac rf parameters.*

correlation of energy with time during the pulse. In order to maintain a small enough energy spread to achieve an acceptable emittance growth it appears necessary to distribute the power from one klystron running off frequency to a number of short accelerator sections, so that each correction is acceptably small. The compensation sections should be placed so that the beam-loaded particles have the same total betatron phase advance as the first particles in the bunch train which have no beam loading. This is approximately equivalent to saying that they should spend as much time (or distance along the accelerator axis) high in energy as they do low in energy. That is the spatial average along the accelerator axis of the difference in energy between the first particle and the last particle should vanish. One can either choose to have short ΔF sections after each regular sections or to have one or two long ΔF sections in each of the accelerator modules. It appears to us that the high power microwave distribution system to many short ΔF compensation sections becomes unreasonably complicated and expensive. Using long ΔF sections will, however, result in poor energy spectrum and associated emittance growth. Therefore, we have chosen to use ΔT compensation for most of the low frequency linacs. The positron capture linac (inside the solenoid) requires ΔF compensation to achieve a acceptable multibunch energy spread because of the extremely high current. It is desirable to use ΔF compensation on the S-band bunch compressors, because the tuning to minimize the energy spread is very straightforward for ΔF compensation.

Table 6-1 presents the preliminary parameters for the S-band structure and rf system. The klystrons are assumed to be like the 65-MW 5045 klystrons, 240 of which power the existing SLAC linac. The useful power at the load is assumed to be 60 MW to allow for attenuation.

L-band Positron Linac: The initial acceleration of the positrons, from the target to the Positron Damping ring at 2 GeV, is done in a L-band linac with 1428-MHz rf power. The reason for this is that the four-dimensional phase space acceptance of a linac varies as λ^4. Thus, accelerating in a 1428-MHz linac instead of a 2856-MHz linac increases the phase volume acceptance of the linac by a factor of 16. This makes the production of the high-intensity ($\approx 10^{12}\ e^+$ per rf pulse) positron beam required for NLC comfortable for a conventional electromagnetic shower positron source. The beam-induced voltage per unit length for a beam pulse length much less than the filling time varies inversely as λ^2, so beam loading becomes less of a problem. 75-MW klystrons are used for accelerating the beam. For the first 200 MeV, the beam is focused by a solenoid. In this region we assume the beam loading is produced by a beam current of up to 14 A, since both electrons and positrons will be captured on the appropriate E field crests. The L-band linac rf parameters are given in Table 6-2.

Frequency	1428 MHz
Structure type	Gaussian detuned
Structure length	5 m
# sections per module	4
Filling time	675 ns
Group velocity	0.046 c to 0.0115 c
# klystrons per module	2
Klystron power	75 MW
Rf pulse length (flattop)	5 μs
Pulse compression	SLED I

Table 6-2. *L-band linac rf parameters.*

6.3 Numerical simulation of beam-loading compensation

The low-frequency linacs can be grouped into four groups based on their commonalities in structure and beam loading. In the following subsections we discuss the beam-loading compensation of each of the linac groups.

Two kinds of klystrons, S65 and L75, will be used to power the low-frequency linacs. The S65 klystron will be SLAC 5045 65-MW-like klystrons (see Section 6.4.1 for details), and capable of delivering 60-MW useful power, allowing for about 10% loss in the rf components. The L75 klystrons are L-band 75-MW klystrons (see Section 6.4.2 for details) which can deliver 67.5-MW useful power. We assume that the klystrons run at full power except when otherwise specified.

The numerical simulations in this section assume that the accelerator structures are dispersionless, which, we believe, is a good approximation. In this section, "energy spread" refers to the relative rms energy spread, and it is relative to the energy gain of that linac, except when other wise specified. The actual energy spread will be a factor of $\Delta E_{\rm gain}/(\Delta E_{\rm gain} + E_0)$ smaller, where E_0 is the initial beam energy. The energy spread obtained in this section is the residual of the beam-loading compensation. The total energy spread of the beam should be obtained by adding in quadrature the initial energy spread to the energy spread obtained in this section.

6.3.1 The e^+ Drive, the e^- Booster and the (e^-, e^+) Prelinacs

The e^+ drive linac, the e^- booster linac, and the (e^-, e^+) prelinacs use the same kinds of klystrons and acceleration structures. It is adequate to discuss the beam-loading compensation together in this section. The only differences among these linacs are the final energy and the beam current, as shown in Table 6-3. The acceleration structure will be S-band Damped Detuned Structure (DDS) ($F = 2.856\,$GHz). The beam-loading voltage is compensated by use of the ΔT compensation method.

The ΔT compensation requires that the beam be injected before the structure is full. The condition

$$\frac{dV_k}{dt} = -\frac{dV_b}{dt} \tag{6.6}$$

is essential for compensating the beam-loading voltage for the later bunches. The profile of the acceleration voltage, as a function of time, of a SLED-I driven structure is as shown in Figure 6-3. The bunch train occupies the region marked by the dotted lines. The slope of the voltage is inversely proportional to the filling time of the structure and

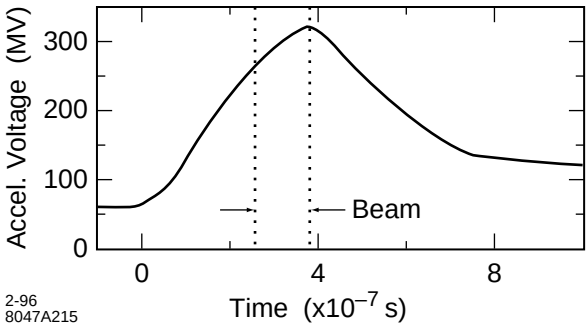


Figure 6-3. *Acceleration voltage of a SLED-I driven structure.*

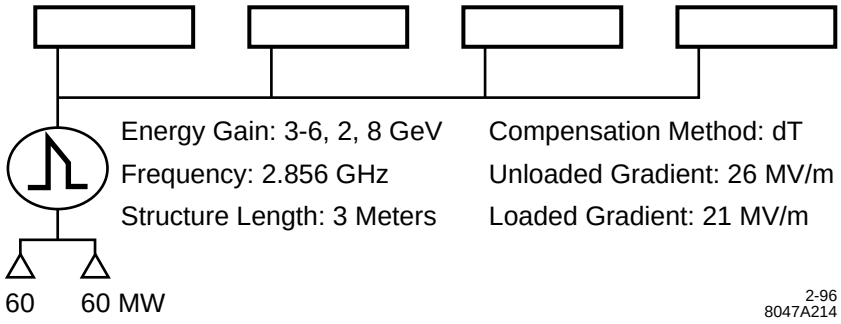


Figure 6-4. *An accelerator module for the e^+ drive, e^- booster, and the (e^-, e^+) prelinacs.*

depends on the profile of the SLED-I output. In order that Eq. 6.6 be satisfied, the filling time of the structure needs to be chosen properly, and the SLED output may also need to be modulated. The filling time of the structure in concern is chosen to be 0.371 μs, which is about three times the length of the bunch train. The beam loading is about 30% of the acceleration for a beam current of 2.2 A. Amplitude modulation of the SLED-I output is required to increase the derivative of the acceleration voltage. The amplitude modulation of the SLED-I output is obtained by amplitude modulating the klystron output. Our design uses a pair of klystrons to drive one SLED-I cavity. The phases of the two klystrons are modulated in opposite directions to obtain a combined amplitude modulated output. The klystrons will be S65 klystrons. A schematic drawing of an accelerator module is shown in Figure 6-4.

The energy spread due to the beam loading of a 2.2-A beam current can, in principle, be compensated to less than 2.4×10^{-5}, and less than 1.1×10^{-5} for a 1.5-A beam current. Since the energy spread is relative to the energy gain of the linac, the actual energy spread will be lower if the initial energy of the beam is not zero. The average gradients in the structure for the two cases are 17.3 MV/m and 21.2 MV/m respectively. The energy spectrum and the SLED-I wave form used for compensating the beam loading of a 2.2-A beam current are shown in Figure 6-5.

6.3.2 The e^+ Capture Linac

The accelerator structure for the e^+ capture linac will be detuned L-band structure (F=1.428 GHz) in order to have a large acceptance. The total energy gain of this linac is about 240 MeV. The beam is focused by a solenoid. We assume the beam loading in this linac is produced by a beam current of up to 14 A. For a typical disk-loaded L-

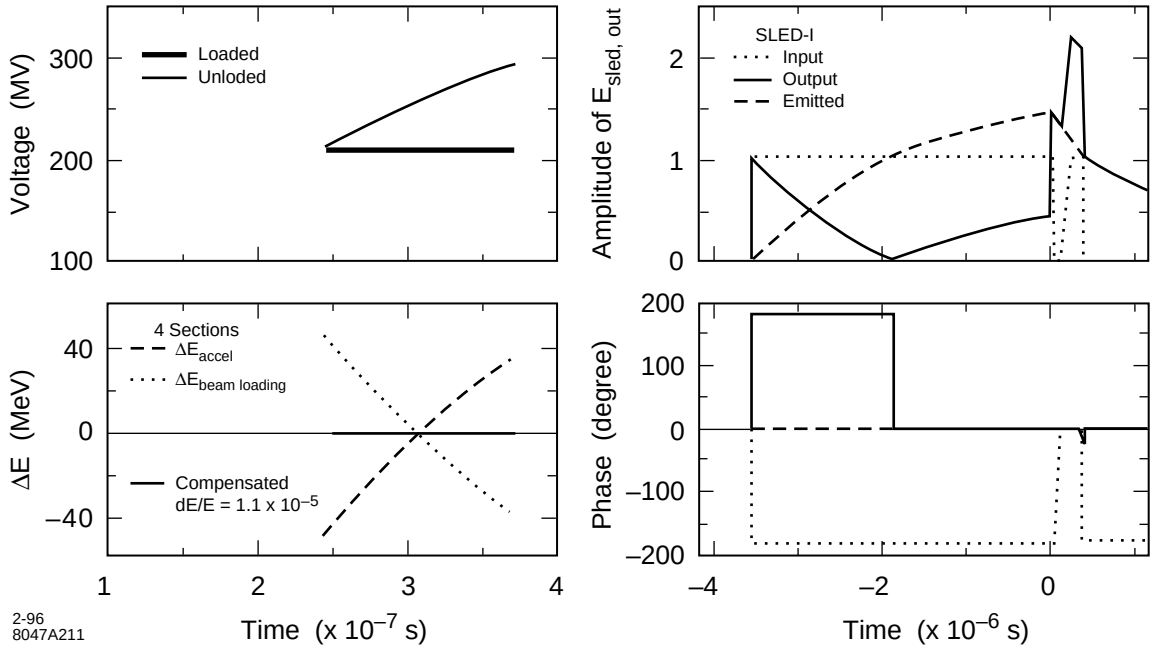


Figure 6-5. *Beam-loading compensation for a 2.2-A beam current in S-band linac by using amplitude modulation of the SLED-I wave form.*

band accelerator structure, the beam-loading voltage can be over 50% of the acceleration. The ΔF compensation is necessary for compensating such a high beam loading.

Since the linac is heavily loaded, short accelerator length is desirable, which can cut the power requirement for the klystrons as well as the cost of building the wrap-round solenoid. However, the length of the accelerator is limited by the maximum average gradient of the structure. Our criterion is that the average gradient in the L-band linac may not exceed 25 MV/m. In our design, we use two 5-m structures for acceleration and two 3-m structures for ΔF compensation. Two compensation sections operate at 1428 ± 1.428 MHz respectively to reduce the single-bunch effect. The filling time of the acceleration and compensation structures as simulated are 0.463 μs and 0.422 μs respectively. The compensation sections are required to have certain acceleration in order to reduce the gradient requirement for the acceleration sections. Each of the acceleration and compensation sections will be driven by two L75 klystrons and one SLED-I cavity. A schematic drawing of the e^+ capture linac is shown in Figure 6-6.

The acceleration klystrons will operate at 62.5-MW useful power while the compensation klystrons will operate at 47-MW useful power. This produces an average gradient for both the acceleration and compensation sections of about 25 MV/m, and a total energy gain of about 240 MeV. The center bunch runs 45° off-crest in the compensation sections, which provides a 53-MeV acceleration in each of the compensation sections. With the ΔF compensation, the beam-loading energy spread can be reduced to 0.82%, which is only a small fraction of energy spread in a single bunch. In Figure 6-12 are shown the energy spectrum of the beam-loading compensation and the SLED-I wave form for the capture linac.

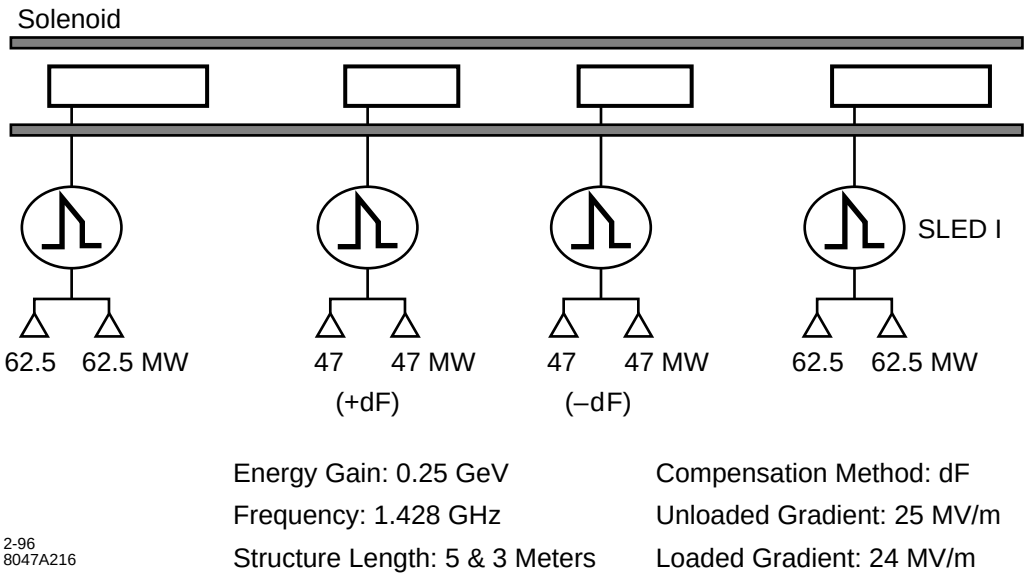


Figure 6-6. *The e^+ capture linac.*

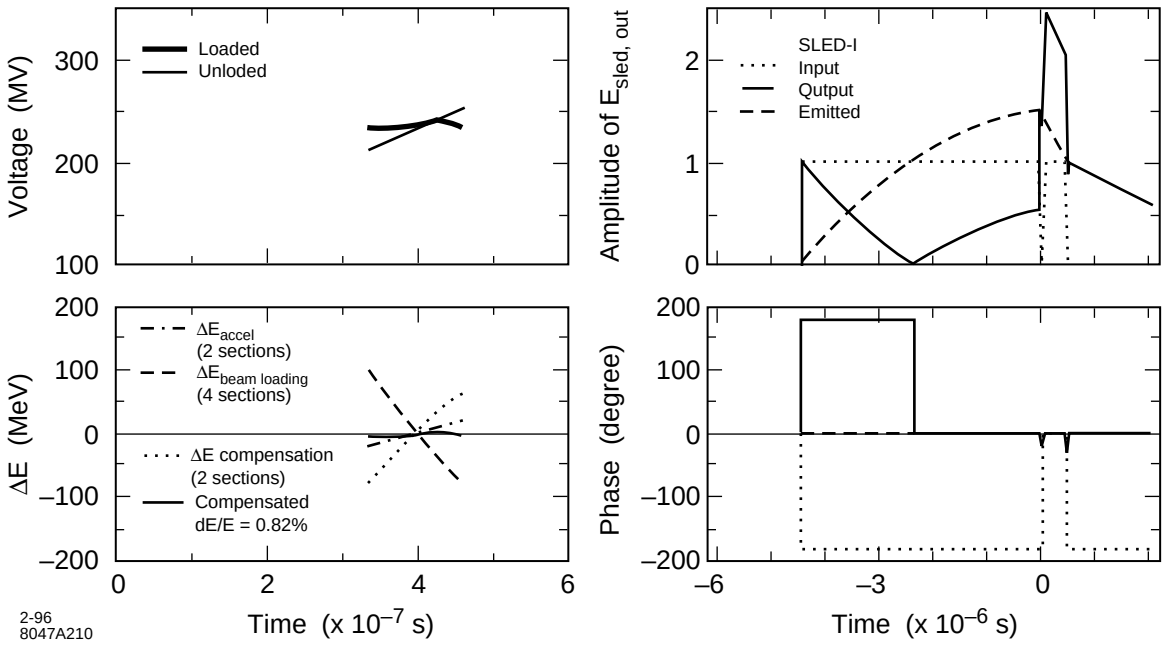


Figure 6-7. *Beam-loading compensation for the e^+ capture linac.*

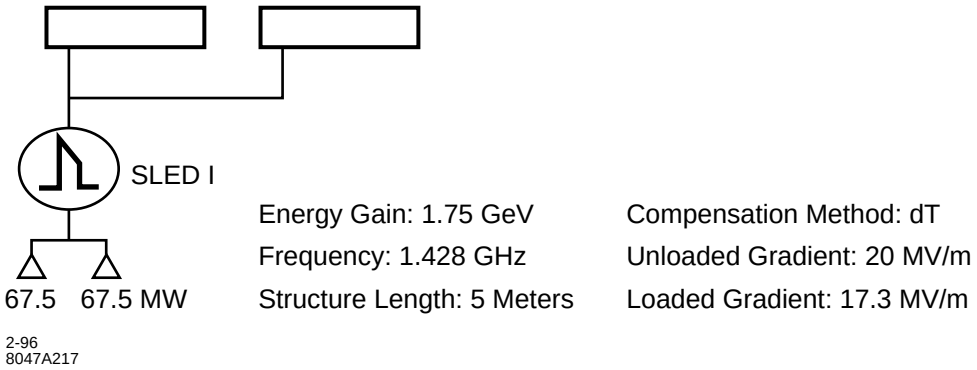


Figure 6-8. *An accelerator module for the e^+ booster linac.*

6.3.3 The e^+ Booster Linac

The e^+ booster linac accelerates the positron beam up to 2 GeV. The acceleration structure will be a detuned L-band structure ($F = 1.428$ GHz). The beam current in the linac is 2.75 A. The beam-loading compensation method is ΔT.

While the beam loading in the L-band structure is lower than in the S-band structure, beam-loading compensation still prefers a short filling time. The filling time of the L-band structure is chosen to be 0.675 μs. The beam loading is about 14% of the acceleration for a beam current of 2.75 A. Amplitude modulation of the SLED-I output is required to increase the derivative of the acceleration voltage. Two L75 klystrons will be used to drive one SLED-I cavity, which in turn drives two accelerator structures. An accelerator module is shown in Figure 6-8. The SLED-I wave form for the L-band accelerator is similar to Figure 6-5. The energy spread after beam-loading compensation is less than 5×10^{-5}. The loaded gradient of the structure is 17.3 MV/m.

6.3.4 L-band (e^-, e^+) Bunch Compressors

After the damping rings, the bunch length of the (e^+, e^-) beams is compressed. Two L-band linacs are used to produce the linear energy spread required for bunch length compression. The beam current is 1.6 A. The bunches run in phase quadrature with the rf crest phase. The compression voltage, which is the unloaded zero phase acceleration, is required to be 136 MeV and with less than 5% variation along the bunch train. The 5-m L-band detuned structures, same as the one used in the e^+ booster linac, will be used. Each compressor linac only needs two of such acceleration structures. Two L75 klystrons running at 35-MW useful power and one SLED-I cavity will be used to drive the whole L-band compressor linac. A schematic drawing is shown in Figure 6-9.

The beam-loading voltage is 90° off the rf voltage. Simple ΔT compensation method cannot be used for beam-loading compensation. A scheme that combines the ΔT method and phase modulation was studied, and shown to be effective for beam-loading compensation. The ΔT-amplitude modulation will be used for obtaining a uniform compression along the bunch train. A phase offset to the SLED-I output during the beam time is introduced to compensate the beam-loading voltage, which is 90° off-crest. The phase offset provides an in-phase (actually 180°-out-of-phase) component of rf voltage that cancels the beam-loading voltage.

Figure 6-10 shows the energy spectrum of the compressor linac and the SLED-I input and output wave forms. The amplitude and phase modulation of the SLED-I input can be obtained by properly phase-modulating the two klystrons.

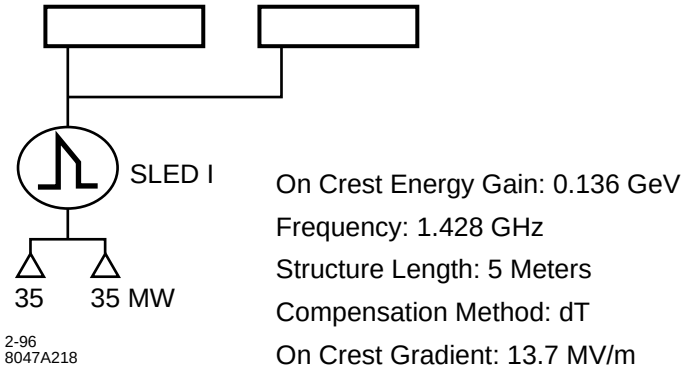


Figure 6-9. *An accelerator module for the L-band compressor linacs.*

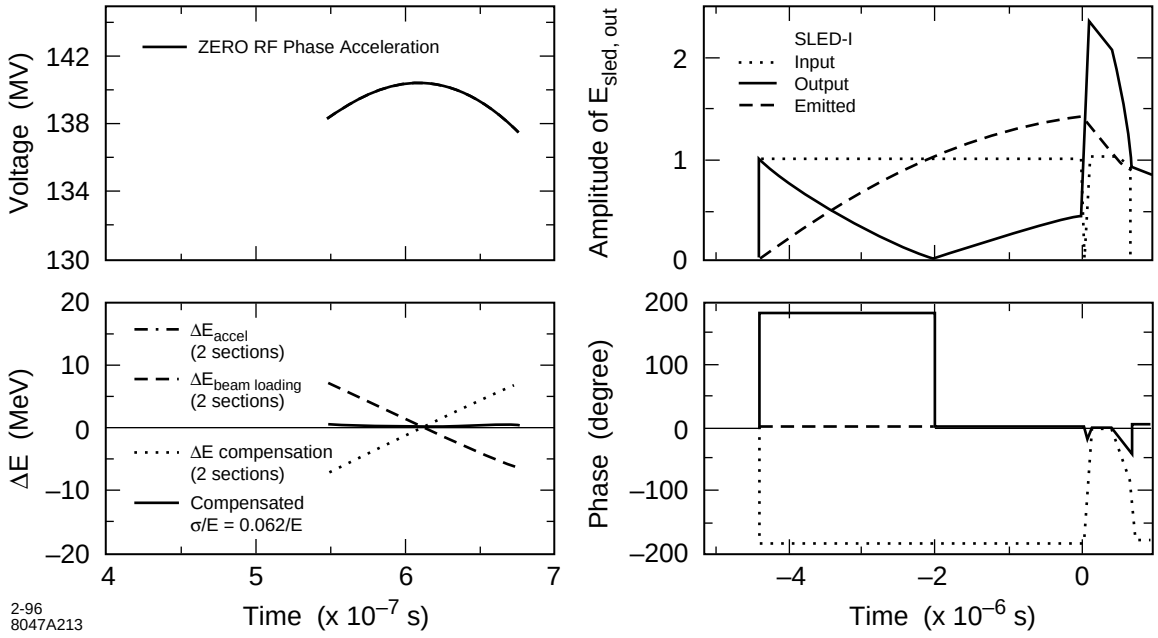


Figure 6-10. *Beam-loading compensation for the L-band compressor linacs.*

The amplitude modulation is obtained by modulating the phases of the two klystrons in opposite directions. Whereas, the phase offset is obtained by phase modulating the two klystrons in the same direction. The beam loading in these linacs can be reduced to $6\times10^{-2}/E_0$. The compression voltage deviates less than 3% along the bunch train.

6.3.5 S-band (e^-, e^+) Bunch Compressors

Before entering the main X-band linacs, the (e^+, e^-) beams are further compressed in bunch length in each of the respective accelerators. S-band linacs are used to produce the linear energy spread required for bunch length

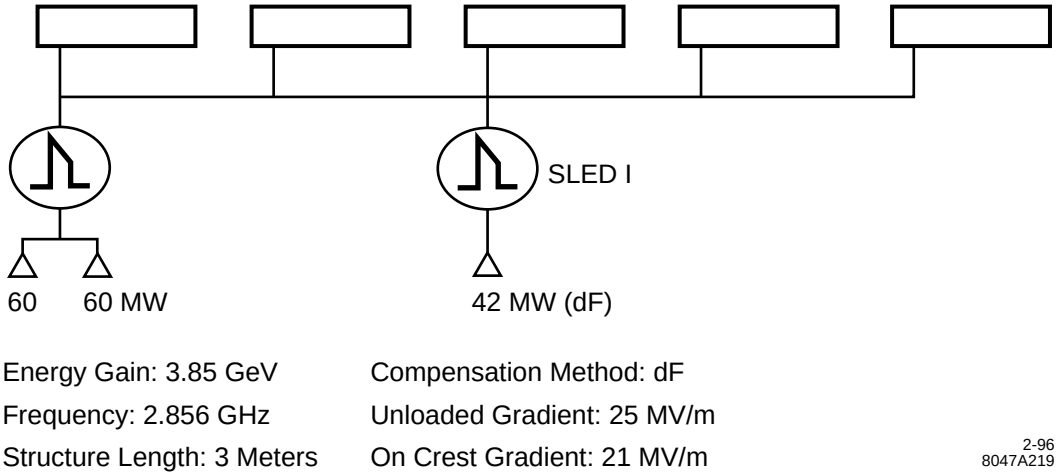


Figure 6-11. *An accelerator module for the S-band compressor linac. The four regular sections are used to generate a linear correlated energy spread within a bunch. One ΔF section is used to compensate the beam loading.*

compression. The beam current in these linacs is 1.5 A. The compression voltage, or the unloaded zero phase acceleration, is required to be 3850 MeV, and the deviation of the compression along the bunch train should be less than 5%. The beam loading in these linacs is relatively high. The ΔF sections will be used for beam-loading compensation, while the ΔT scheme will be used in the regular sections to obtain a uniform compression voltage. Both the ΔF and ΔT sections will be 3-m S-band DDS structures.

Two L65 klystrons and one SLED-I cavity will be used to drive four regular acceleration sections. Amplitude modulation of SLED-I output is needed to generate a uniform compression. The SLED-I wave form for the regular sections is shown in Figure 6-12. One L65 klystron and a SLED-I cavity will be used to drive one ΔF compensation section. The compensation klystron only needs to run at 42-MW useful power for an optimal beam-loading compensation. No amplitude modulation is needed for the compensation SLED-I. The ΔF sections operate alternatively at 2856 ± 1.1424 MHz through the linac to reduce the single bunch effect. A schematic drawing of an accelerator module as described above is shown in Figure 6-11.

In Figure 6-12 are shown the compression voltage and the beam-loading compensation results for an accelerator module. The optimal rf phase for the ΔF sections is 44° off-crest. At this rf phase, the ΔF sections also provide some acceleration that nearly cancels the deceleration (half the beam-loading voltage) due to the beam loading. With ΔF compensation, the beam-loading energy spread can be reduced to 0.62×10^{-3} at an energy of 10 GeV. The average gradient in the structure is about 22 MV/m.

An alternative ΔT compensation method, similar to the one used in the L-band compressors, was also investigated. It has been shown that the ΔT compensation can give better energy spectrum compared to the ΔF compensation. However, the ΔF method is more straightforward and operationally more convenient.

6.3.6 The (e^-, e^+) energy compressors

There are energy compressors associated with the electron and positron beams before they enter the damping rings. The compressor will be S-band for the electron beam and L-band for the positron beam. The compression voltage requirements for the S-band and L-band compressors are 42 MeV and 80 MeV respectively. The deviation of the

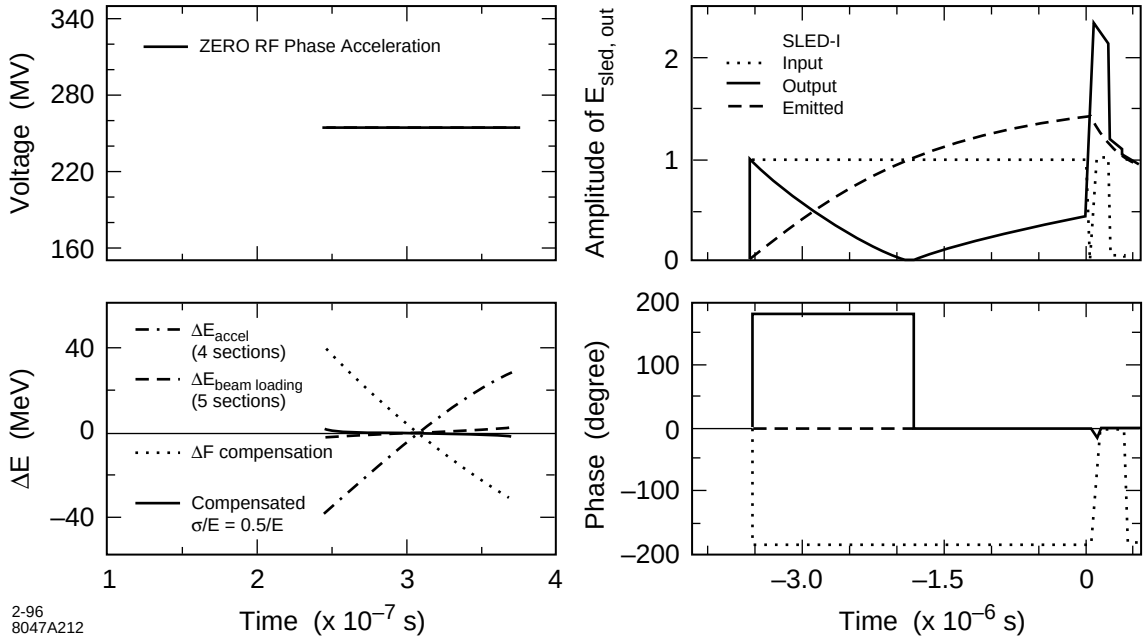


Figure 6-12. *Beam-loading compensation for the S-band compressor linacs.*

compression voltage along the bunch train is required to be less than 10%, and the compensated beam-loading energy spread less than 0.1% in full width. Both compressors will use ΔT and phase modulation method to obtain a uniform compression voltage and beam-loading compensation.

S-band e^- energy compressor. The beam current in the electron energy compressor is 2.2 A. A single 2.5-m S-band accelerator section is used to generate the energy correlation needed for energy compression. The filling time of the structure is 0.371 μs. The compressor will be driven by one unsledded S65 klystron, with a 130-ns rising time in amplitude. The maximum phase offset for beam-loading compensation is 60°. The resultant full energy spread along the bunch train is 0.73×10^{-3}, and the deviation of the compression voltage is less than 5%.

L-band e^+ energy compressor. The beam current in the positron energy compressor is 2.75 A. The compressor will be composed of one 4-m L-band accelerator section. The compressor will be driven by two L75 klystrons with SLED-I pulse compression. The filling time of the structure is 0.537 μs. The maximum phase offset for beam-loading compensation is 80°. With ΔT and phase modulation, the full energy spread can be reduced to less than 0.49×10^{-3} and the compression voltage is within 5% deviation.

6.3.7 Beam-loading Compensation in the Spare Modules

Each of the linacs discussed in this chapter is designed to have about 10% spare accelerating modules. The beam induces the same transient beam-loading voltage in these spare accelerator sections as in a powered section. In order to maintain localized compensation, we will drive one of the two klystrons which drives a module with a switched drive signal at $F_0 \pm \Delta F$. With appropriate choice of ΔF, 1.428 MHz for L-band and 1.1424 MHz for S-band, this drive frequency will produce enough voltage in the standard accelerating structures when powered by one klystron to compensate for the beam loading in the same structure. Since it is very unlikely that both klystrons in a module

	e^+ drive e^- booster	(e^-, e^+) prelinac	e^+ capture	e^+ booster	(e^-, e^+) L-band compressor		(e^-, e^+) S-band compressor	
					Bunch	Energy	Bunch	Energy
Frequency (GHz)	2.856	2.856	1.428	1.428	1.428	1.428	2.856	2.856
E Gain (GeV)	3-6, 2	8	0.250	1.75	0.136	0.080	3.85	0.042
Current (A)	2.2	1.5	14	2.75	1.6	2.75	1.5	2.2
$L_{\rm struct}$ (m)	3	3	5	5	5	4	3	2.5
T_f (μs)	0.371	0.371	0.464	0.675	0.675	0.537	0.371	0.371
N Struct/module	4	4	2	2	2	1	4	1
$P_{\rm klystron}$ (MW)	2×60/4	2×60/4	2×62.5/1	2×67.5/2	2×35/2	2×35/1	2×60/4	56[b]/1
Klystron Pulse (μs)	4	4	5	5	5	5	4	0.5
Average Gradient (MV/m)	17.3	21.2	24.2	17.3	13.7	20.0	21.1	16.8
(MV/module)	208	254	242	173	137	80.0	253	42
Compensation Method	ΔT	ΔT	$\Delta T/\Delta F$	ΔT	ΔT	ΔT	$\Delta T/\Delta F$	ΔT
$L_{\rm struct}$ (m)			3				3/3	
T_f (μs)			0.422				0.371	
N Struct/module			2				1	
$P_{\rm klystron}$ (MW)			2×47/1				42/1	
Klystron Pulse (μs)			5				4	
$\sigma/(\Delta E(\text{ MeV}))$	2.4×10^{-5}	1.1×10^{-5}	8×10^{-3}	5×10^{-5}	$\frac{0.062}{E_0}$	$\frac{0.49^a}{E_0}$	$\frac{6.2}{E_0}$	$\frac{0.73^a}{E_0}$

[a] Maximum energy spread.
[b] Not SLEDed.

Table 6-3. *A summary of low-frequency linac beam-loading compensation.*

will fail at the same time, this technique will usually maintain localized beam-loading compensation even with failed klystrons. Since ΔF is greater than the bandwidth of the SLED-I cavities, they will behave as if they are detuned.

6.3.8 Summary of the Beam-loading Compensation

A summary of the beam-loading compensation is shown in Table 6-3. In the compensation sections, the frequency offset ΔF is 1.428 MHz for the e^+ capture linac and is 1.1424 MHz for the S-band compressor linac. The notation, such as 2×60/4, in the klystron power specification means to use two klystrons running at 60-MW useful power to drive one SLED-I cavity, and the SLED-I cavity drives four accelerator sections.

Beam voltage	350 kV
Beam current	400 A
Rf pulsewidth @ rep rate	4 μs @ 180 Hz
Cathode loading	2:1 (8 A/cm2 max)
Cathode convergence	18:1 (3.525" dia.)
Rf output power	65 MW
Saturated gain	$\sim$55 dB
Efficiency	$\geq$40%
Operating frequency	2856 $\pm$1.4 MHz
3 dB Bandwidth	15 MHz
Solenoidal focusing field	1400 Gauss

Table 6-4. *Design parameters for a 65-MW S-band klystron*

Beam voltage	388 kV
Beam current	483 A
Rf Pulsewidth @ rep rate	5 to 6 μs @ 180 Hz
Cathode loading	2:1 (6 A/cm2 max)
Cathode convergence	11:1 (4" dia.)
Rf output power	75 MW
Saturated gain	$\sim$55 dB
Efficiency	$\geq$40%
Operating frequency	1428 $\pm$1.4 MHz
3 dB Bandwidth	8 MHz
Solenoidal focusing field	980 Gauss

Table 6-5. *Design parameters for a 75-MW L-band klystron.*

6.4 Klystrons

6.4.1 S-band S65 klystron

Klystrons operating at 2856 MHz with 65 MW of available rf output power will be required for the NLC. Such klystrons already exist at SLAC as the 5045 klystron tubes. The 5045 klystrons will need slight modifications to improve pulsewidth (from 3.5 μs to 4 μs) and average power handling capability. Table 6-4 summarizes the parameters for NLC S-band klystrons.

6.4.2 L-band L75 klystron

Klystrons operating at 1428 MHz with 75 MW of available rf output power will also be required for the NLC. Table 6-5 summarizes the parameters for NLC L-band klystrons. A conventional klystron as described above would be approximately 123-in long and weigh approximately 900 lb. Such a klystron can be built and operated using technology currently available and proven. Indeed, 150-MW klystrons operating at twice the frequency have already been built at SLAC. However, interesting designs for multiple-beam klystrons have been studied at SLAC which indicate a substantial reduction in beam voltage without sacrificing efficiency. For example, a 10-beam klystron operating at only 273 kV would have 114 A in each beam which leads to a theoretical efficiency of 64% (giving 200 MW).

6.5 Dipole Wakefields

The S-band accelerator sections will be detuned by about 6% in a Gaussian density distribution to diminish the effect of dipole wakefields. This will cause the structure to be what might inarticulately be called "over-constant-gradient", *i.e.*, the gradient rises from the input to the output. Damping will be desirable. A manifold damped structure similar to the X-band damped detuned structure is being studies and it appears to have dipole wakefields greatly reduced from the structure without damping. Figure 6-13 shows a MAFIA [Mafia Collab.] representation of 1/4 of a cell in the proposed three-m S-band structure. In order to enhance the damping, the S-band structure was modified from a simple scaling from the X-band design: a narrow (5-mm-wide) slot extends from each manifold into each disk, extending about half the distance from the manifold to the iris in the disk. This greatly enhances the coupling to the $TM11$-like mode which is predominant in the lowest dipole band for the parameters chosen for the S-band structure. The ringing time of microwave structure scales as $\lambda^{3/2}$, and consequently the natural scaling for accelerator structure length is $\lambda^{3/2}$. However, this scaling from the NLC X-band structure would give a cumbersome 14-m-long S-band accelerator section. Instead, a 3-m length has been chosen tentatively. With the filling time chosen for ΔT compensation, we get an average group velocity which is a factor of 2.2 smaller for the S-band structure. This change makes the lowest passband dipole predominantly $TM11$-like, even at the zero phase advance of the passband. Slots in the disks were added to improve the coupling. Narrow slots have negligible effect on the fundamental mode shunt impedance since they are parallel to the current for that mode.

A preliminary study indicates that just detuning will be adequate for the positron L-band linac because of the strong focusing produced by the many wrap-around quads and because of the very low dipole impedance of the large aperture L-band linac.

6.5.1 Detuned Structure

The most important wakefield in the S-band structure is the first passband dipole wakefields. One finds that by detuning the dipole frequency in a Gaussian density distribution, the wakefield can be dramatically suppressed. The NLC S-band structure will be detuned by about 6% in full width with a three-sigma cutoff. The detuning can be obtained by properly varying the iris aperture and the cell radius while keeping the fundamental frequency unchanged. Figure 6-14 shows a preliminary estimation of the dipole spectrum for the S-band structure. Each frequency shown in the spectrum corresponds to one cell in the structure.

The dipole wakefield estimated by using a equivalent circuit model [Thompson] for a S-band Gaussian detuned structure is shown in Figure 6-15. It is shown that the detuned structure can strongly suppress the wakefield. However,

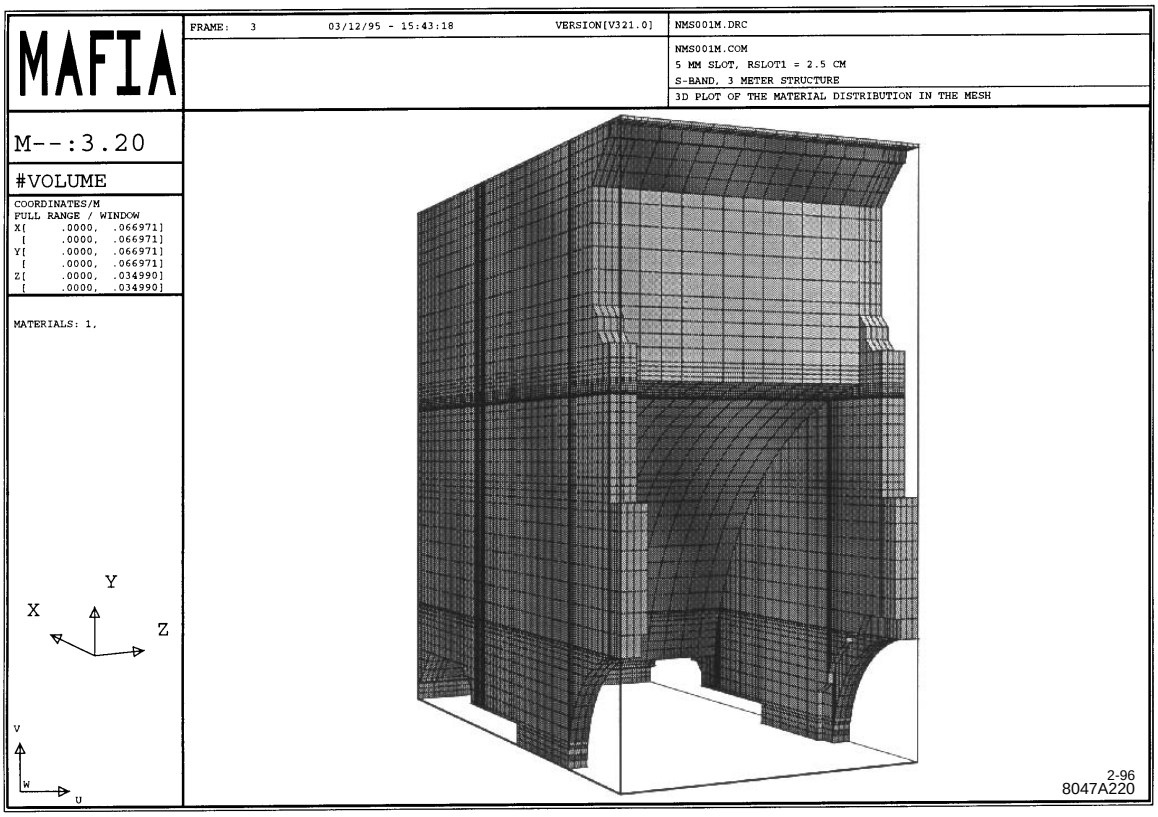


Figure 6-13. *The 1/4 cell of the S-band structure.*

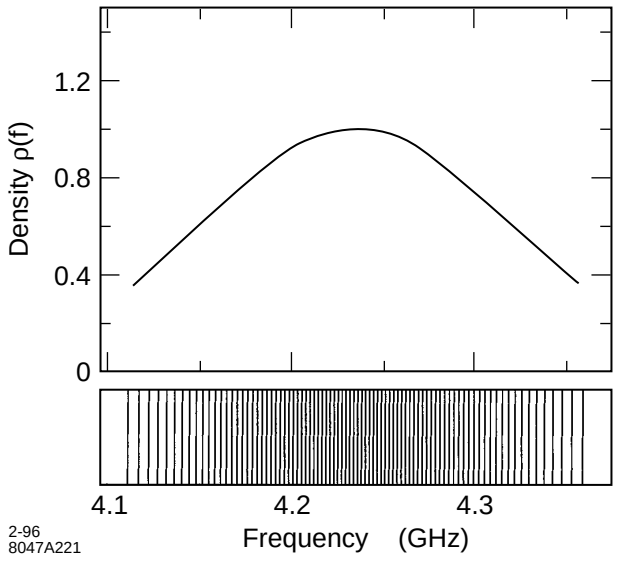


Figure 6-14. *Gaussian detuned dipole spectrum for the S-band structure.*

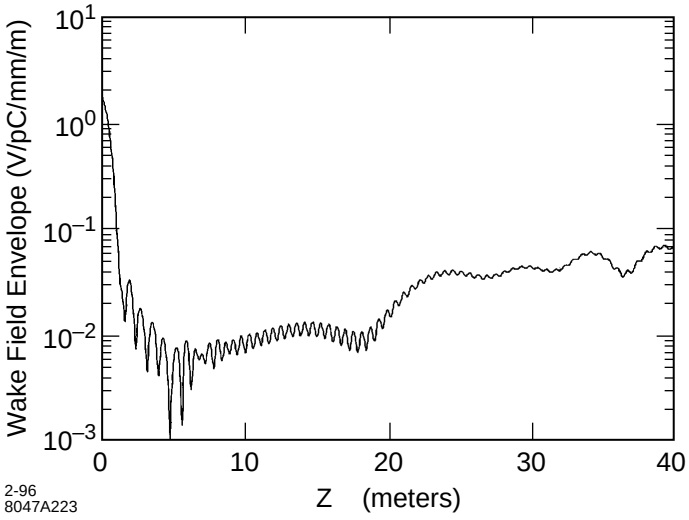


Figure 6-15. *Wakefields for a detuned S-band structure.*

since the dipole spectrum is a discrete and truncated Gaussian, the decoherence among the modes only helps to reduce the short-term wakefields. At a longer time scale, but shorter than the pulse length, the modes re-cohere, which results in strong long-term wakefields. Certain damping scheme is needed to minimize the long-term wakefields.

6.5.2 Damped Detuned Structure

A damping scheme similar to the X-band damping manifold will be used to damp the long-term dipole wakefield. In addition to the damping manifold, a narrow slot extends from each manifold into each disk is added to enhance the coupling, see Figure 6-16. The manifold runs parallel to the structure, and is 1.39-cm wide and 2.12-cm high. The wall between the cell and the manifold is cut open with the full manifold width. The narrow slot, 0.5 cm in full width, cuts about halfway into the disk. The full height of the slot-manifold is about 3.834 cm.

The coupling between the slot-manifold and the cell is estimated by use of the 3-D MAFIA and code MDAMP [Kroll 1994] derived from the equivalent circuit model. For a Gaussian detuned structure, the iris radii a and the cell radii b are tapered. With the additional slot-manifold, the structure is three-dimensional. It is not practical to model the whole structure by use of the 3-D MAFIA. However, since the tapering is gradual and smooth, we can use a cell-to-cell approach, *i.e.,* to model the structure cell by cell. For each cell, the two irises at the two ends of the cell are assumed to have the same radius, which takes the average of the two actual radii of the irises. Periodic boundary conditions are applied to the two ends, which is equivalent to model an infinitely long constant impedance structure. The 3D MAFIA is used to model a number of selected cells along the structure. Interpolation method is used to obtain the parameters for the rest of the cells.

In Figure 6-17 is shown a MAFIA result of the dispersion relation for a cell with $a = 1.150$ cm and $b = 4.071$ cm. At zero phase advance, the lower frequency mode is the manifold mode and the upper mode is the dipole mode of the cell. As the phase advance increases, the two modes come closer to each other. The modes, however, do not cross. At a particular phase advance, 120° for this case, the field patterns of these modes switch and the dispersion curves start to separate further apart. The frequency separation of the modes at the avoided crossing point indicates the coupling

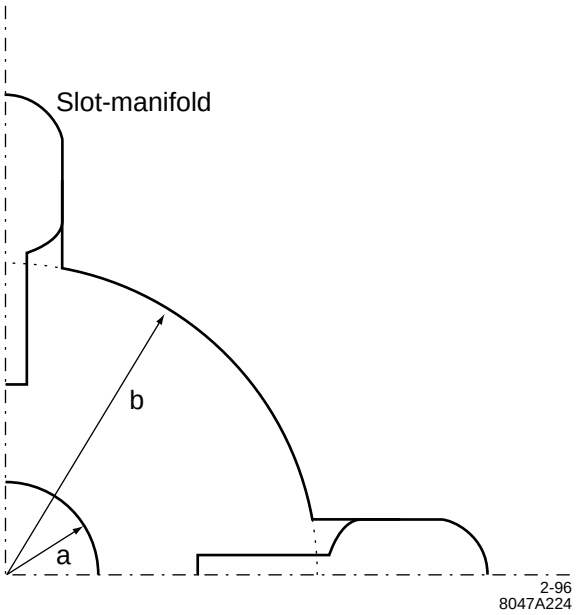


Figure 6-16. *Slot-manifold damping structure for the S-band accelerator.*

strength, with large separation corresponding to large coupling. The simulation shown in Figure 6-17 has about 6% of coupling.

The coupling is insensitive to the width of the slot. In Figure 6-18 are shown the couplings for different slot widths. As one can see that the coupling only increases by 0.5%, from 5% to 5.5%, as the slot width increases from 3 mm to 5 mm. On the other hand, the coupling is sensitive to the depth of the slot. We calculated the couplings for different slot positions while keeping the whole slot-manifold geometry unchanged. Figure 6-19 shows the coupling as a function of the nearest distance of the slot to the axis. The coupling varies almost linearly to the distance and the slope is about 1% per millimeter.

The coupling not only depends on the depth of the slot, but also depends on the iris opening. The iris opening is tapered in the detuned S-band structure, with large opening in the front and smaller opening toward the end. The radial position of the slot will also be tapered to get similar coupling for all cells.

The slot-manifold also preserves the Q_0 and the shunt impedance (r) of the structure, as shown in Figure 6-18. The reason underlying this is the small wall loss of the fundamental mode induced by the slot. Even though the slot-manifold couples strongly to the $TM11$ dipole field, it has very weak coupling to the fundamental mode which is a $TM01$ mode. For the $TM01$ mode, the current in the disk is in the radial direction. The narrow radial slot in the disk has small perturbation to this current. The E and B fields of the fundamental mode are weak in the slot, which results in small wall loss and associated Q_0 and r loss.

The dipole wakefield, estimated by using equivalent circuit model, for a damped-detuned structure with 6% cell-manifold coupling is shown in Figure 6-20. The long-term wakefield is strongly minimized by the damping structure. This indicates that the slot-manifold can provide the coupling needed for damping the wakefields.

It is worthwhile to mention that the equivalent circuit model used so far is the single-band model. Recently, a multi-band equivalent circuit model has been developed, which promises to give better accuracy for wakefield estimation.

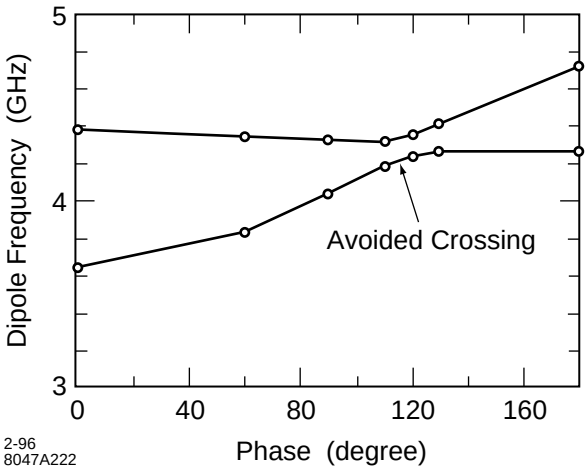


Figure 6-17. *MAFIA simulation of the dispersion curve of a slot-manifold coupled S-band accelerator cell. The avoided crossing indicates the strength of the coupling.*

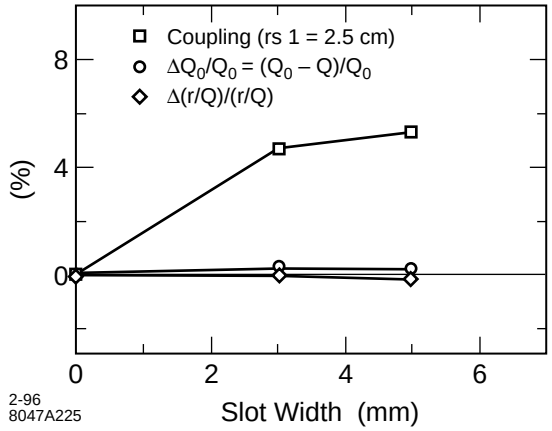


Figure 6-18. *Cell-slot-manifold coupling, $\Delta Q_0/Q_0$ and $\Delta(r/Q_0)/(r/Q_0)$ for different slot width.*

Nevertheless, the single-band model has been proven experimentally in the X-band damped-detuned structure measurements to give reasonably good results.

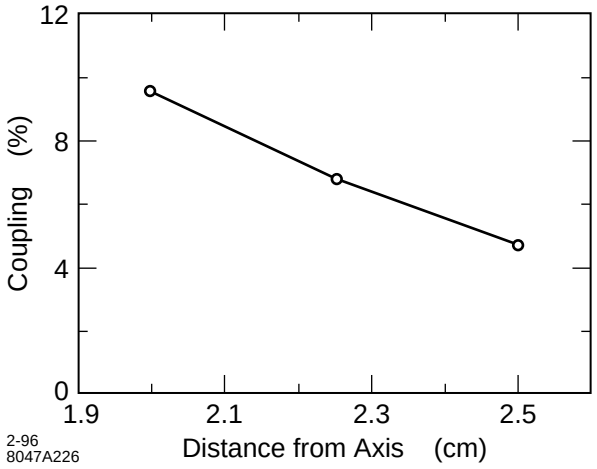


Figure 6-19. *Cell-slot-manifold coupling as a function of the distance of the slot from the axis.*

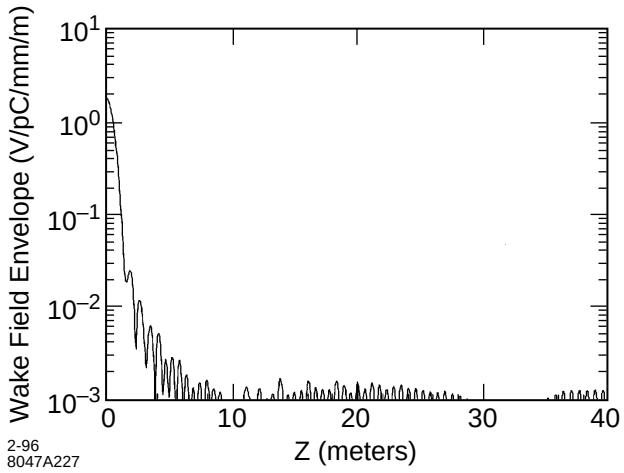


Figure 6-20. *Wakefields for a damped-detuned S-band structure.*

References

[Mafia Collab.] The Mafia Collaboration, MAFIA—The ECAD System.

[Thompson] K. Thompson, private communications.

[Kroll 1994] N. Kroll *et al.*, "Manifold Damping of the NLC Detuned Accelerating Structure", SLAC–PUB–6660 (1994), in Proc. of the Sixth Workshop on Advanced Accelerator Concepts, Lake Geneva, WI, June 12-18, 1994.

Contributors

- David Farkas
- Kwok Ko
- Zenghai Li
- Roger Miller
- Tor Raubenheimer
- Daryl Sprehn
- Huan Tang
- Kathy Thompson
- Dian Yeremian

7

Main Linacs: Design and Dynamics

Contents

7.1 Introduction

The main linacs will accelerate electrons and positrons from an energy of 10 GeV to 250 GeV in the initial stage of running, and from 10 GeV to 500 GeV after upgrades to the rf system. For initial (upgraded) operation, trains of 90 bunches with 0.75×10^{10} (1.1×10^{10}) particles per bunch, and a 1.4-ns spacing between bunches, will be accelerated at a repetition rate of 180 (120) Hz.

There are two major design issues for the main linacs: the efficient acceleration of the beams and the preservation of the small beam emittances that are required to achieve the desired collision luminosity. The acceleration of the beams will be accomplished using X-band rf technology which is described in detail in Chapter 8. In this chapter, we discuss the design and operation of the linacs with the emphasis on the underlying beam dynamics that have driven the choices made in these regards.

The NLC linac design builds on the experience that has been gained over the last ten years from the operation of the SLC, in particular, in the development of beam-based alignment, tuning, and feedback techniques. Information from beam measurements are now routinely used in the SLC to align magnets, tune dispersion and chromaticity, compensate for accelerator structure misalignments, and to stabilize beam energy and orbits, including the beam-beam offset at the collision point. Such methods will also be crucial for the NLC operation as will become clear in this chapter. Their success in the SLC has given us confidence they will work in the NLC as well.

The NLC design also draws on the SLC operation experience in such areas as the transport of flat beams, the use of BNS damping, and linac energy management. From this, we have identified many things that need to be improved for the NLC. For example, drifts in the energy profile along the SLC linac have made tuning difficult, so in the NLC we are considering the use of beam-derived signals to keep the rf phases synchronous with the beams.

In areas where we cannot directly verify NLC operational issues with the SLC, we have built test facilities. These include ASSET for X-band structure wakefield measurements, ASTA and NLCTA for X-band accelerator technology testing, and the FFTB for final focus evaluation, although some of its features, such as quad movers, test linac concepts as well. Of these facilities, the NLCTA will provide the most direct test of the basic linac design in its ability to accelerate multibunch beams and to control the effects of beam loading. These facilities are discussed further in Chapters 8 and 11.

In this chapter, we begin with a discussion of parameters and specifications for the main linacs, followed by an overview (Section 7.3) of their design and layout. Many of the issues touched upon here (*e.g.,* alignment and stability, design of individual components, and instrumentation) will be treated in greater detail later in the chapter and elsewhere in this report, but Section 7.3 will provide a general orientation to the main issues and tradeoffs that have led to the present design.

Following the section on design and layout, we begin a detailed discussion of linac dynamics issues most closely related to emittance preservation. Of major importance are issues associated with the wakefields generated in the X-band accelerator structures by the beams. Although these structures will provide a high accelerating gradient, they will also heavily load the bunch train and will generate strong short- and long-range transverse wakefields when the bunches travel off-axis though them. Of particular concern is the long-range wakefield, which if not heavily suppressed, will resonately amplify the betatron motion of the bunch train by many orders of magnitude. Much work to date has been devoted to solving this problem, the outcome of which is an accelerating structure that incorporates both detuning and damping of the wakefield. Most of the details of the structure design are given in the next chapter: here, at the beginning of Section 7.4, we review how the wakefields, both longitudinal and transverse, were computed for this structure.

With this foundation, we next consider the dynamics issues associated with single-bunch and multibunch phase space, both longitudinal and transverse. These include operational issues such as the optimal single-bunch energy spread profile for BNS damping and the methodology used to compensate long-range beam loading. Most of the discussion, however, is devoted to the tolerances required to prevent significant emittance dilution from such sources as misalignments and jitter (the latter can be thought of as producing a dilution of the time-averaged emittance). In particular, we consider only bare tolerances in that we assume no benefit from a beam-based types of correction, including feedback.

In the next section, 7.5, we consider issues of alignment, stability, and feedback, with emphasis on the beam-based methods that will be required in cases where the bare tolerances cannot be met. Where data is available on error sources, we compare their size to the tolerances presented in Section 7.4 We will see, for example, that beam jitter due to ground motion should not be a major problem for the linacs.

Maintaining the quadrupole and accelerator structure alignment, however, will be a key aspect of NLC operation. Quad offsets will lead to dispersive emittance growth while the structure offsets will lead to wakefield related emittance growth. In both cases, beam-based methods will be required as the bare tolerances would be impossible to meet otherwise. The proposed methods will require accurate measurements the beam position. How this will be achieved is discussed later in this chapter.

The quad and structure alignment algorithms are examples of local types of correction in that they correct the problem at the source. Also discussed are non-local correction methods where trajectory bumps, kickers, etc., are used to cancel the dilutions after they have accumulated. We generally consider them fall-back methods since they are sensitive to variations in beam transport properties between the sources of the dilution and the correction point.

Another important feature of the NLC operation will be the use of feedback systems to maintain stability on timescales of seconds and longer. They will be used to stabilize both bunch energies and trajectories at various locations along the linac. Because of their importance, we review the performance of the SLC feedback systems.

The remaining sections of the chapter contain information on beam transport simulations, tolerances, operations, linac components, and instrumentation. Of particular interest, due to the very large number of components in the main linacs, is to quantify the reliability of the various systems; studies of this are just beginning and will be discussed in future revisions of this chapter.

7.2 Parameters and Specifications

7.2.1 System Boundaries

The main linac systems boundaries are from the exit of the bunch compressor complex to the entrance of the collimation system that leads to the final focus. The main linacs consist basically of X-band acceleration sections, interspersed with quadrupoles to focus the beam, along with the necessary diagnostics and supporting systems. The rf system feeding the main linac accelerating structures and the details of the design of those structures are discussed in the next chapter.

On each rf pulse, a train of up to 90 bunches enters the main linacs from the bunch compressor. The two main tasks of the main linacs are to accelerate the bunches to high energy and to minimize the growth of the beam emittance. Even with care taken to minimize emittance growth and development of transverse “tails” on the beam, it will be necessary to clip the tails to minimize backgrounds in the detector; the collimation systems needed to do this are discussed in Chapter 9.

	NLC-I			NLC-II		
	a	b	c	a	b	c
N (10^{10})	0.65	0.75	0.85	0.95	1.10	1.25
Bunches/Train		90			90	
Repetition Rate (hz)		180			120	
σ_z(μm)	100	125	150	125	150	150
Unloaded Gradient (MV/m)		50			85	
Multibunch Loading (%)	25.5	29.4	33.34	22.0	25.5	28.9
Multibunch Loading (MV/m)	-12.8	-14.7	-16.7	-18.7	-21.5	-24.6
Single-bunch Loading[a] (MV/m)	-0.2	-0.3	-0.4	-0.4	-0.5	-0.6
Average $\phi^b_{\rm rf}$ (°)	-15	-15.5	-16	-7.5	-8.2	-11.2
BNS Overhead (%)	3	3	3	3.5	3.5	3.5
Feedback Overhead (%)	2	2	2	2	2	2
Repair Margin (%)	3	3	3	3	3	3
$L_{\rm acc}$ (m)		8150			8834	
Number Structure		4528			4908	
$L_{\rm total}$ (m)		8807			9550	
$E^c_{\rm max}$ (GeV)	266.7	250	232.2	529	500	468

[a] Single-bunch loading only includes HOM contribution.

[b] Average rf phase for 0.8% FWHM energy spread.

[c] Including 10-GeV initial energy.

Table 7-1. *NLC linac parameters.*

7.2.2 Main Linac Parameters and Requirements

Parameters of the main linacs are listed in Table 7-1 for NLC-I, which has a 500-GeV center-of-mass energy, and NLC-II, which has a 1-TeV center-of-mass energy. The length of each of the two main linacs in NLC-I is 8.8 km. This will increase only slightly in the upgrade to NLC-II, with the bulk of the energy gain coming from an increase in the unloaded gradient from 50 to 85 MeV/m. This increase in gradient will be performed adiabatically by upgrading and doubling the number of klystrons which supply the rf power. It will be accompanied by an increase in beam charge which will leave the fractional beam loading roughly constant. However, the repetition rate will be decreased to keep total site power consumption from becoming too large. Since the total linac length required for NLC-II is not much longer than that for NLC-I, we will probably construct the full NLC-II linac length initially and use the additional length in NLC-I to attain beam energies slightly higher than 250 GeV.

For both NLC-I and NLC-II, three different parameter sets are listed. These correspond to the same rf power and linac length, but assume different bunch charges and bunch lengths. This leads to different loading through the linacs and therefore different final beam energies. The total beam energy is determined from the accelerating gradient, $Grad_{\rm rf}$, the average beam phase, $\phi_{\rm rf}$, and the single and multibunch beam loading:

$$\Delta E = L_{\rm acc} Grad_{\rm rf} \cos\phi_{\rm rf} (\%fdbk)(\%BNS)(\%off) - L_{\rm acc} q (k_{||\,\rm multi} + k_{||\,\rm single}) \quad . \tag{7.1}$$

In addition, there must be sufficient energy overhead for the energy feedback systems, the phase profile required for BNS damping, and the energy loss from klystron stations that have to be repaired. The value assumed for the klystron repair margin is roughly 50% larger than that needed in the SLC linac. The overhead assumed for the energy feedback is similar to that used in the SLC. Finally, the additional energy needed for BNS damping arises because the linac will

	NLC-I			NLC-II		
	a	b	c	a	b	c
N (10^{10})	0.65	0.75	0.85	0.95	1.10	1.25
σ_z (μm)	100	125	150	125	150	150
$\gamma\epsilon_{x\,\mathrm{inj}}$($10^{-6}$ m-rad)		3.6			3.6	
$\gamma\epsilon_{y\,\mathrm{inj}}$($10^{-8}$ m-rad)		4			4	
$\sigma_{\epsilon\,\mathrm{inj}}$(%)		<1.5			<1.5	
$\gamma\epsilon_{x\,\mathrm{ext}}$($10^{-6}$ m-rad)		4			4	
$\gamma\epsilon_{y\,\mathrm{ext}}$($10^{-8}$ m-rad)	7	9	11	9	11	13
$\Delta E/E_{FWHM\,\mathrm{ext}}$ single bunch (%)		0.8			0.8	
$\Delta E/E_{FWHM\,\mathrm{ext}}$ train (%)		1.0			1.0	

Table 7-2. *NLC linac emittance requirements*

not be operated at a constant rf phase. Instead, the rf phases are set in the first part of the linac to produce the desired BNS bunch energy spread; in the remainder of the linac, they are set to achieve a final energy spread that is within the final-focus bandwidth. This will be discussed further in Subsection 7.4.3.

The design-normalized emittance of the bunches entering the main linacs is $\gamma\epsilon_x = 3.6 \times 10^{-6}$ m-rad, $\gamma\epsilon_y = 4.0 \times 10^{-8}$ m-rad, and the energy is 10 GeV. It is the task of the main linacs to accelerate the bunches to ∼250 GeV (NLC-I) or ∼500 GeV (NLC-II), while preserving the low transverse emittance and maintaining the small final beam energy spread; these requirements are listed in Table 7-2.

The three parameter sets listed for NLC-I and NLC-II reflect a tradeoff in bunch charge versus beam size where the luminosity is kept nearly constant. As discussed in the introduction to this document, this is the main tradeoff that remains in a linear collider design when all constraints are taken into account. Larger beam sizes are desirable since they loosen many of the tolerances. However, increasing the beam charge to compensate the loss in luminosity produces problems related to beam power, and generally leads to tighter wakefield-related tolerances. The three parameter sets each represent compromises among these competing effects: together they span the range of operating conditions that we consider reasonable.

In this chapter, we will concentrate the mid-range parameters, that is, sets NLC-Ib and NLC-IIb. Note that we have allowed for over 100% vertical emittance growth in the linacs in these cases. This lets us set fairly conservative tolerances on the accelerator structure alignment. However, we believe that we will do better than these tolerances as we gain more experience with X-band accelerator construction and operation.

7.3 System Design

7.3.1 Design Overview

Although the layout of the main linacs is fairly simple, an array of X-band accelerator structures interleaved with a FODO quadrupole lattice, one faces a number of challenges to accelerate low-emittance bunch trains without significantly degrading the beam phase space. The X-band structures that provide a high-acceleration gradient also heavily load the beam, so careful control of the rf pulse shape will be needed to achieve the small beam energy spread that is required at the end of the linac. The X-band structures also have the drawback that large transverse wakefields, both long-range (bunch-to-bunch) and short-range (intra-bunch), are generated when the bunch train passes off-axis through the structures. Although damping and detuning of the dipole modes of the structure will be used to suppress the long-range wakefield, and autophasing will be used to offset the short-range wakefield effect on betatron motion, the structures will still have to be aligned precisely to limit emittance growth from both short-range and long-range wakefield effects. The small bunch sizes also make the beam sensitive to other emittance growth mechanisms such as dispersion from quadrupole misalignments, focusing from the ions generated in the residual gas of the beam-line vacuum chambers, and dilution from beam trajectory jitter, like that caused by vibrations of the quadrupole magnets.

Most of the deleterious effects will be limited by design. For example, the quadrupole magnet power supplies will be chosen to be stable enough not to cause significant beam trajectory jitter. For some of the effects, however, we will rely mainly on measurements of the beam properties to control emittance growth. In particular, beam trajectory measurements will be used to adjust the positions of the quadrupoles and structures to minimize the emittance growth caused by their misalignment. For this purpose, each quadrupole and structure in the linacs will contain a beam position monitor (BPM), and will be supported by movers that can be remotely controlled (*e.g.,* like the magnet movers in the FFTB).

A beam-based approach will also be used in a trajectory feedback system to suppress low-frequency jitter of the bunch train as a whole, and if necessary, bunch-to-bunch position variations within the train (special multibunch BPMs will be used in this case). These feedbacks will be included in each of five instrumentation regions that will be located along the linacs. These regions will also include high-dispersion sections so that bunch energy and energy spread can be measured and used for rf control feedback. As with the operation of the SLC, if the feedback and alignment algorithms fail to suppress emittance growth to the desired level, non-local types of tuning will be used, such as orbit bumps. To aid in this tuning, and to monitor the beam phase space, the instrumentation regions will also contain beam size monitors.

Finally, we note some of the machine protection strategies that will be used to deal with the high-power beams. Initial tuning and recovery from beam shut-off will be done by running with a single bunch per pulse, and with a larger than nominal emittance so that the bunch will not damage the beam pipe if it were steered into it. A series of thin spoilers, one per structure, will be the first material intercepted by the bunch, and will disperse the resulting shower of secondary particles to prevent damage to other beam line elements. For nominal beam operation, a series of safeguards will be used to shut off the beam before it hits the beam pipe, and to prevent fast changes ($<$ pulse period) that could steer the bunches into the beam pipe.

7.3.2 Linac Optics

The linac focusing system is designed to contain the beam transversely as it is accelerated. In the NLC linacs, there are many effects that will dilute the transverse projected emittance of the beams. The three primary sources of emittance dilution are:

- Short- and long-range transverse wakefields.
- Dispersive and chromatic effects.
- Transverse jitter induced by quadrupole vibration.

Ideally, one would like to design a lattice that will minimize all three of these effects without significantly increasing the length of the linac. Unfortunately, these three sources of dilution impose opposing constraints on the focusing lattice. The wakefield dilutions are minimized by increasing the focusing strength, while the other dilutions are increased by the stronger focusing. In the NLC main linacs, we will use beam-based alignment of the quadrupoles and BPMs to minimize the dispersive emittance growth; the technique for this is straightforward. Thus, in principle, the optimization becomes a balance between the dispersive and jitter dilutions and the wakefield dilutions.

In practice, the optimization is actually a balance between the additional length due to the strong focusing and the wakefield dilutions. That is, we want to have a sufficiently strong focusing system so that the wakefield dilutions are small but not an excessively strong lattice that would unnecessarily increase the length of the linacs due to the additional quadrupole magnets that would be required. A parameter that characterizes the strength of the wakefield relative to the focusing is the BNS energy spread needed for autophasing: autophasing is the condition where the chromatic growth of a beam performing a coherent betatron oscillation exactly cancels the wakefield growth and thus the beam oscillates as a rigid body.

The energy spread required for autophasing can be estimated from a two-particle model as:

$$\sigma_{\delta\,\mathrm{BNS}} \approx \frac{N r_0 W_\perp(2\sigma_z)}{8\gamma}\frac{L_c^2}{6}\left(1+\frac{3}{2}\mathrm{ctn}^2\frac{\psi_c}{2}\right) \ , \tag{7.2}$$

where N, r_0, and $W_\perp$ are, respectively, the number of particles per bunch, the classical electron radius, and the transverse wakefield. In addition, L_c and ψ_c are the FODO focusing cell length and the betatron phase advance through the cell. A shorter cell length and larger phase advance yield stronger focusing and smaller BNS energy spreads.

In the NLC linacs, we have limited the BNS energy spread to less than 1.6%. This yields tolerances on the trajectory control and alignment of the accelerating structures that are reasonable using high-resolution BPMs and a beam-based alignment technique for the structures. For comparison, in most of the SLC linac $\sigma_{\delta\,\mathrm{BNS}} \gtrsim 4\%$.

Lattice Design

The main linacs are based upon simple FODO cells where the quadrupoles are separated by an integral number of pairs of accelerator structures—see discussion in Section 7.3.3. Two structures are paired together since this simplifies the waveguide routing that is required. For the NLC linacs, we considered two basic approaches: (1) placing quadrupoles, all having the same length, after every structure pair and (2) increasing the number of structure pairs between quadrupoles as $(E/E_0)^\alpha$ where α typically varies between 0.2 and 0.8.

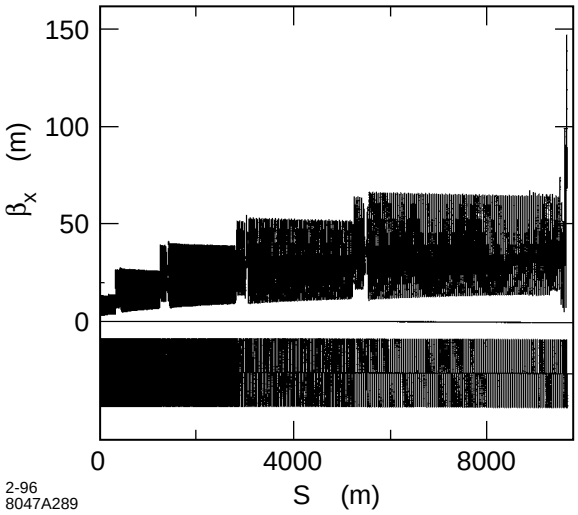


Figure 7-1. *Horizontal beta function in the NLC-II linacs.*

While the first option has the most flexibility, allowing one to set the "optimal" lattice regardless of the final beam energy, it is very space inefficient. This arises because many quadrupoles are located in regions where the horizontal and vertical beta functions are nearly equal. Thus these quadrupoles do not contribute to the net focusing and are wasting space.

In the cases we studied, the length of the linac occupied by the quadrupoles was more than 15% with the first option while it was roughly 7% with the second option. For this reason and because we feel that it can provide sufficient flexibility, we have chosen to adopt the second approach.

For NLC-I and NLC-II, the primary center-of-mass energy range that needs to be considered is between 500 GeV and 1 TeV. We have designed a lattice where the quadrupole spacing is increased by pairs of accelerating structures such that the cell length scales with $\alpha \sim 0.5$ when the center-of-mass energy is roughly 700 GeV; this is a compromise between the 500-GeV center-of-mass energy and the 1-TeV center-of-mass energy. This lattice will provide sufficient focusing at even lower center-of-mass energies, and the focusing can be decreased at the higher energies to decrease the sensitivity to quadrupole vibration.

To decrease the final beam energies much below 250 GeV while constraining the beam emittance dilution due to the transverse wakefields would be difficult. In this case, one would likely add a bypass line to skip part of the linac. For example, to operate the collider at 92 GeV to study the "Z_0" boson would only require the first fifth of the accelerating structures. The wakefields in the remaining structures simply increase the longitudinal and transverse emittance dilution and thereby decrease the performance of the collider.

In the lattice, the cell lengths increase in five steps, from roughly 8-m cells to 40-m cells, and the beta functions must do the same. We have matched across these transition regions without any significant increase in sensitivity. The horizontal and vertical beta functions for the NLC-II linacs are plotted in Figures 7-1 and 7-2. Notice that immediately following a transition point, the lattice contains a diagnostic section where the focusing is slightly weaker and is less uniform to provide optimal conditions for the emittance measurements; these diagnostic sections are discussed in Section 7.3.4. Figure 7-3 shows the horizontal and vertical phase advance per cell, and Figure 7-4 shows the horizontal and vertical BNS autophasing energy spreads. Figure 7-5 shows the location of the diagnostic stations on a plot of the horizontal dispersion.

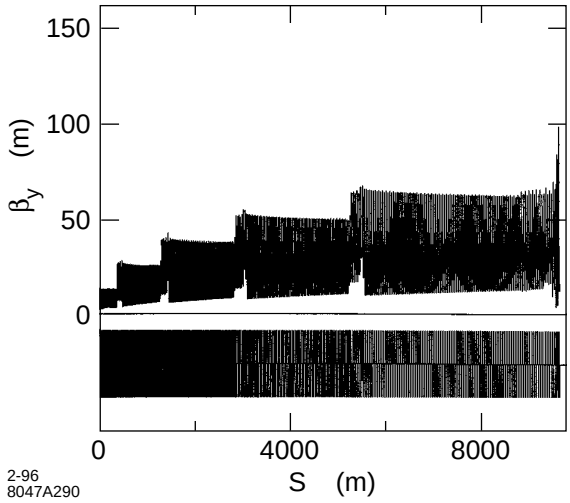


Figure 7-2. *Vertical beta function in the NLC-II linacs.*

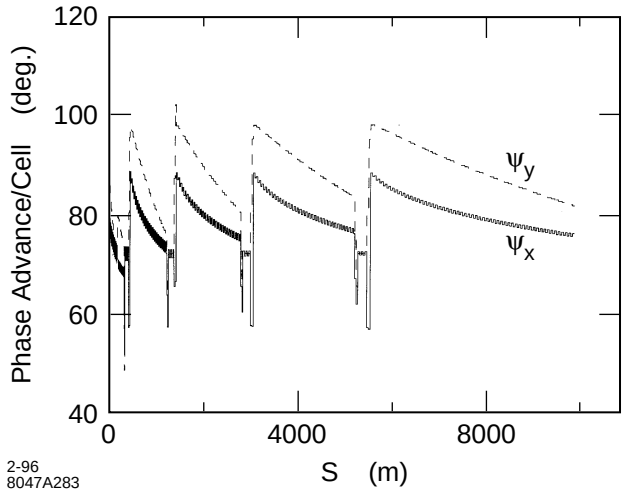


Figure 7-3. *Horizontal and vertical phase advance in the NLC-II linacs; the vertical phase advance is larger than the horizontal to optimize the focusing in the vertical plane.*

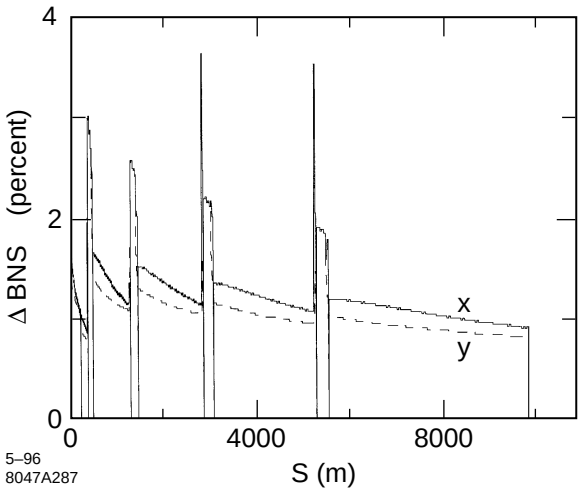


Figure 7-4. *Horizontal and vertical BNS autophasing energy spreads for NLC-IIb; the large values occur in the diagnostic regions.*

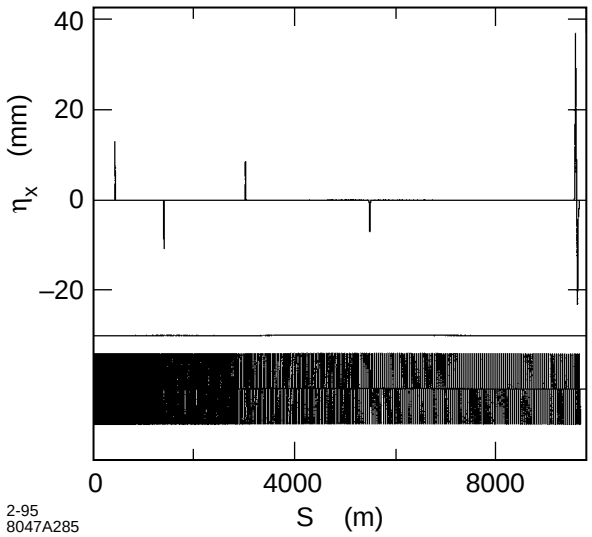


Figure 7-5. *Horizontal dispersion function in the NLC-II linacs; note that the diagnostic stations are located at the non-zero dispersion regions.*

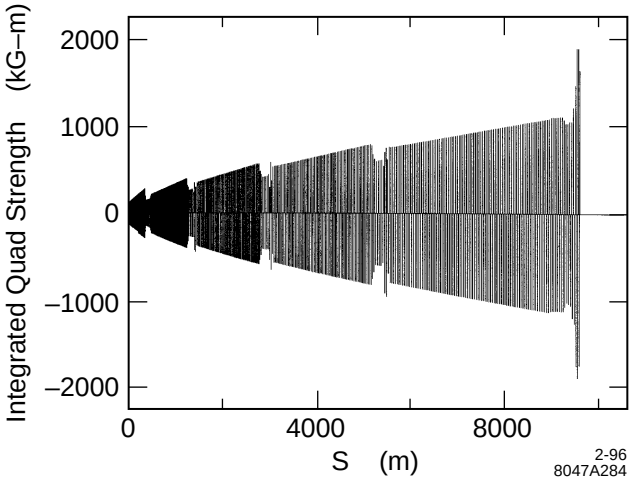


Figure 7-6. *Integrated quadrupole strengths in NLC-IIb lattice.*

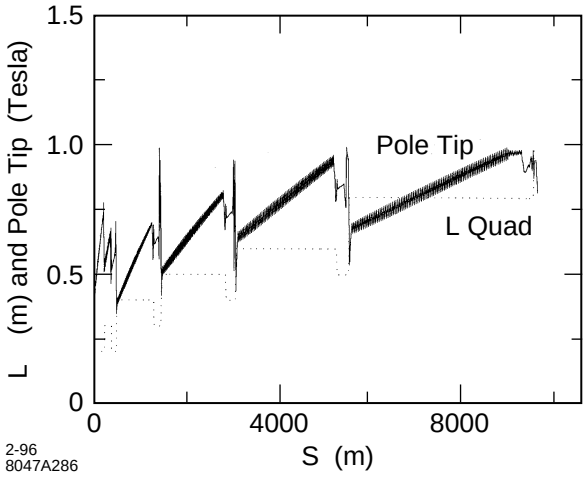


Figure 7-7. *Quadrupole lengths (dotted) and pole-tip fields (solid) in NLC-IIb lattice.*

The integrated quadrupole strengths required for the NLC-IIb lattice are plotted in Figure 7-6. The quadrupoles are assumed to have a 0.7-cm pole-tip radius. To keep the pole-tip fields reasonable, the quadrupole lengths are increased as the cell lengths are increased. The linac lattice uses five different quadrupole lengths: 0.2 m, 0.3 m, 0.4 m, 0.6 m, and 0.8 m in the regular lattice, while the final energy diagnostic section assumes a 1-m quadrupole with a 5-mm pole-tip radius; these longer quadrupoles are similar to those required in the collimation and final focus sections. The assumed lengths and the corresponding pole tip fields are plotted in Figure 7-7; note that some of the matching quadrupoles in the diagnostic sections are far too strong and must be lengthened.

Finally, we have studied the bandwidth of the linacs assuming that injection energy was correct but the accelerating gradient was too high or too low. The results are plotted in Figure 7-8 as a function of $B_{\rm mag} - 1$. The parameter $B_{\rm mag}$ is equal to:

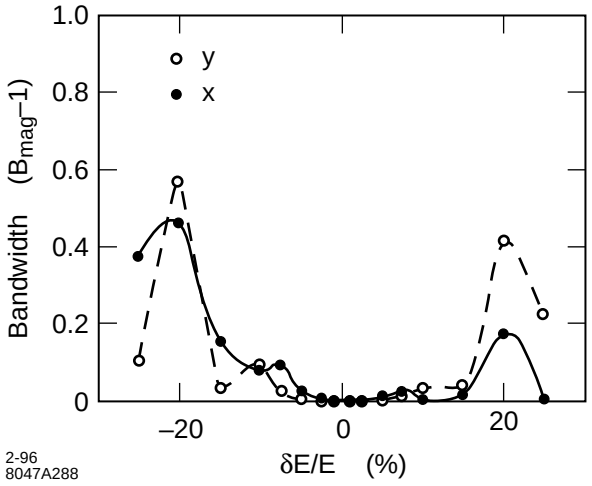


Figure 7-8. *Bandwidth of the NLC-II linacs plotted in terms of $B_{\rm mag}-1$ which is proportional to the emittance increase after filamentation. The energy deviation is due to an increased or decreased acceleration gradient.*

$$B_{\rm mag} = \frac{1}{2}\left[\frac{\beta^\star}{\beta} + \frac{\beta}{\beta^\star} + \left(\alpha\sqrt{\frac{\beta^\star}{\beta}} - \alpha^\star\sqrt{\frac{\beta}{\beta^\star}}\right)^2\right] \quad , \tag{7.3}$$

where $\beta^\star$ and $\alpha^\star$ describe the beam ellipse and are equal to the lattice functions β and α when the beam is "matched" to the lattice. With this definition, $B_{\rm mag}$ times the beam emittance is equal to the emittance after filamentation. Thus, when $B_{\rm mag} - 1$ equals one, there is 100% emittance growth. Figure 7-8 indicates that the optical bandwidth of the linacs is more than sufficient.

7.3.3 Beam Line Layout: Normal Sections

The main linacs are basically an array of X-band accelerator structures interleaved with a FODO quadrupole lattice. In this section, we describe the general layout and components of these "normal" beam line regions. The special diagnostic regions of the main linacs are described in Section 7.3.4. Further details on components and instrumentation are given in Sections 7.9 and 7.10, including newer concepts for the support system for the beam line components.

Structures

Figure 7-9 shows the beam-line layout near the beginning of the linacs where there is a quadrupole between each pair of accelerator structures; Figure 7-10 shows the end view and a close-up side view of the linac components. Note that the accelerator structures are supported in pairs on a common strongback, and that the strongback itself is supported by mechanical movers, which we call structure movers. Grouping the structures in pairs is motivated by the fact that they have a common rf feed and therefore are mechanically constrained, at least for large movements. The two structures will each be supported on the strongback at six locations and will be aligned on the strongback before installation (five of the six supports can flex to allow for thermal expansion of the structure along the beam axis). It will still be possible to adjust the six supports *in situ*, although not easily, given that an alignment precision of about 10 μm will be required.

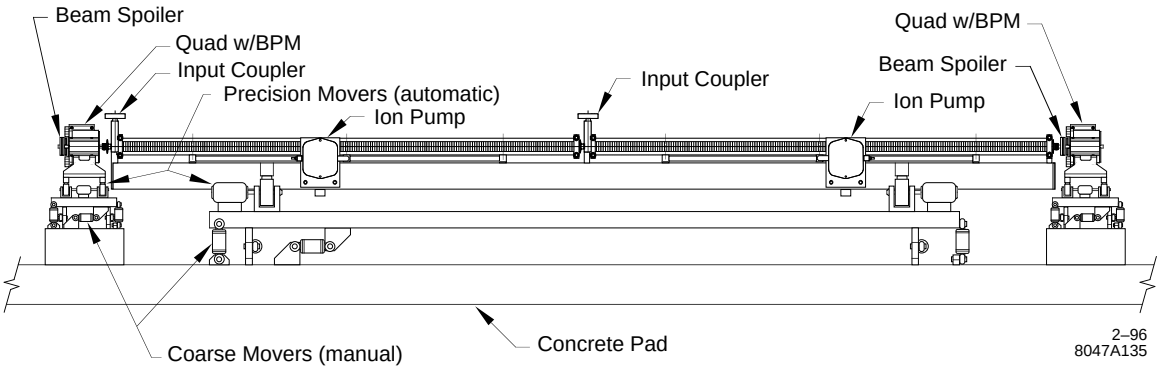


Figure 7-9. *Side view of linac components near the beginning of the linacs, where there is a quadrupole between each pair of accelerator components.*

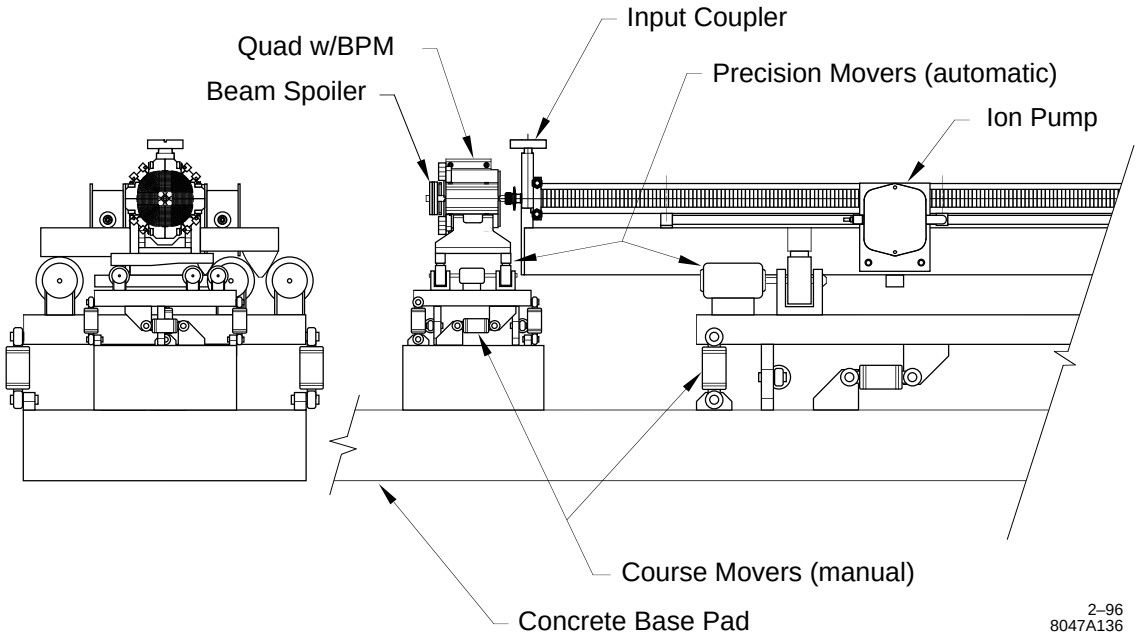


Figure 7-10. *End view and close-up side view of the linac components.*

The structure mover system will be similar to that used to move the SLC final focus triplet. The movers will provide independent horizontal and vertical positioning control at two points along the strongback, and will be used to remotely align the two structures as a whole relative to the beam trajectory. The alignment correction will be based on measurements of beam-induced signals from the structures, and the position changes will be made via five stepper motors that drive off-axis cams. Steps as small as 1 μm will be achievable over a range of ± 1 mm. Structure roll can be adjusted as well but will not be changed. This capability is a byproduct of using a mounting system that does not over-constrain the object that is being supported.

The structure movers will be attached to a common girder that will be attached to the beam line support pedestal via manually adjustable supports. These girder supports will be used to set the initial position and orientation of the structure pairs, and will have a ± 1-cm range of travel. They will also be used if a global realignment of the beam line is necessary due to ground settling.

Quadrupoles

Near the beginning of each linac, there will be a quadrupole magnet after each structure pair (Figure 7-9). However, the separation of the quadrupoles will increase in two-structure increments along the linac, from one pair of structures to five pairs at the end of the linac. These quadrupoles will form a FODO optical array as discussed in Section 7.3.2.

The quadrupoles, each of which will contain a BPM, will be moved as one unit for alignment purposes. Each quadrupole/BPM unit will be mounted on a magnet mover, which in turn will be mounted on a manually adjustable support. This system will have the same adjustment capability as was described for the structures, but the movers will be similar to those used in the FFTB.

BPMs

A stripline BPM will be located inside each quadrupole and will contain four strips at an inner radius of 5 mm with strip lengths of 105 mm. This length was chosen because it makes the period between the forward and reflected pulses from a single bunch equal to half of the bunch spacing, which is optimal for multibunch measurements. Since this length is shorter than the minimum quadrupole length, the addition of the BPMs will not significantly increase the beam-line length.

Spoilers

After each structure, there will be a spoiler that will be the first material hit if the beam is steered far off-axis. The spoiler will be a titanium disk that is 0.2-radiation-lengths long (about 8 mm), with a 2-mm inner radius and a 1-cm outer radius. For normal beam operation, it will do little to prevent damage to the beam pipe, but for tuning, it allows a single high-emittance bunch to intercept the beam pipe without causing damage; the Machine Protection System (MPS) is discussed further in Section 7.8. The integrated transverse wakefield kick due to a spoiler is about 10% that of the wakefield due to a single accelerator structure.

7.3.4 Beam Line Layout: Special Diagnostic and Control Sections

In a linac, it can be difficult to localize sources of emittance dilution, optical errors, and beam energy errors. If these errors propagate a sufficient distance, the emittance dilutions and optical errors will filament into unrecoverable

emittance dilution and the beam energy errors will lead to large variations in the betatron phase. The relevant length scale is the number of betatron oscillations. The NLC linacs are sufficiently long (roughly 80 betatron wavelengths compared to 30 wavelengths in the SLC linac) that it is important to measure the beam emittance and energy at multiple points along the linacs. In addition, because trajectory errors will also filament into unrecoverable emittance dilution, it is important to provide fast feedbacks that will stabilize the trajectory at multiple points along the linacs.

To monitor the beam emittance, energy, and trajectory, there are four diagnostic stations located in each linac and one station located at the end of each linac. The stations within the linacs are located at 400, 1300, 2900, and 5300 m; they are concentrated more closely at the beginning of the linacs because the beam is more sensitive at the low-energy end. In terms of betatron phase advance, these placements (including the station at the end of the linac) are located at 10, 22, 36, 54, and 78 oscillations. The diagnostic stations will provide five main functions:

- Beam emittance measurements.
- Energy and energy spread measurements.
- Fast feedback correctors.
- Bunch-by-bunch measurement and trajectory control.
- Beam dump.

Each of these will be described after briefly discussing the layouts.

The diagnostic stations are constructed from six FODO cells, each of which is roughly one-m longer than a normal FODO cell. This additional space is used by the diagnostics and kickers which are located near the quadrupole magnets. The accelerating structures occupy the middle of the cells just as they do in normal FODO cells. The quadrupole and accelerating section supports and girders are also identical to those of the main linac sections.

In addition to the six lengthened FODO cells, each of four diagnostic sections within each main linac includes a chicane which introduces dispersion for the energy and energy spread measurements; the length of the chicanes and matching quadrupoles is roughly equal to that of a single FODO cell. Finally, a much higher-resolution energy diagnostic section with a length of roughly 100 m is located at the end of each linac. This final energy diagnostic section also provides a net bending angle to compensate for the collimation section that follows the linac.

Emittance Measurement

The beam emittance measurements are performed using laser wire scanners located at different betatron phases which allows the beam matrix to be measured without changing the optics between scans. Only three independent X and Y wire scanners are needed to reconstruct the uncoupled 2×2 X and Y beam matrices. This effectively measures the projected horizontal and vertical emittances along with the beta and alpha parameters that describe the beam. If one also wants to measure the off-diagonal X and Y coupling terms, then at least four scanners with X, Y, and diagonal wires (U-wires) are needed since there are ten rather than six independent parameters that describe the beam. In this case, the horizontal and vertical phase advances must be separated to get independent measurements of the coupling terms.

In the NLC linacs, we do not expect betatron coupling to be a significant source of emittance dilution; the roll tolerances on the quadrupoles are relatively loose. Regardless, when commissioning the linacs or tracing sources of emittance dilution, it is important to be able to verify the degree of coupling in the beams. Finally, experience with the SLC emittance measurements suggests using an additional wire scanner to over-constrain the measurement

Station		Energy (GeV)	S (km)	σ_x (μm)	σ_y (μm)
NLC-Ib	Diag 1	20	0.4	26	7
	Diag 2	45	1.3	21	6
	Diag 3	90	2.9	18	5
	Diag 4	154	5.3	15	4
	Diag 5	245	8.6	12	3
NLC-IIb	Diag 1	30	0.4	22	6
	Diag 2	75	1.3	17	5
	Diag 3	155	2.9	14	4
	Diag 4	280	5.3	11	3
	Diag 5	490	9.4	8	2

Table 7-3. *Beam sizes at laser wire scanners in linac diagnostic stations for the NLC-I and NLC-II.*

and provide an estimate of the measurement error; this additional scanner also provides a backup if one scanner fails. Thus, each of the diagnostic stations includes five laser wire scanners which can measure the emittances and coupling while providing an estimate of the measurement accuracy.

The wire scanners are located near maxima of the vertical beta function where the vertical beam size is largest and the aspect ratio σ_x/σ_y is smallest. The beam energy and beam sizes at the wire scanners are listed in Table 7-3 assuming normalized emittances of $\gamma\epsilon_x = 3.6 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 4 \times 10^{-8}$ m-rad; these emittances are smaller than those expected through most of the linacs. To optimize the diagonal measurements, the U-wires should be aligned at an angle given by the inverse tangent of the beam aspect ratio. For the NLC linacs, this corresponds to an angle of roughly 15 degrees. The laser wire scanners should be optimized to measure vertical beam spots in the range $1.6\,\mu\text{m} \lesssim \sigma_y \lesssim 20\,\mu\text{m}$; these spot sizes correspond to a vertical emittance ranging from $\gamma\epsilon_y \lesssim 6 \times 10^{-7}$ m-rad at 20 GeV to $\gamma\epsilon_y \gtrsim 3 \times 10^{-8}$ m-rad at 500 GeV.

Energy and Energy Spread Measurement

The beam energy and energy spread is determined by measuring the trajectory and beam size in dispersive regions. The dispersion is generated by chicanes in the linacs and by short arc sections at the end of the linacs. The chicanes produce a maximum transverse displacement of roughly 1 cm and do not have a net bending angle or net transverse displacement. The arc sections at the ends of the linacs provide a total bending angle of roughly 0.5 mr to compensate for bending in the collimation sections that follow the linacs.

The locations, maximum dispersion values, and lengths of the energy diagnostic sections are listed in Table 7-4 for the NLC-II; the stations have the same length and dispersion in the NLC-I, but the beam energy differs because of the lower accelerating gradient. In all cases, the lengths and strengths of the systems were adjusted to limit the emittance growth due to synchrotron radiation to ~1%. Finally, the sections will cause the beam phase and bunch length to change as a function of the beam energy and energy spread. These changes are quantified by the value of the R_{56} and are very small. Assuming a 10% energy error through the entire linac, the bunch length will be increased by less than 2% and the beam phase will be changed by 0.3° which is insignificant compared to the 10% energy error. The values of the R_{56}, as well as the emittance growth due to the synchrotron radiation, are also listed in Table 7-4; in the NLC-I, the R_{56} values are the same, but the emittance growths are negligible since the growth depends upon the sixth power of the beam energy.

Station	E (GeV)	η_x (mm)	Length (m)	R_{56} (μm)	I_5 (1/m)	$\Delta\epsilon_x/\epsilon_x$ (%)
Diag 1	32	13	8	-125	1.6×10^{-10}	0.1
Diag 2	80	11	16	-41	4.7×10^{-13}	0.1
Diag 3	165	8.6	23	-17	1.5×10^{-14}	0.3
Diag 4	290	7.0	30	-9	8.4×10^{-16}	0.4
Diag 5	500	37	110	-4	7.9×10^{-17}	1.2

Table 7-4. *Parameters of linac energy diagnostic stations in the NLC-II.*

Station	η_x (mm)	$\delta E/E$ Res. (%)	$\sigma_{x\beta}$ (NLC-I)(μm)	σ_ϵ Res. (%) (NLC-I)	σ_ϵ Nom. (%)
Diag 1	13	0.08	37	0.34	0.8
Diag 2	11	0.09	28	0.30	1.0
Diag 3	8.6	0.12	25	0.32	1.1
Diag 4	7.0	0.14	24	0.34	0.8
Diag 5	37	0.03	21	0.06	0.3

Table 7-5. *Energy resolution of energy diagnostic stations for NLC-I and NLC-II.*

In the chicanes, the beam energy measurement will be performed using a single BPM located at the maximum of the dispersion, and several BPMs located in nearby nondispersive regions. These latter BPMs will be used to remove the dependence on the incoming trajectory. The BPMs at the high dispersion points will differ from the normal linac BPMs in that they will need roughly twice the horizontal aperture. The primary purpose of the energy measurement is to track changes in the beam energy over time. These energy changes lead to variations in the phase advance along the linacs which can lead to changes in the beam emittances and trajectories, and the feedback system efficiencies. In the NLC, we would like to limit the phase advance variations to roughly 15°. This requires limiting the beam energy changes to 0.5% at the beginning of the linacs and 0.25% at the end of the linacs.

The expected resolution of the energy measurements is listed in Table 7-5 for the NLC-I and NLC-II, assuming a systematic error of 10 μm on the BPM measurement. This error is assumed to arise from both the stability of the BPM measurement and the resolution on determining the incoming trajectory from the surrounding BPMs which are in nondispersive regions. These different measurements are insensitive to the absolute value of the bending fields or the horizontal alignment of the BPM. In addition, assuming that the bending fields have been measured with a relative error of 10^{-3}, the absolute beam energy can be calibrated to less than 1% by varying the chicanes.

In the final energy diagnostic, the absolute beam energy can be calibrated much more accurately by measuring the angular separation of the synchrotron radiation generated at the beginning and end of the bending magnets. In this case, it should be straightforward to attain an absolute measurement with a relative error less than 10^{-3}. A similar procedure is used in the SLC final focus beam dump with a relative accuracy of roughly 2×10^{-4}.

In addition to the beam energy measurements, the chicanes and arc sections will provide energy spread measurements. The energy spread will be determine from either the synchrotron light spot size or a laser wire scanner. The resolution, estimated from the ratio of the dispersion to the betatron spot size for a horizontal emittance of $\gamma\epsilon_x = 4 \times 10^{-6}$ mr, is listed in Table 7-5 for the NLC-I; the values for the NLC-II are roughly 30% smaller. The nominal beam energy spread, which is much larger than the expected resolutions, is also listed.

It is envisioned that the chicanes will be used during normal operation to continuously measure the energy and energy spread. Regardless, the vacuum chamber through the chicanes will be designed to allow the bending fields to be varied continuously from zero to full excitation and to accept beams with an energy as low as 50% of nominal. This range

Station	Energy (GeV) (NLC-I/NLC-II)	Corrector Strength [G-m] (NLC-I/NLC-II)
Diag 1	20 / 30	3 / 4
Diag 2	45 / 75	4 / 6
Diag 3	90 / 155	6 / 10
Diag 4	155 / 280	8 / 14
Diag 5	245 / 490	12 / 25

Table 7-6. *Fast Feedback corrector strength in the NLC-I and NLC-II*

of conditions requires a horizontal aperture that exceeds twice the nominal beam displacement in the chicane. In a chicane, the maximum displacement is equal to the peak dispersion function which is listed in Table 7-5 and varies between 13 mm and 7 mm. A vacuum chamber aperture of 3 cm by 1 cm in cross-section is more than sufficient.

Fast Feedbacks

At each diagnostic station, there are fast trajectory feedbacks that will control the horizontal and vertical trajectories. The feedback systems consist of horizontal and vertical pairs of dipole correctors followed by 10 high-resolution BPMs. The horizontal correctors are located at the first two horizontally focusing quadrupoles in the diagnostic sections while the vertical correctors are located at the adjacent defocusing quadrupoles. The BPMs are located at the subsequent focusing and defocusing quadrupoles.

To prevent the dipole correctors from driving the beam into the vacuum chamber, the corrector strength will be limited so the maximum trajectory deflections are less than 100 μm; the maximum required strengths are listed in Table 7-6. The feedback systems need to operate at the full 180-Hz repetition rate. Thus, the correctors and the vacuum chamber should be designed so the fields attain a consistent slew rate.

For more details of the fast feedback system see Section 7.5.3.

Bunch-by-Bunch Trajectory Control

In addition to the fast feedback for the bunch train trajectory, we will implement high-speed vertical kickers to realign the bunches in a train. This system is designed to compensate for the long-range transverse wakefields induced by accelerator structure offsets. The high frequency kickers will be short striplines separated by roughly 90° in betatron phase; details are described in Section 7.5.4. Because the system is designed to correct static offsets of the structures, it will be controlled by a slow feedback system.

Beam Dumps

At this time, the only beam dumps we are planning for the linacs are ones that will be located at the ends. Beams would be deflected to the dump by a 180-Hz kicker magnet so that the beam direction can be switched from the dump line to the collimation section in a single interpulse period. The dump line begins in the beta matching section for the collimation region; this is located after the final energy diagnostic station and consists of drifts roughly 100 m in length. Although a detailed design does not yet exist, at 500 GeV, an integrated field of 5.6 kg-m will deflect the beam by 2 cm after a drift of 60 m. This would allow a simple septum to further deflect the beam into the dump line. The kicker field of 5.6 kg-m is roughly 75% larger than the field of the 2–9 single-beam dumper in the SLC.

For more details of the beam dump, see Chapter 11.

7.3.5 Site Requirements and Tunnel Layout

Since the main linacs are the largest components of the NLC, the impact of the site on their cost and operation will be the dominant factor in choosing a location to build the accelerator. Although a systematic study of the site requirements has yet to be done, several factors that will impact the site selection can be readily identified. For example, the beam line must be located at least 20 ft below ground to provide radiation shielding. Access shafts to the surface will probably be required about every kilometer to provide cooling tower and substation connections, and to allow maintenance crews to reach the beam line. A flat terrain is desirable to minimize the depth of such access shafts and would also facilitate the establishment of a global coordinate system for alignment on the surface and the subsequent transfer of the reference system to the beam line.

There are a number of ground composition issues that will influence the site location. The general experience in constructing long beam line tunnels such as LEP is that the ground should be compact, homogeneous, hard and dry. The hardness of the ground has a number of tradeoffs. Tunnels bored in hard rock, such as granite, will generally be more stable in terms of ground vibrations and long-term settling than those bored in softer rock. However, the tunneling costs may be prohibitive and the site would likely be located in a mountainous region. Areas with softer rock such as sandstone and limestone, which are likely to have flatter terrain, would make the tunneling easier, but the tunnel would have to be better reinforced, which may offset much of the cost savings. Also, ground water variations are likely to have a bigger effect on the alignment of a tunnel built in softer rock. In this regard, the tunnel movement over several years should be < 1 cm on kilometer-length scales. Much larger variations will require that a global realignment of the beam line be performed.

Seismic activity on all scales will be a factor in the site selection. While one certainly wants to avoid areas with a high-rate of major earthquakes, even areas with low seismic activity should be carefully considered since minor fault lines may act as "slip-joints" for any geo-mechanical forces. On a smaller amplitude scale, ground motion will impact the machine operation due to the pulse-to-pulse beam position jitter generated by the movements of the quadrupole magnets. The ground motion frequencies of concern are those above about 0.1 Hz where the feedback systems that will be used to stabilize the beam may not produce the suppression that is required. The large micro-seismic activity (≈ 0.14 Hz) that originates from ocean waves tends to be fairly directional so its effect on beam motion can be minimized by orienting the beam line perpendicular to the direction of these ground waves (*i.e.,* a correlated motion of all quadrupoles produces less net beam motion). At higher frequencies, where the correlation lengths generally become shorter, the beam motion becomes more sensitive to quadrupole vibrations although the amplitude of ground motion tends to decrease as well. The tolerances in this regard depend on several factors and are discussed in Section 7.4. In general, however, the preference is for a naturally "quiet" site that is not subject to large "cultural" noise such as that produced by heavy machinery.

So far we have considered the site requirements without a specific model of the tunnel layout. However, the tunnel layout will influence the site selection, and vice-versa. The minimal requirement for the cross-sectional tunnel layout is that there be two areas, one that houses the beam line, and one that houses the klystrons, modulators and control electronics, which we call the gallery. We want to be able to inhabit the gallery while the beam line is in operation which will require that the two areas be separated by at least an equivalent of 10 ft of concrete for shielding purposes. With these constraints, we have begun to examine more seriously the two layout configurations shown in Figures 7-11 and 7-12. The KEK-type, so named because the JLC group at KEK has proposed this design, houses both the gallery and beam line areas in one tunnel, but with a 10-ft-wide concrete wall separating them. The semicircular shape of this tunnel would require the use of specialized boring machines. The second design is modeled after the SLAC linac in which the gallery is separated from the beam line with the ground in between serving as shielding. In its

construction, the gallery would be located just below the surface and would follow the terrain where possible. It would be built using a "cut-and-cover" method whereas the beam line area would be constructed with circular tunnel boring machines. Having the gallery underground will aid in stabilizing the gallery air temperature to within the 1°C requirement. In the beam-line area, we want a 0.1°C stability, which should not be too difficult to achieve with either design since the beam line tunnels will be well isolated.

Careful evaluation of a number of tradeoffs will be required prior to selecting a tunnel design. The SLAC design will likely be cheaper to build and will allow easier access to the gallery where most of the maintenance work will be done. The beam line is also better isolated, thus reducing the vibrations there due to the maintenance work, pumps, and pulsed-power systems in the gallery. The impact of the shielding requirements on the design, including the complications caused by the numerous penetrations to the beam line, will also be less significant. However, the SLAC design requires longer rf runs with more bends between the pulse compression system and the structures, which will reduce the rf power transmission efficiency by a few percent. Also, the long penetrations through the ground between the two tunnels will have to be well-sealed to prevent water seepage into the beam line.

For costing purposes, we have developed in more detail the SLAC-type layout, including the utility distribution. This is presented in Chapter 18 which describes the conventional facilities.

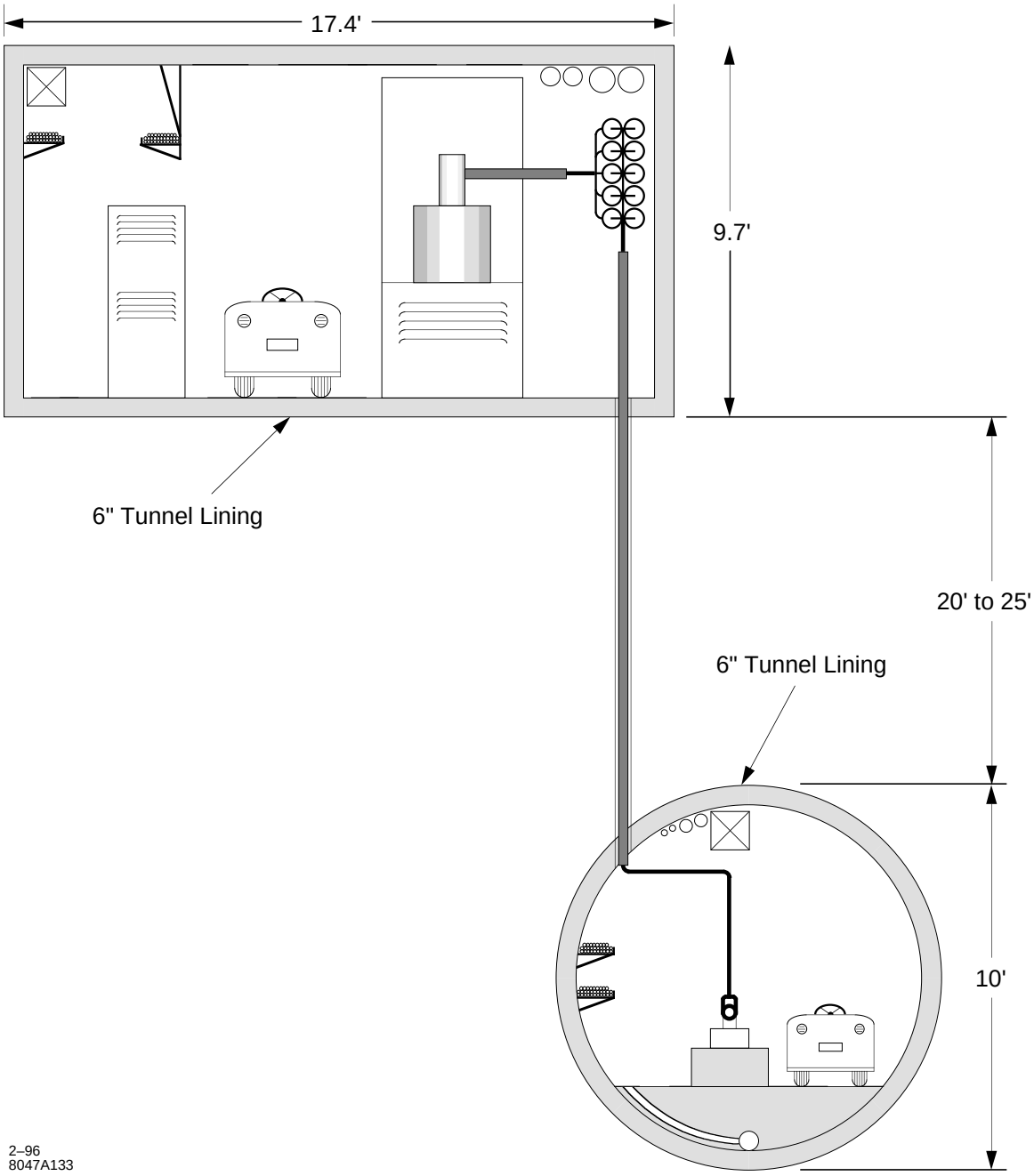


Figure 7-11. *Klystron gallery and beam line tunnel layout ("SLAC type").*

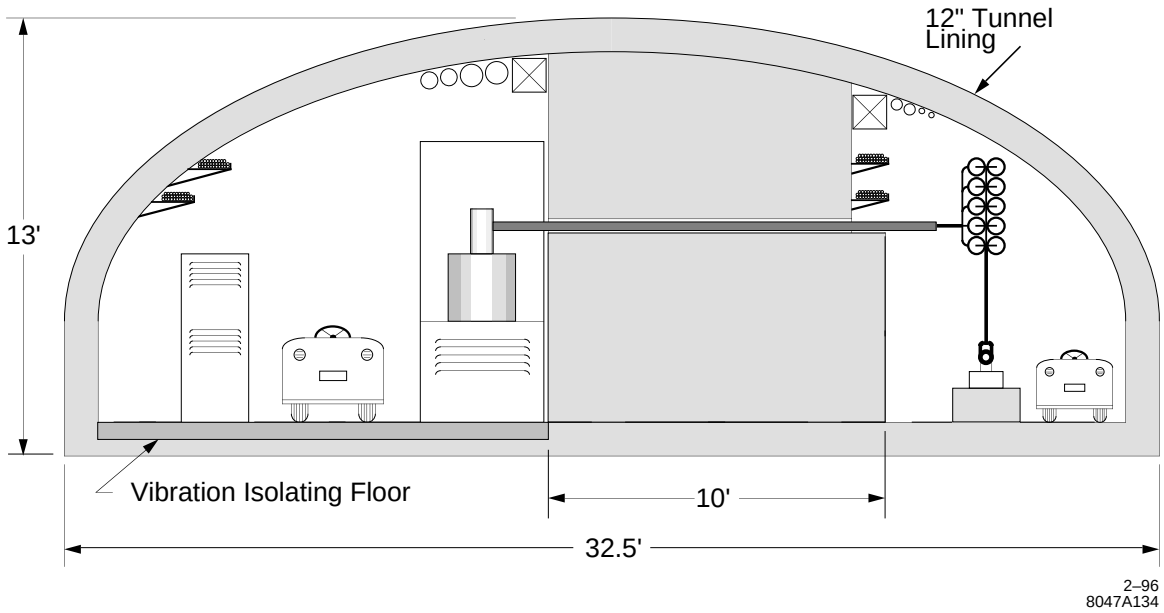


Figure 7-12. *Semi-elliptical highway-style tunnel layout ("KEK-type").*

7.4 Emittance Dynamics

The preservation of the transverse beam emittances and the control of the beam energy spread in the main linacs will require tight alignment and rf control tolerances. In this section, we is describe these tolerances with an emphasis on the underlying dynamical effects. We consider only the bare tolerances which are derived assuming no benefit from beam-based correction methods, including feedback. In the next section, we continue with discussions of stability, alignment, and feedback, where the emphasis is on the errors sources that set the scale of the problem, and the beam-based methods that will be required in cases where the bare tolerances cannot be met by other means.

We start this section by discussing the calculation of the short-range and long-range wakefields for the X-band accelerator structures. Although the intra-bunch dynamical effects of the wakefields in the NLC linacs are weak when compared with the SLC linac, they are still very important. Control of long-range wakefield effects is crucial to the operation of the NLC, due to its multibunch beam. Following the discussion of wakefield calculations, we discuss single bunch phase space issues (longitudinal and transverse), then multibunch phase space issues (longitudinal and transverse).

The discussion of single-bunch longitudinal phase space focuses on the need to adjust the average rf phase of the linac so the final energy spread is within the final-focus bandwidth. This average phase puts constraints on the energy spread profile in the linac which impacts wakefield and dispersive effects.

We next turn to single-bunch transverse effects where we present the bare tolerances for injection jitter, ground motion, power supply stability, quad, BPM and structure alignment, and rf deflections. In the case the quad and structure alignment, the results clearly indicate that conventional techniques will not be applicable as micron-level alignment will be required on lengths scales over 100 m.

For simplicity, the single bunch tolerances were computed assuming that the linac is autophased, that is, the energy spread of the bunches are adjusted so that the effect of the differential focusing along the longitudinal bunch profile exactly cancels the head-to-tail driving effect of the short-range transverse wakefields (using a correlated energy spread to offset resonate betatron growth is generally referred to as BNS damping). Because this requires energy spreads that are larger than the final-focus energy bandwidth, we will not be able to achieve this condition everywhere in the linac. Thus we consider how to optimize the rf phase profile along the linac to minimize betatron growth and yet keep the final bunch energy spread within the final-focus bandwidth.

Another option, that of using rf quadrupoles instead of (or in addition to) introducing a correlated energy spread to produce BNS damping, is also discussed. This would bring the linac closer to being autophased, but would require either a more complicated accelerating structure design or the introduction of special lossy rf quadrupole cells that would increase the length of the linacs.

The discussion of multibunch longitudinal phase space focuses upon the compensation of the multibunch beam loading, which varies from 20% to over 30% for the various linac parameters. The compensation scheme that has been chosen is "local" in that the rf input pulse is shaped to keep the energy of all the bunches in a train the same to within a few tenths of a percent throughout the linacs. The bunch-to-bunch charge variations and the ripple in the rf phase and amplitude must be well-controlled; simulation studies to set tolerances on these and other parameters are discussed in this subsection.

The discussion of multibunch transverse phase space includes the effects of the long-range transverse wakefield on emittance growth due to injection jitter and structure misalignments. The structure alignment tolerances are characterized by the length scale of the misalignments and are compared to the tolerances due to single-bunch effects. Finally, tolerances associated with the effect on ions on multibunch transport are presented.

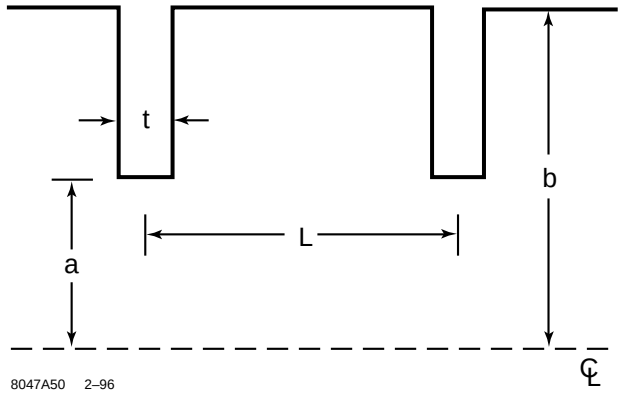


Figure 7-13. *The model of one cell of the NLC structure that is used in the calculations.*

7.4.1 Short-range Wakefields

Accelerator Structures

The accelerator structures in the main linacs will each consist of 206 cells and will operate at a frequency of 11.424 GHz. The structure design incorporates damping and detuning which is described in detail in Chapter 8. When computing the short-range wakefields generated in these structures, we ignore the manifolds that are used for damping but not the cell geometry variation that produces the detuning. Figure 7-13 illustrates the cell geometry (note the actual irises are rounded). Each cell is slightly different from its neighbors. The cell period is fixed at $L = 8.75$ mm and the cavity radius varies slightly but is roughly $b = 11$ mm. As one moves from the beginning to the end of the structure the iris radius varies from $a = 5.90$ mm to $a = 4.14$ mm and the iris thickness from $t = 1.26$ mm to $t = 2.46$ mm. The average value of a/λ for the cells is 0.187 where $\lambda = 26.25$ mm is the wavelength of the fundamental mode.

For our wakefield calculations we locally approximate the structure by a periodic model. The calculation method is as follows [Bane 1980a]: For the geometry of five representative cells, numbers 1, 51, 103, 154, and 206, we use the computer program KN7C [Keil 1972] to obtain the synchronous frequencies and loss factors of the first 250 or so monopole modes, which give us the low-frequency impedance. We approximate the high frequency contribution using the Sessler-Vainsteyn optical resonator model [Keil 1972, Brandt 1982]. Fourier-transforming these two contributions, we obtain the longitudinal wakefield for the given geometry. To calculate the wakefield of the entire NLC linac structure, we average the wakes obtained in this way for each of the five representative cells. The transverse (dipole) wakefield is obtained in an analogous manner, but using the computer program TRANSVRS [Bane 1980b] to obtain the dipole mode frequencies and loss factors. We first obtain the longitudinal dipole wakefield $W_z^{(1)}(s)$, then applying the Panofsky-Wenzel theorem [Panofsky 1956], we obtain the transverse dipole wakefield: $W_x(s) = \int_0^s W_z^{(1)}(s')\,ds'$.

The short-range longitudinal wakefield for cells 1, 51, 103, 154, and 206 of the NLC structure are shown in Figure 7-14 as a function of the distance behind the particle that is generating the wake. We should point out that the values at the origin, which should equal $W_z(0) = Z_0 c/(\pi a^2)$ [Palmer 1990, Gluckstern 1989] are about 5%–10% low, indicating some calculation error. The average wakefield for the entire structure is given by the dashed curve. A fit to the average wakefield, given by

$$W_z = 1388(V/pC/m) \cdot \exp\left[-1.16(s/mm)^{0.55}\right] \quad , \tag{7.4}$$

is shown by the dots. The net intra-bunch wakefield is obtained by convoluting this function with the longitudinal charge profile of the bunch. To set the length scale, the rms bunch length in the NLC will be in the 100- to 150-micron range.

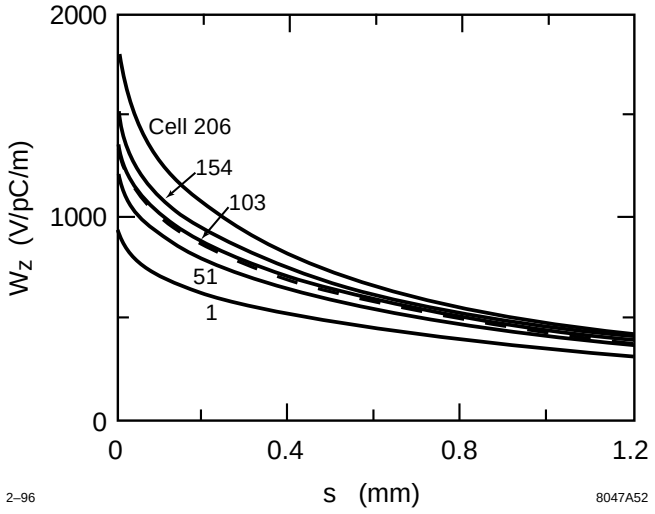


Figure 7-14. *The longitudinal wakefield of cells 1, 51, 103, 154, and 206 of the NLC structure. The average is the same as that of cell 103. The dashed curve represents the average wake, the dots, the empirical fit (see text).*

The short-range transverse wakefield for the representative cells are given in Figure 7-15. In this case, the slopes at the origin should equal $W_x'(0) = 2Z_0c/(\pi a^4)$; our results are slightly higher, but in no case by more than 4%. The average wakefield for the entire structure is indicated by the dashed curve in Figure 7-15. A fit to the average wakefield, given by

$$W_x = 88(V/pC/mm/m) \cdot \left(1 - \exp\left[-0.89(s/mm)^{0.87}\right]\right) \quad , \tag{7.5}$$

is indicated by the dots.

MPS Spoilers

There is one 8-mm-long MPS spoiler with an iris radius of 2 mm following every accelerator structure. Because the aperture is roughly a factor of two smaller than that of the accelerator structures, the transverse wakefield of these spoilers can be significant even though there is only one per structure. Figures 7-16 and 7-17 show the spoiler geometry and transverse wakefield kick on a 100-μm-long bunch; this kick is roughly 10% of the kick due to a single accelerator structure and thus has a small, but noticeable, effect on the dynamics. Finally, the relative magnitude of the longitudinal wakefield and the long-range wakefields are much smaller than those of the structures and can be neglected.

7.4.2 Long-Range Wakefields

In this section we discuss how the wakefield produced by a single bunch varies on the length scale of the bunch train. Given the transverse positions of the bunches in the train, these results can be used to compute the net wakefield acting on any bunch.

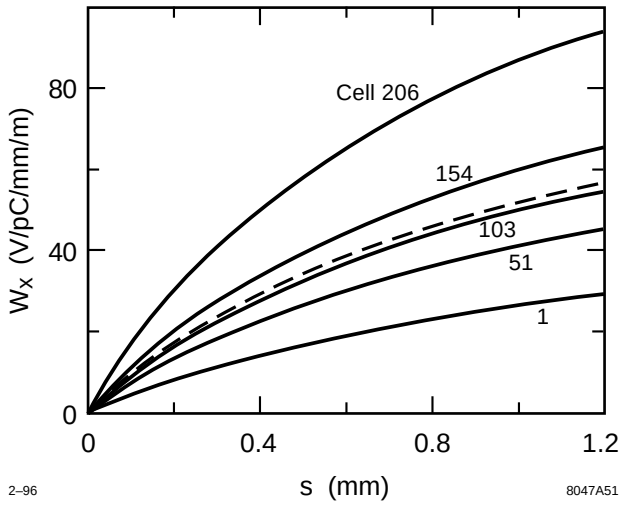


Figure 7-15. *The transverse (dipole) wakefield of cells 1, 51, 103, 154, and 206 of the NLC structure. The dashed curve gives the average wakefield; the dots, the empirical fit.*

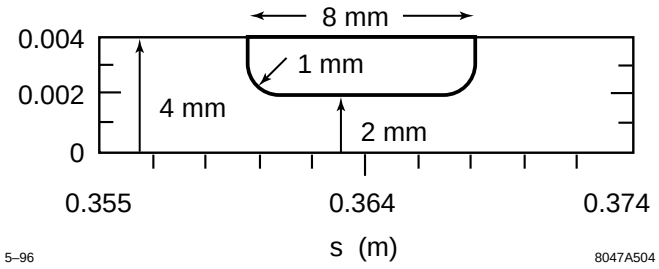


Figure 7-16. *Geometry of MPS spoiler located at the ends of the accelerator structures.*

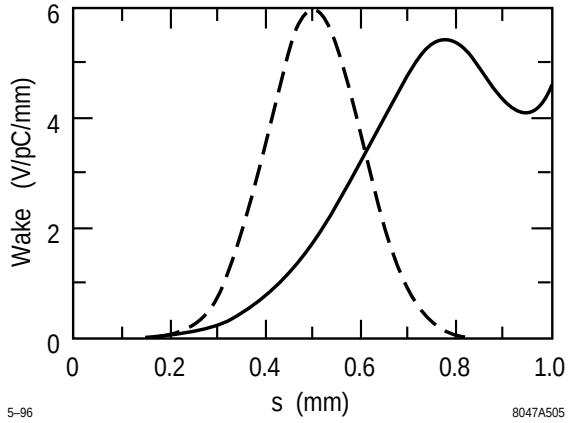


Figure 7-17. *Transverse wakefield (solid) due to a 100-μm-long bunch passing an MPS spoiler; this is roughly 10% of the transverse wakefield due to a single 1.8-m accelerator structure. Dashed curve is bunch profile with 100-μm σ_z.*

Detuned Structures

We first discuss the calculation of the long-range wakefield for a structure with detuning alone [Bane 1993]. The detuning involves varying the cell geometry to produce a Gaussian distribution in frequency of the product of the dipole mode density and the mode coupling strength to the beam. This yields an initial Gaussian fall-off of the integrated wakefield through the structure, which eventually levels off due to the discreteness of the modes. Further details of the detuning are given in Section 8.2. Here we outline the computational methods used to obtain the wakefield.

In this derivation, we consider each cell to be excited in some combination of cell modes (in practice, from the lowest or the lowest two dipole bands) and that the coupling holes determine the relation between the coefficients of these modes in two adjacent cells. Since the longitudinal electric field vanishes on the axis, we expect the coupling to be magnetic. We represent the structure by a single or double chain of circuits, where each circuit corresponds to a single-cell mode.

The uncoupled frequencies and the coupling parameters are obtained by applying the program TRANSVERS to periodic structures with the local cell dimensions at several points in the structure, and interpolating to obtain the values of these parameters at each cell. The coupled-circuit equations, supplemented by the appropriate boundary conditions in the first and last cells, are solved numerically by matrix inversion. The kick factors are then calculated from the resulting eigenfrequencies and eigenfunctions.

Most of the modes are localized in the structure in the lowest band. The results are shown in Figures 7-18 and 7-19 for the double-band model where only the two lowest bands are modeled. Note that the modes of the second band (roughly, those above 17 GHz) have low-kick factors and therefore do not contribute significantly to the wakefield. The wakefields obtained from the single- and double-band models do not differ greatly.

So far we have considered only the lowest two bands of dipole modes. When we perform a calculation that ignores cell-to-cell coupling but includes the effects of the modes of bands 3–8 [Bane 1994], we find that the wakefield amplitude is now an unacceptable 10% at the position of the second bunch, and it decreases only slowly as we move further back in the bunch train. This results from lack of detuning in some of the higher bands, especially the third and the sixth. Running URMEL [Weiland 1983], we find that near the light line, band 3 is a TM111-like mode and band 6 is a TM121-like mode. This suggests that by varying the iris thicknesses along the structure, with the thinner irises in cells with the larger radii—*i.e.*, near the beginning of the structure, and the thicker ones in cells with smaller radii—*i.e.*, near the end of the structure, we could detune these modes more. This variation has been incorporated in the our structure designs: a Gaussian distribution of iris thicknesses with an average of 1.5 mm, an rms of 0.25 mm, and a total variation of 1 mm is used. This is expected to reduce to 1% the contribution of Bands 3–8 to the wakefield at the subsequent bunch locations.

Damped Detuned Structure

Although the structure detuning significantly reduces the long-range wakefield, the amplification of the betatron motion of the bunches in an NLC built with these structures would be larger than can be tolerated. One solution would be to use different structure types with interleaved modes to improve the long-range decoherence of the wakefield. Another solution is to also damp the dipole modes, which has the added benefit of loosening the internal alignment tolerances for the structure. This approach is being actively pursued; the basic method is to channel the dipole mode energy to an external load through four manifolds that run along structure and couple to each of the cells (see Chapter 8 for details).

The equivalent circuit models used for the detuned structure may be extended to this damped detuned structure (DDS). Although the theoretical treatment of this problem is still being refined, the results so far are encouraging. The initial calculations were performed using a single-band model for the accelerating structure and coupling it to a transmission

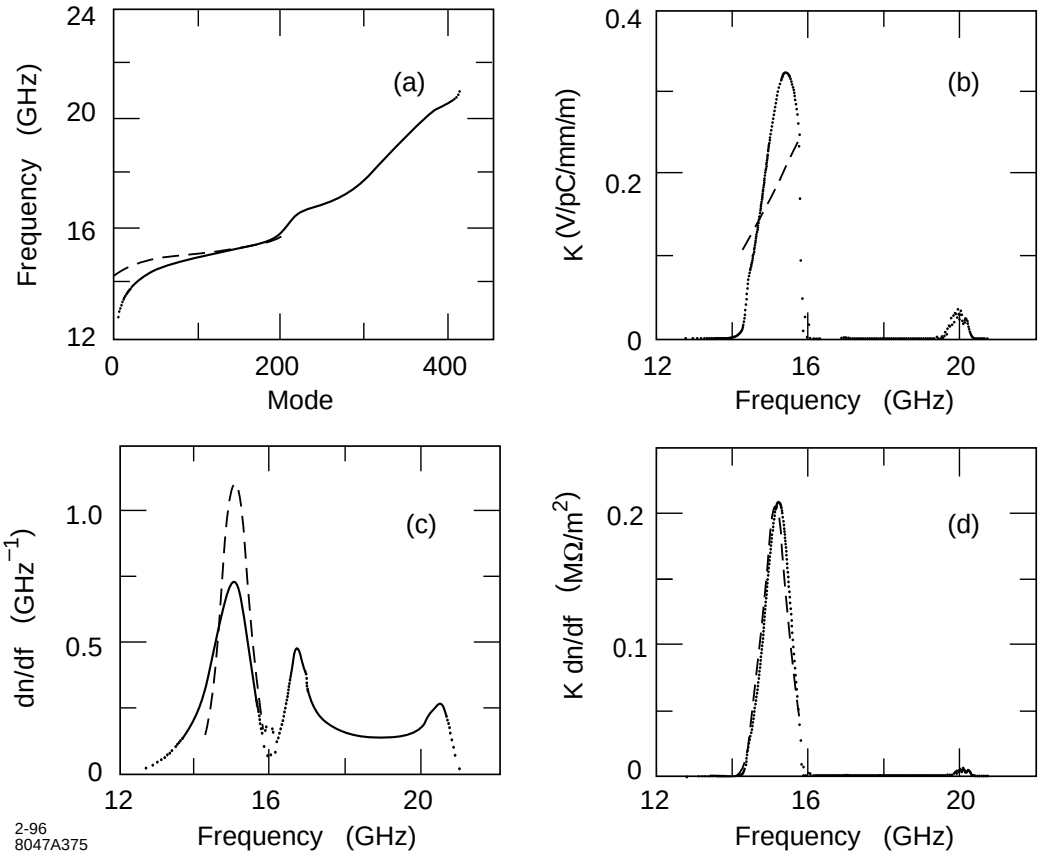


Figure 7-18. *Results of double-band coupled model: (a) the mode spectrum, (b) the kick factor, (c) the mode density (normalized so that the integral of dn/df over the first half of the modes is one), and (d) the product of kick factor and mode density. For comparison, dashed lines show the results from a model that does not include the effect of cell-to-cell coupling.*

line representation of each manifold mode that is included. The Qs of the modes from this treatment are shown in Figure 7-20 and the envelope of the resulting wake function is shown in Figure 7-21. The effect of copper losses have been included by combining the Q values shown in Figure 7-20 with an assumed copper loss value of 6500.

7.4.3 Single-Bunch Longitudinal Phase Space

The bunches will enter the NLC linacs with a 1% rms energy spread, which will be uncorrelated with longitudinal position along each bunch. In the linacs, the bunches will acquire a longitudinally-correlated component of energy spread due to the short-range wakefields and the curvature of the rf wave. Whereas the uncorrected component decreases inversely with energy, the correlated component will vary depending on how the phases of the accelerating rf field are set with respect to the beam timing.

There are three main constraints that one must consider when choosing the rf phase profile along the linacs. (1) The final energy spread of the beam must fit within the energy bandwidth of the final focus systems that are downstream of the linacs. (2) A correlated energy spread is needed in each bunch, with the bunch head at a higher energy than

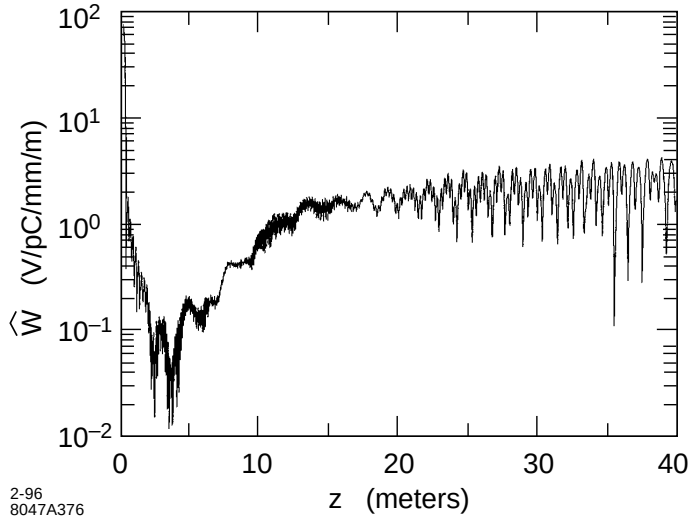


Figure 7-19. *The envelope of the wake function calculated using the double-band model.*

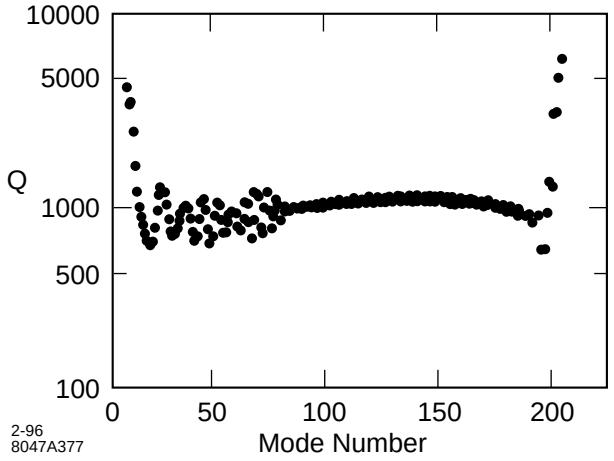


Figure 7-20. *The Qs of the modes in the lowest dipole passband, calculated for the damped detuned structure (DDS).*

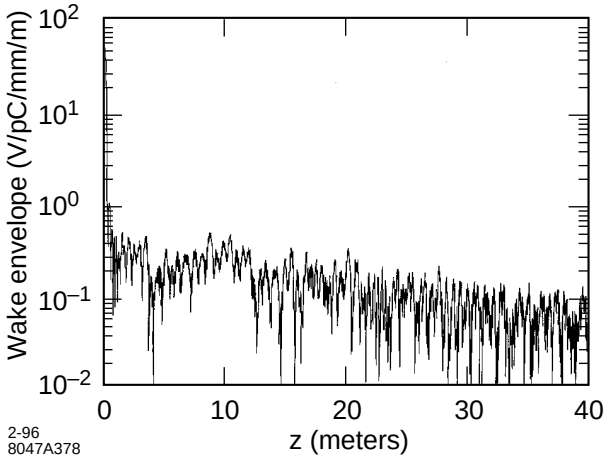


Figure 7-21. *The envelope of the wakefield function calculated for the damped detuned structure (DDS).*

the tail, to counteract the amplifying effect of the short-range wakefield on the betatron motion of the bunch. This approach to suppressing resonant growth is usually referred to as BNS damping; using an energy spread to exactly cancel the resonate driving term, so all parts of the bunch oscillate uniformly, is generally referred to as autophasing [Balakin 1983]. We want to run as close to the autophasing condition as possible. (3) rf power is precious so the phasing must be done efficiently to maintain a high accelerating gradient.

Satisfying these constraints will require some compromises. For example, in our linacs where the beta function scales roughly as the $\sqrt{E}$, the optimal energy spread for autophasing is nearly constant along the machine. This could be achieved by running the beams at a constant phase behind the rf crest but the resulting final energy spread would be larger than the final-focus bandwidth. Thus the beams will need to run more forward on the crest, on average, than is optimal for autophasing. The proposed solution is to run the beams somewhat behind the crest early in the linacs, and then sufficiently in front of the crest near the end to reach the desired final energy distribution, as is done in the SLC. This method requires some extra energy overhead so we must be efficient in selecting the phase profile.

The tradeoffs in this regard are discussed in Section 7.4.4 where we examine single-bunch transverse phase space effects in general. Another approach to this problem that uses rf quadrupoles is also described. Below we discuss the constraints on the average rf phase due to the final-focus bandwidth.

The Average Linac RF Phase

Let us consider the NLC-I parameters: initial energy $E_0 = 10\,\text{GeV}$, final energy $E_f = 250\,\text{GeV}$, and total structure length $L_{\text{acc}} = 8150\,\text{m}$. Let us assume that the bunches have a Gaussian longitudinal profile with rms length $\sigma_z = 125\,\mu\text{m}$, a population $N = 0.75 \times 10^{10}$ particles, and an initial uncorrelated energy spread of 1%. For this case, the single-bunch loading is 590 keV/ m. Also, the minimum rms energy spread is obtained when the average linac rf phase $\langle\phi\rangle = -18.0°$ (a negative phase represents a position in front of the rf crest); in this case the energy distribution is double-horned, with an rms spread of 0.27% and a horn spacing of 0.56%. Taking advantage of the larger energy bandwidth of the final focus systems, we choose to run the beams at $\langle\phi\rangle = -15.5°$, where the horn separation becomes 0.81%. In this way we introduce a slight, longitudinally correlated, energy variation, which will reduce the energy overhead needed for autophasing. Figure 7-22 shows the resulting energy distribution and energy correlation at the ends of the linacs.

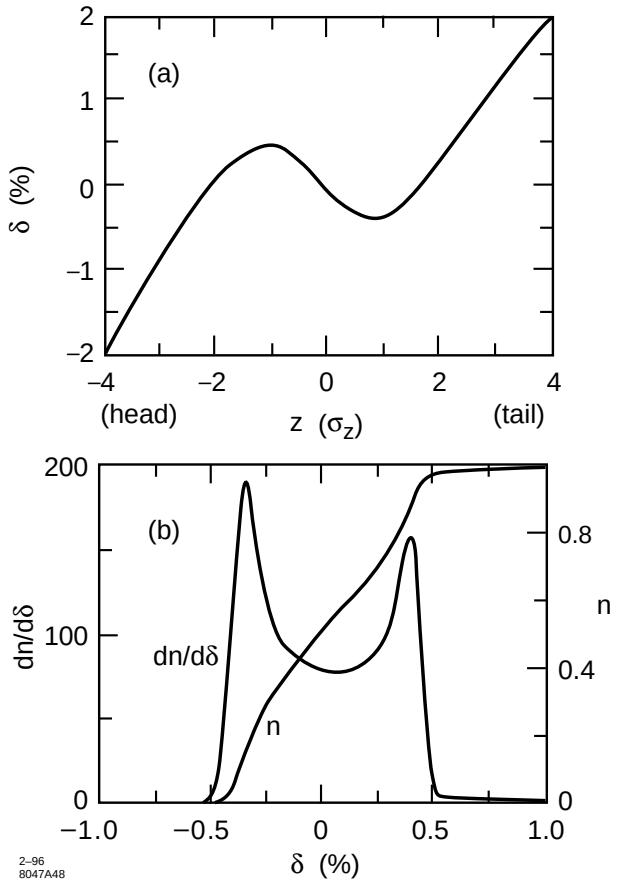


Figure 7-22. *For the NLC-I when $\langle\phi\rangle = -15.5°$: the correlated energy variation along the bunch (a), and the energy distribution and its integral (b).*

For NLC-II with $E_f = 500\,\text{GeV}$, $L_{\text{acc}} = 8150\,\text{m}$, $\sigma_z = 150\,\mu\text{m}$ and $N = 1.1 \times 10^{10}$, the single-bunch loading is 830 keV/m. In this case, we will run with $\langle\phi\rangle = -8.5°$ to achieve the energy distribution width of 0.8%.

7.4.4 Single-Bunch Transverse Phase Space

To transport the micron-size bunches through the main linacs without significantly degrading the bunch emittances will require that certain tolerances be met in regard to stability and alignment. In this section we discuss those tolerances that result from effects which degrade the transverse emittances of the all bunches equally, and ignore those effects which degrade the emittance of the bunch train as a whole (*e.g.*, kicks from long-range transverse wakefields). In doing so, we effectively treat the linacs as if only a single bunch were being transported per rf pulse. With this caveat, we examine jitter tolerances on injection and quadrupole motion; alignment requirements for the quadrupoles, BPMs, and accelerator structures; and limits on the transverse components of the rf field. In particular, we present only the bare requirements in that no beam-based alignment or feedback techniques are assumed to be used. In some cases, we

will see that the bare tolerances will be nearly impossible to meet, and so the use of these techniques will be essential. A description of these techniques and their effect on the tolerances are presented in Section 7.5.

To compute the tolerances, a number of simplifying assumptions were made. An important one is that the linacs are operated in the autophased condition so that the effect of short-range transverse wakefields on betatron motion can be ignored. For the linac lattice, it was assumed that the beta function and the spacing between quadrupoles scale as $\sqrt{E}$, and that the horizontal and vertical phase advance is 90° per cell. For the NLC Ib (NLC IIb) configuration, an initial quadrupole spacing of 4.2 m (3.2 m) was assumed which yields a beta function profile and quadrupole magnet count that are similar to that of the proposed optics (Section 7.3.2). Only the NLC IIb results for the vertical plane are presented here: the vertical tolerances for the NLC Ib configuration are generally about 35% looser, and the tolerances for the horizontal plane in either configuration are about a factor of three looser. The emittance growth values that are quoted are averages and are relative to $\gamma\epsilon_y = 4 \times 10^{-8}$. In most cases, the emittance growth is distributed as χ^2 with two degrees of freedom, which is exponential.

We next turn to a more detailed consideration of how one implements the desired BNS damping, *i.e.,* how one obtains as nearly as possible the autophased condition while staying within acceptable bounds of energy overhead and final energy spread. Also we examine the use of rf quadrupoles as a possible alternative to achieving autophasing.

Injection Jitter

Pulse-to-pulse variations of the bunch trajectories at injection to the linacs will degrade the machine performance in two ways. In traversing the linacs, the energy spread of the off-axis bunches will lead to filamentation, and thus increase the bunch emittances. Even in the absence of this effect, the bunches will collide off-center at the IP, thus reducing the collision luminosity. The effect of this jitter can be characterized by an equivalent increase in the bunch emittance, but requires an assumption about the transverse phase space of the jitter. If we assume that the jitter phase space is matched to the beam, as is generally seen in the SLC, then the emittance growth is equal to the square of the ratio of rms jitter to beam size. Hence, to keep the emittance growth below 3%, one must limit the jitter to less than 17% of the beam size.

Including the effect of filamentation does not change this result since the jitter phase space is matched to the beam. However, its effect would be a limiting factor if one could remove the jitter in a feed-forward manner between the end of the linacs and the IP. If the filamentation in the linac was complete, the jitter tolerance would not change, but we expect only partial filamentation, about 50° rms assuming a 0.1% uncorrelated energy spread that is independent of energy, plus a 1% uncorrelated energy spread component at the beginning of the linacs that decreases as 1/E (note the BNS energy spread can be ignored for this calculation if the beams are autophased). With this partial filamentation, the injection jitter tolerance would increase to about 26% of the beam size, which would hardly justify an effort at a feed-forward correction.

Quadrupole Vibration and Ground Motion

Another source of bunch trajectory jitter is from quadrupole magnet vibrations. If the transverse position of a quadrupole is displaced, it will kick the beam by an angle equal to the offset divided by the focal length of the quadrupole. If all quadrupole positions vary pulse-to-pulse, then the bunch trajectories at the end of the linacs will vary as the sum of the betatron trajectories produced by the individual kicks during each pulse. The size of this jitter depends on the spectrum of the individual quadrupole motions as well as the correlations of the quadrupole motions. As a simplification, we assume isotropy in that we suppose that all quadrupole motion spectra are equal and that the correlations depend only on the distance between quadrupoles.

A major concern in the NLC design is the effect that ground motion will have on beam transport via the quadrupole vibration it generates. When computing ground motion tolerances in this regard, one can treat each region of the NLC, such as the electron linac, independently from the rest of the machine to a good approximation. The quantity of interest is the deviation of the beam motion from the ground motion at the end of the region when the initial beam trajectory follows the ground motion at the beginning of the region. When computing the effect on luminosity, the deviation at the end of the region is assumed to remain unchanged relative to the beam size as the beam travels to the IP, where by symmetry, it can be shown to be related to that of the opposing beam. Thus the effect of ground motion on luminosity can be computed for both halves of the machine from the relative beam motion it induces in one half.

If the ground motion is uncorrelated quad-to-quad, then the tolerance one finds on the vertical rms motion is 6 nm for a 1.5% reduction in the time-averaged luminosity ($\gamma\epsilon_y = 4\times10^{-8}$ is assumed at the IP). In the NLC linac tunnels, the rms vertical ground motion, when computed over periods of several seconds, will likely be larger than this tolerance. However at low frequencies, where the ground motion is largest, the motion is generally correlated over long distances which loosens the tolerance.

While much data exists on ground motion spectra, little is available on long-range correlations. To see how typical these measurements are, and to better understand the dynamics of ground motion, we did a series of seismic measurements in the SLAC linac tunnel. These are described in Appendix C together with the formalism needed to analyze the data and to compute the effect of ground motion on beam transport.

From this study, it was found that the vertical ground motion can be reasonably well described by a model in which the motion is due to horizontally-traveling waves that are isotropically distributed in direction. For two points separated by distance Δz, the model predicts that the correlation in the vertical motion equals $J_0\left(2\pi f\Delta z/v\right)$ where J_0 is the zero-order Bessel function, f is the frequency of the waves, and v is the wave phase velocity, which we find depends on f:

$$v(\mathrm{km/s}) = 0.45 + 1.9\exp(-f(\,\mathrm{Hz})/2.0) \quad . \tag{7.6}$$

Using this correlation function, the tolerance on the rms vertical ground motion for a 1.5% reduction in the time-averaged luminosity was computed as a function of frequency. As a worst case, the tolerance was computed assuming that in addition to wave-like motion, there is an uncorrelated component of the motion that is equal to the seismometer electronic noise contribution to our measurements (see C.4 for details). The results for these two cases are shown in Figure 7-23. One sees that for frequencies above 4 Hz, the tolerance is roughly equal to the 6-nm value for uncorrelated ground motion. Below 4 Hz, the tolerance rises rapidly due to the fact that the wavelengths of the ground motion, v/f, are longer than the betatron wavelengths, so that the sum of the quadrupole kicks average out. In fact, it is the 180-m betatron wavelength at the end of the linac that determines this transition point. With the inclusion of uncorrelated motion, the tolerances at low frequency reflect the fraction of the measured power that is due to seismometer noise (*e.g.,* 1% noise power yields a $6/\sqrt{.01}$ = 60 nm tolerance).

In Section 7.5.1, we compare these tolerances to measured ground motion spectra, and discuss ways to actively suppress the effects of quadrupole vibrations. We should note, however, that ground motion will not be the only source of quadrupole vibration. Vibrations caused by the water flow through the cooling circuits for the quadrupoles will have to be controlled, as will the motions of the structures if they couple through the supports and vacuum pipe to the quadrupoles (the structures themselves have a much looser vibration tolerance than the quadrupoles). Also, care will have to be taken in designing the support systems to avoid mechanical resonances.

Finally, we consider the beam trajectory jitter that results from random fluctuations in the quadrupole power supply currents. Since the beam will not travel exactly on-axis through the quadrupoles, pulse-to-pulse changes in the quadrupole magnetic fields will produce an effect similar to uncorrelated ground motion. We expect that the beam will be centered in the quadrupoles to better than 100-μm rms (10 μm if no non-local steering is used), so the quadrupole

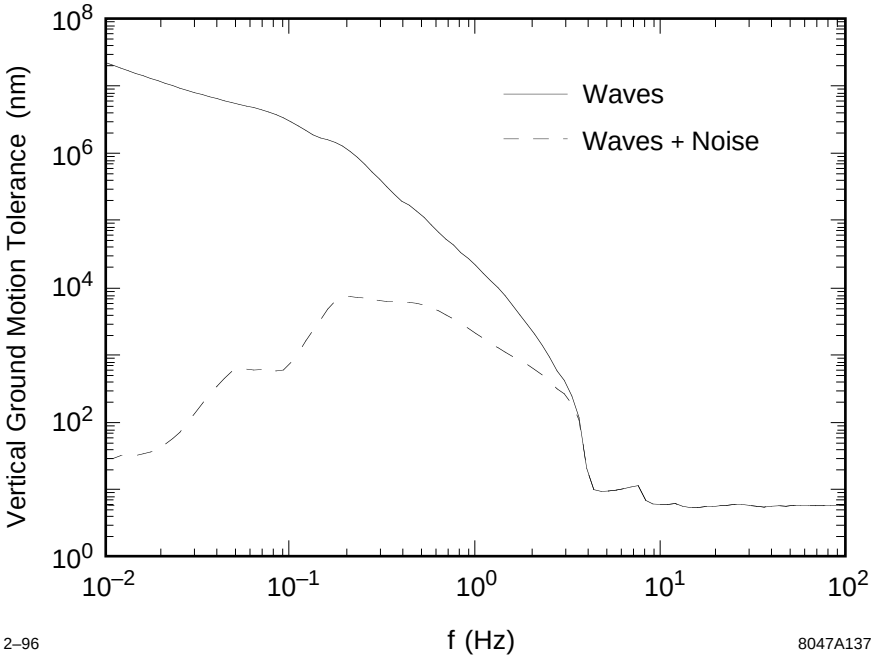


Figure 7-23. *Linac rms vertical ground motion tolerance for a 1.5% luminosity reduction as a function of frequency for pure wave-like motion (solid) and for wave-like plus uncorrelated motion equal to the seismometer electronic noise (dashed).*

power supplies must be regulated to 6×10^{-5}, assuming they are independent, to limit the luminosity reduction to less than 1.5%.

Quadrupole and BPM Offsets

If the quadrupoles are perfectly aligned along the linac axis and a bunch is injected on-axis, no dispersive emittance growth will occur. However, if the quadrupoles are offset from this axis, and the bunch is steered to the centers of the quadrupoles, then the differential kicks that the bunch receives due to its energy spread will generate emittance growth. This growth will be worse if the BPMs themselves are not centered within the quadrupoles since they will be used to steer the beam. The BPM offsets, which we define as the difference between the BPM centers (mechanical and electronic) and the quadrupole magnetic centers, will add in quadrature to the quadrupole offsets in their effect on dispersion if they are random. In this case, the tolerance on the sum is 3 μm for a 25% emittance growth in the vertical plane assuming an effective bunch energy spread of 0.6%.

Since the quadrupole alignment accuracy will depend on the distance scale over which the alignment is done, a better way to characterize the tolerance is in terms of the wavelength of the misalignments. For this purpose, we assume that the quadrupole offsets are sinusoidal with wavelength λ, and that bunches are steered to zero in the BPMs, which themselves are centered in the quadrupoles. Figure 7-24 shows the tolerance on the offset amplitude for a 25% vertical emittance growth as a function of λ assuming that the beam is injected on-axis. One sees that for $\lambda < 180\,\mathrm{m}$, the tolerances are nearly all less than 10 μm, and at larger λ, they rise rapidly. This rise occurs because λ becomes larger than the longest betatron wavelength in the linac, so the dispersion averages out (*i.e.*, for the same reason that the ground motion tolerances rise below 4 Hz in Figure 7-23). Using conventional survey methods, one would have a

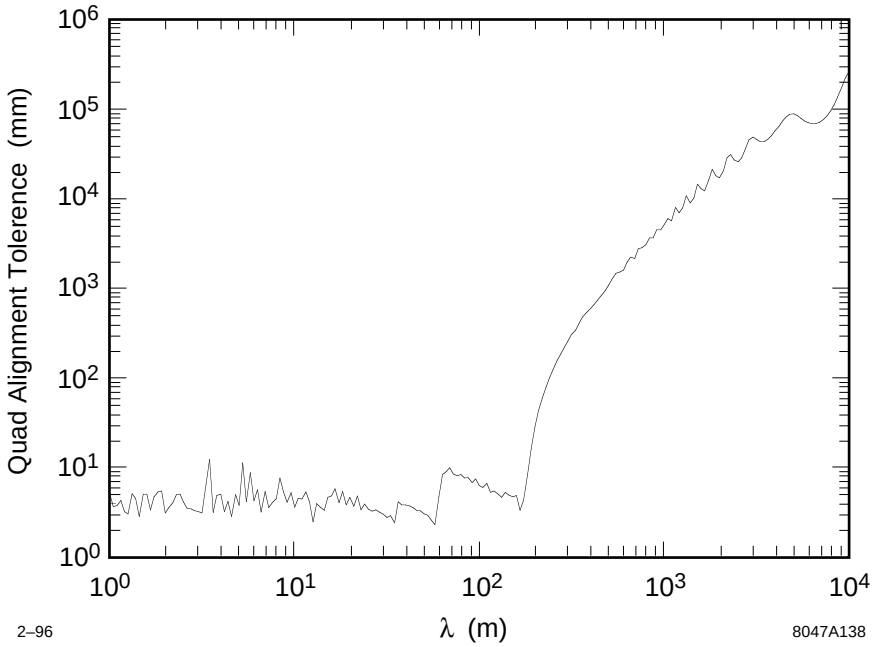


Figure 7-24. *Quad alignment tolerance for a 25% vertical emittance growth as a function of the wavelength* (λ) *of the misalignment. A 0.6% bunch energy spread is assumed.*

difficult time achieving a 3-μm alignment on a 1-m scale, let alone on a 180-m scale. The use of beam-based alignment will alleviate the need for such tight tolerances; this is discussed in Section 7.5.2.

Structure Offsets

If a bunch travels off-axis through a structure, the transverse wakefield that is generated will produce a differential kick along the longitudinal profile of the bunch. Like the differential kicks due to the bunch energy spread, the wakefield kicks will generate an emittance growth. To a good approximation, the effect of the wakefield kicks can be treated to first order where one ignores the position variation along the bunch when computing the generated wakefields. In this approximation, the emittance growth depends on the rms variation of the wakefield kick along the longitudinal profile of the bunch, just as dispersive growth depends on energy spread. The rms wakefield kick used to obtain the results presented here is 6.2 V/pC/ m/ mm, which assumes a 150-μm bunch length.

To compute the structure alignment tolerances, we assume that the quadrupoles and BPMs are perfectly aligned, and that the net wakefield kick to the bunch is removed locally by steering the bunch centroid to zero in the BPMs. In this case, a 6-μm rms random misalignment of the structures parallel to the beam axis yields a 25% vertical emittance growth. Like the dispersion analysis, we also computed the tolerance for sinusoidal misalignments of the structures, where both intra-structure and inter-structure alignment varies. Figure 7-25 shows the results for a 25% vertical emittance growth as a function of the wavelength of the misalignment. One sees that the most sensitive wavelength is about 180 m. This wavelength is resonant with the betatron motion near the end of the linacs where the weaker focusing leads to a larger growth. Like the quadrupole alignment tolerances, achieving these values would be extremely difficult with conventional alignment techniques. Instead, we will rely on beam induced signals from the structures to keep them centered about the beam trajectory. This method is described in Section 7.5.2. Note that we revisit structure tolerances in Section 7.4.6, where we consider multibunch effects as well.

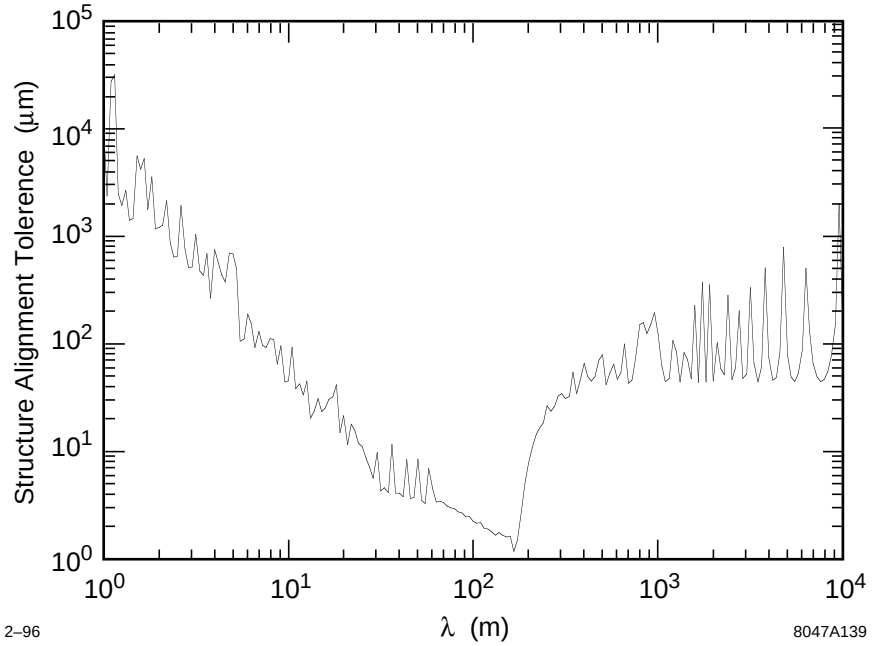


Figure 7-25. *Structure alignment tolerance for a 25% vertical emittance growth as a function of the wavelength* (λ) *of the misalignment.*

RF Deflections

Internal misalignments of the structures (*e.g.*, bookshelving) and dark currents within the structures can generate transverse rf fields that will deflect the beam. These fields, which can be static or vary pulse-to-pulse, will produce both a net and differential kick to the bunches. Using the wakefield tolerance result for uncorrelated structure offsets, one can readily compute the tolerance on the allowed integrated transverse rf field per structure assuming that the rf kicks are also uncorrelated structure-to-structure, as is likely to be the case.

To prevent an emittance growth of more than 3% as the result of the differential kicks produced along the longitudinal profile of a bunch, the tolerance on the integrated rf field amplitude is $1.2\,\text{keV}/\sin(\psi)$ where ψ is the phase offset of the bunch from the crest of the rf field. This tolerance holds whether the kicks are static, or vary pulse-to-pulse in amplitude but with a constant phase relative to the beam. Fields that are 90° out of phase with respect to the beam produce the largest slope and hence the tightest tolerance; for the transverse rf field effect on the net motion of the beam, the opposite is true. Here the tolerance is $0.04\,\text{keV}/\cos(\psi)$ for a 3% emittance growth if the transverse field jitters pulse-to-pulse, and $80\,\text{keV}/\cos(\psi)$ if the field is static but there is 10^{-3} pulse-to-pulse energy gain jitter in the groups of four structures that are powered in common.

For either emittance growth mechanism, it will be especially important that the tolerances be met if the rf transverse fields vary independently pulse-to-pulse since there will be no means to offset their effect. Currently we do not have much information on the size and temporal nature of the transverse rf fields in the X-band structures that have been built. (This is also true for the transverse field jitter in the SLC S-band structures, although there is a hint that it could be as high as 1-keV rms per 12-m section based on the observed white-noise-like beam jitter growth in the SLC linac.) From observing the power in the dipole modes that are excited in the DDS X-band structures without beam, we should be able to put limits on the transverse kicks. Also, from operating the NLCTA, we should learn enough to know whether the tolerances for most cases would be met in an NLC built with the same rf system.

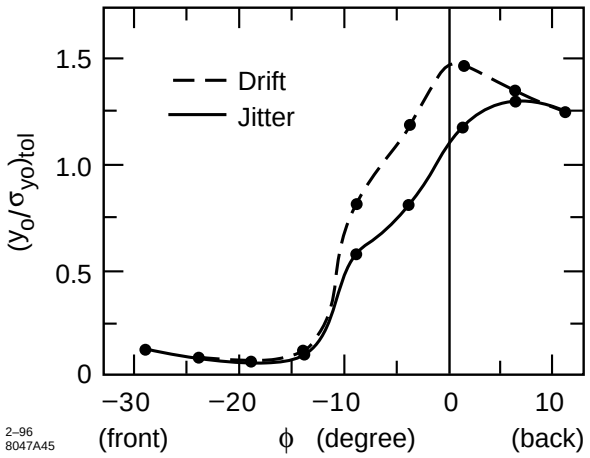


Figure 7-26. *Tolerance to oscillations at injection, i.e., the ratio (y_0/σ_{y0}) that gives 25% emittance growth at the end of the linac, when all klystrons are set to phase ϕ. For the jitter tolerance (solid curve) the emittance is calculated with respect to the beam line axis; for the drift tolerance, with respect to the bunch centroid.*

The RF Phase Profile Along the Linac

To suppress betatron amplification due to short-range wakefields, we plan to vary the klystron rf phase profile along the linacs to best achieve autophasing while keeping the average phase such that the final-bunch energy width is 0.8%, as was discussed previously. To decide which particular rf phase profile to use, we studied the sensitivity of the final beam emittance to betatron oscillations in an error-free linac. As figures of merit, we define t_0 as the injection jitter tolerance, and t as the injection drift tolerance. Specifically, t_0 is the ratio of the initial vertical bunch jitter, y_0, to the initial vertical bunch size, σ_{y0}, that results in a time-averaged emittance growth of 25% with respect to the beam line axis; t is defined similarly except that the emittance growth is with respect to the bunch centroid. Note that for the case of no wakefields and no energy spread, $t_0 = 0.75$ and $t = \infty$, and that for $t_0 > 0.75$, there is actual damping of the initial oscillation. The values of $t_0 = 0.75$ and $t = \infty$ are also the case for exact autophasing in the absence of an uncorrelated energy spread.

The oscillation studies were done using the parameters of the NLC-I, but with an active linac length of 7500 m, for which an energy width of 0.8% is obtained when $\langle\phi\rangle = -14.0°$. The lattice used was a piece-wise FODO array with 90° phase advance per cell. Between the quadrupoles there are an integer number of 1.8-m-long accelerating structures, beginning with two structures between the first two quadrupoles. Overall, β scales roughly as $\sqrt{E}$. Initially, $\langle\beta_0\rangle = 8$ m.

First, let us consider the simple case where all rf phases are equal. Figure 7-26 shows the tolerances as a function of this phase, ϕ. We see that at $\phi = -14.0°$, both the injection jitter and drift tolerances are 0.1. To be near the autophasing condition ($t_0 \approx 0.75$) throughout the linacs, the klystron phases need to be shifted by about $+10°$. Unfortunately, we cannot run this way since the final beam energy spread would be unacceptably large.

To constrain the final energy spread and still allow freedom to adjust the phase profile, we propose a method similar to that used in the SLC. We divide the klystrons into two families, with the first family containing the first n_a klystron units that are set to phase ϕ_a, and the second containing the rest of the units that are set to phase ϕ_b. In all cases the average phase is kept fixed at $\langle\phi\rangle = -14.0°$. For the purposes of this calculation, the i^{th} klystron unit denotes all klystrons that feed into the i^{th} group of four lattice cells; in total there are 88 such units. Injection jitter tolerance curves were calculated for fixed amounts of energy overhead, ΔE_{BNS}, for the phase shifting; here $\Delta E_{BNS} = 2$, 4,

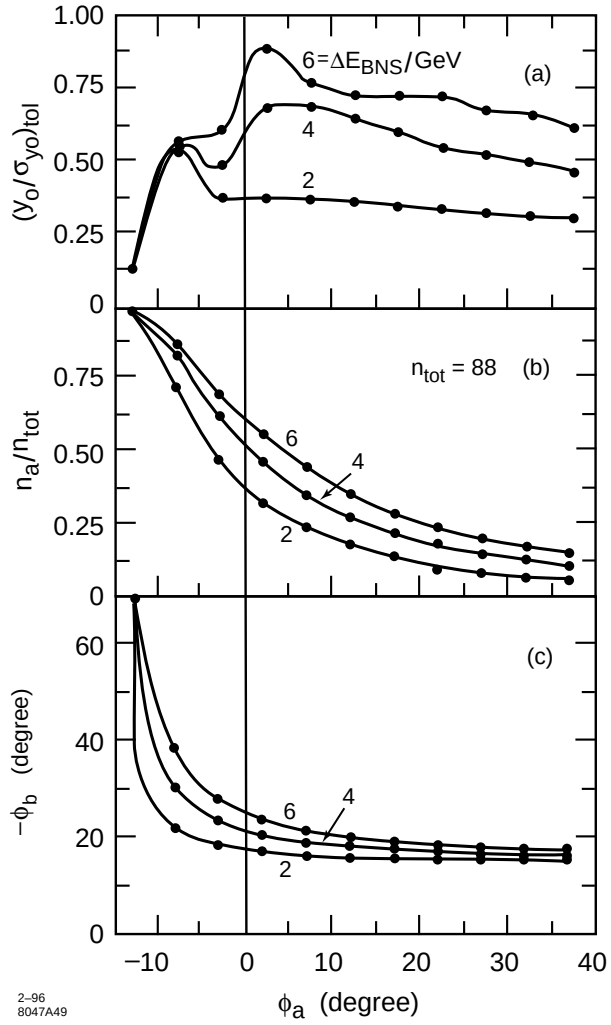


Figure 7-27. *Injection jitter tolerance for 25% emittance growth for klystron configurations employing two families of phases, plotted as (a) function of the phase of the first family ϕ_a for $\Delta E_{bns} = 2$, 4, 6 GeV. In all cases the final energy distribution has a width of 0.8%. Also plotted are (b) the relative number of units in family a, and (c) the phase of the second family ϕ_b .*

6 GeV. Plotted in Figure 7-27 are the injection jitter tolerance, the relative number of a-type klystron units n_a/n_{tot}, and ϕ_b, all as a function of ϕ_a. We see that if we are willing to give up 6 GeV, which is a 2.5% energy overhead, we can reach an injection jitter tolerance of $t_0 = 0.9$ with $\phi_a = 2°$, $n_a/n_{tot} = 0.55$, and $\phi_b = -23.7°$.

Next we studied the effect of jitter that begins at various locations along the linacs, not just at injection. In Figure 7-28 we plot the tolerances as functions of starting location of the jitter for the cases of no BNS damping, and with BNS damping, with $\Delta E_{BNS} = 6$ GeV and $\phi_a = 2°$. Note that BNS damping helps primarily in the first half of the linac, which is precisely where we need it most.

In the tolerance studies described above, we were mainly concerned with the behavior of the core of the beam, which is the part that contributes most to luminosity. In the simulations that were done for these studies, the longitudinal distribution was truncated at $-2.5\sigma_z$ (in front of beam center) and $2.0\sigma_z$ (in back of it), so that a few particles far out in

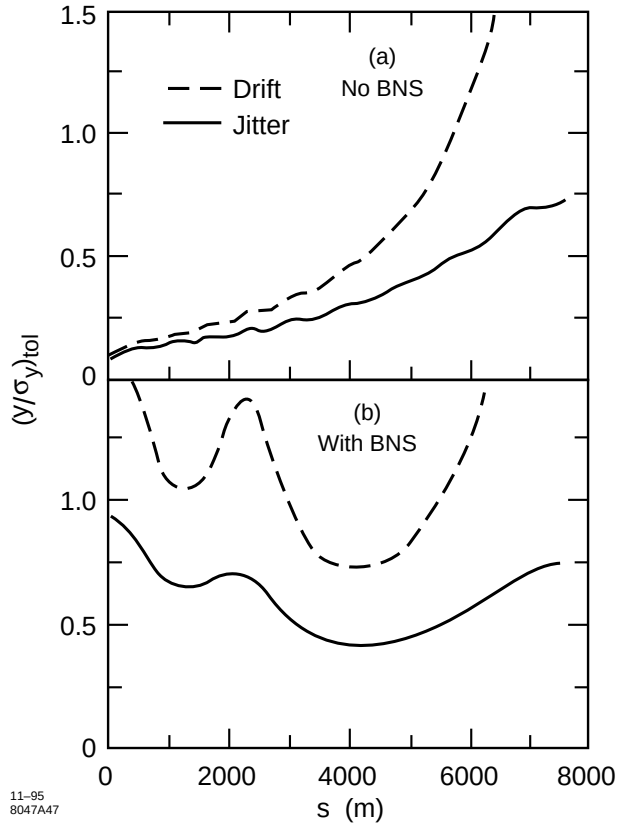


Figure 7-28. *Tolerance to oscillations beginning in the linac as function of the starting location of the oscillation, for the case of (a) no BNS damping, and (b) with BNS damping, with $\Delta E_{BNS} = 6$ GeV and $\phi_a = 2°$. Plotted is the ratio of the initial jitter to beam size at that location that yields a final emittance growth of 25%. For the jitter tolerance (solid curve), the emittance is calculated with respect to the beam linac axis; for the drift tolerance (dashed curve), the emittance is calculated with respect to the beam centroid.*

the tail of the distribution would not significantly affect the results. However, because of their potential for generating backgrounds, we would also like to know the behavior of particles in the tails. To show how a betatron oscillation affects such particles we have repeated the calculation for our BNS example, but now tracking longitudinal slices that cover $\pm 4\sigma_z$ of the bunch. In Figure 7-29, we show the phase space of a bunch that was injected with an initial vertical offset of σ_{y0}. Note that the largest amplitude slices are in the tail of the beam, at $z > 2.5\sigma_z$, followed by slices at the head, at $z < -2.5\sigma_z$. Therefore, it is desirable to truncate the longitudinal distribution of the bunches at $|z| \lesssim 2.5\sigma_z$.

The Option of Using RF Quadrupoles

In the CLIC project, rf quadrupoles have been proposed as a means to suppress wakefield-induced betatron growth as an alternate to the BNS method of using a correlated energy spread [Schnell 1987, Schnell 1991]. A small fraction of the accelerating structure cells (∼10%) are built in a noncylindrically symmetric shape, resulting in a fundamental mode that focuses as well as accelerates. With this system, in principle, BNS-like damping can be effective throughout

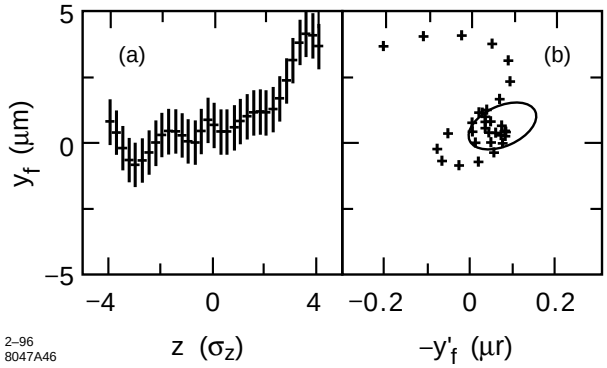


Figure 7-29. *Vertical phase space of a bunch that was injected off axis by σ_{y0} in a linac with $\Delta E_{BNS} = 6$ GeV and $\phi_a = 2^\circ$. Shown are (a) the final centroid positions y_f (the error bars indicate $\pm\sigma_{yf}$) vs. longitudinal position within the bunch z, and (b) the final centroid positions vs. final centroid angles (the curve is the 1σ ellipse of the beam slice at $z = 0$).*

the linacs, with no additional cost in energy overhead. For cavities coupled by narrow slits, the focusing gradient (in T/m) is given by [Schnell 1987]:

$$G = \frac{k_{rf} E_z}{2c} \sin(k_{rf} z + \phi) \quad , \tag{7.7}$$

where k_{rf} is the rf wave number, E_z is the accelerating gradient, and ϕ is the rf phase. Near $\phi = 0$ there is a nearly linear variation of gradient along the bunch. Although the model used to compute G is very simplistic, it has been reported that in realistic accelerator structures, where either the normal irises have been flattened, or where the cavity is elliptical with the ratio of the major to minor axis being about 2, the gradient can be 85% of that given by Equation 7.7 (though, in the latter case, with 25% reduction in fundamental mode R/Q) [Schnell 1991].

From Figure 7-26 we see that we would like an rf quadrupole focusing strength that is equivalent to about a $\Delta\phi = 10^\circ$ shift of average phase, or a focusing strength change at $z = \sigma_z$ of $k_{rf}\sigma_z \tan\Delta\phi \approx 0.5\%$. Assuming that the rf quadrupoles are 80% as effective as Equation 7.7, and using our machine parameters, we find that 4.5% of the structures need to be rf quadrupoles. Considering that we have a piecewise constant lattice, this number increases to 6.3%. Finally, assuming that only two versions of our accelerating structures will be built—one with an x-focusing rf quadrupole and one with an x-defocusing rf quadrupole —and that there will be on average many accelerating structures between neighboring quadrupoles, the fraction of each structure that would be used for rf focusing becomes 10%. This would require a major redesign of our accelerating structures.

Another approach is to build small, lossy rf quadrupole cells, one for each quadrupole (∼350 in total), that are powered to an effective gradient of 100 MV/m. In this case, the total length of the rf quadrupole cells relative to the total accelerator structure length is 1.5%.

7.4.5 Multibunch Longitudinal Phase Space

This section discusses multibunch longitudinal dynamics, and the method of multibunch energy compensation planned for the NLC.

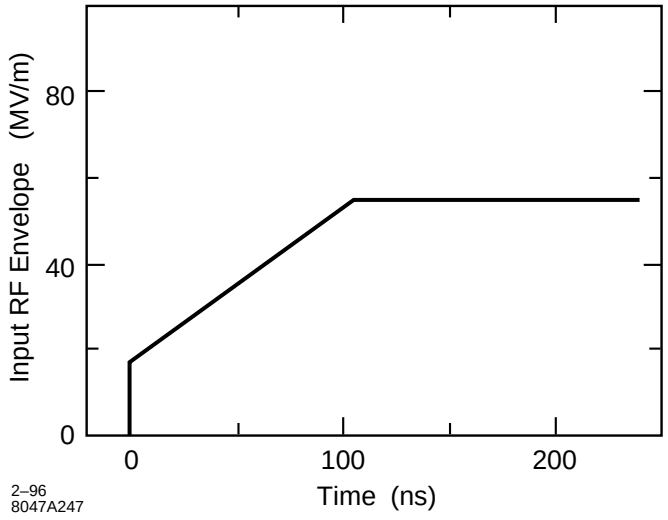


Figure 7-30. *Envelope of input rf pulse, used in long-pulse pre-filling compensation scheme.*

Without some method of compensation, there would be a drop in energy of about 25% from the head of a train to the tail due to beam loading in the accelerating mode. There is also potentially some bunch-to-bunch variation in energy due to HOMs—this must be controlled by detuning and/or damping those modes.

The tolerance on the multibunch energy spread and on the variation in the average energy of the beams have both been taken to be as little as $\Delta E/E \sim 10^{-3}$ in some of the simulations discussed or presented in this section. This may seem somewhat conservative since the energy bandwidth of the final focus is $.8\%$. However, it needs to be kept small due to the residual dispersion in the linac.

The method of beam-loading compensation is to pre-fill the rf structure with a pulse shaped to simulate the steady-state beam-loaded rf profile in the structure. This profile is then automatically maintained if the incoming rf field amplitude is kept constant during the passage of the train. A simple linear ramp yields good compensation, but the rf pulse could be further corrected to increase the compensation (and maintain it via feedback as conditions vary).

The input rf pulse used in our example is shown in Figure 7-30. However, since dispersion creates large amplitude variations on the front of the rf pulse, it is desirable to wait an additional 10 ns before injecting the bunch train, to allow the worst of these dispersion "wiggles" to propagate out of the structure.

We model the linac structures as having a constant unloaded gradient, with $2\pi/3$ phase advance per cell. Parameters are shown in Table 7-7. In this subsection, the parameters of the ramp are chosen to give a loaded average energy gradient of 50 MV/m. Charge per bunch is 1×10^{10} and the bunch spacing is 16 rf buckets ($\sim$1.4 ns).

A detailed simulation program has been written [Thompson 1993] in which one may take into account input rf pulse shaping and timing, the dispersion of the rf pulse as it transits the structure, the longitudinal distribution of charge within the bunches, the long-range wake including both the fundamental (accelerating) mode and higher order modes, the short-range wake, and phasing of the bunches with respect to the crests of the rf.

Figure 7-31(a) shows the average energy gained by each bunch in a single 1.8-m accelerating section. Figure 7-31(b) shows the fractional deviation of the energy centroid of each bunch from the average energy of the bunches. Here we have included the effect of dispersion of different frequencies in the rf pulse as it travels through the structure; this is

rf frequency, f_{rf}	11.424 GHz
Section length	1.8 m
Attenuation τ	0.505
Fundamental mode Q	7107
Fundamental mode loss factor, κ_1	203.75 V/pC
Filling time, T_f	100 ns
Bunch spacing	$16\lambda_{rf} \approx$42 cm
Bunch charge	1×10^{10}

Table 7-7. *Parameters used in energy compensation simulations.*

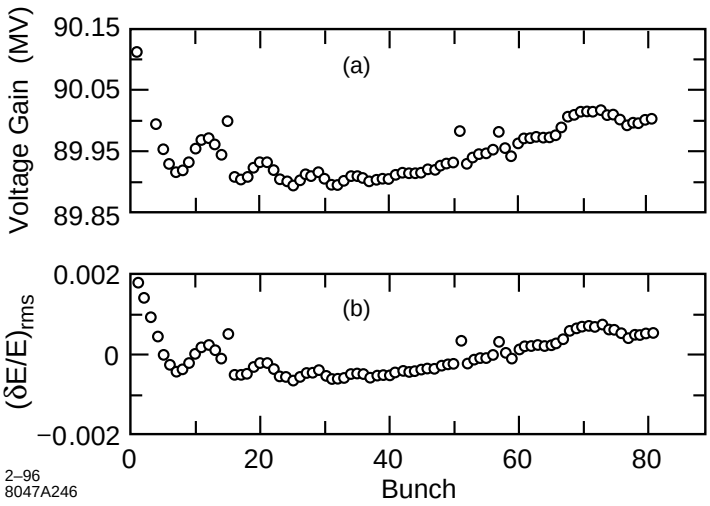


Figure 7-31. *(a) Average energy gained by each bunch in a single 1.8-m accelerating section. (b) Fractional deviation of the energy centroid of each bunch from the average energy of the bunches.*

the cause of the ripples seen in the energy distribution of the bunches. The energy spread can be reduced further by shaping the rf pulse to be slightly different from the linear ramp used here.

We must also run the bunches off-crest to do intrabunch energy compensation of the short- and long-range wakefields. The short-range wakefield used here is an average of the short-range wakes for individual cells of the detuned structure (Figure 7-14).

We assume each bunch has an approximately Gaussian distribution and divide the charge into nine equally-spaced macroparticles with Gaussian weight, assuming $\sigma_z = 100\ \mu$m. We seek the optimum phase for minimizing the energy spread by varying the rf phase, while also adjusting the height of the ramped rf pulse to keep the steady-state bunch centroid energy gain constant; the resulting energy spread as a function of rf phase at the bunch centroid is shown in Figure 7-32. There is some variation in the rms energy spread of individual bunches due to the effects of the long-range wakefields and the compensation scheme; thus optimization of BNS phase will be done for a bunch near the middle of the train which has an average energy spread.

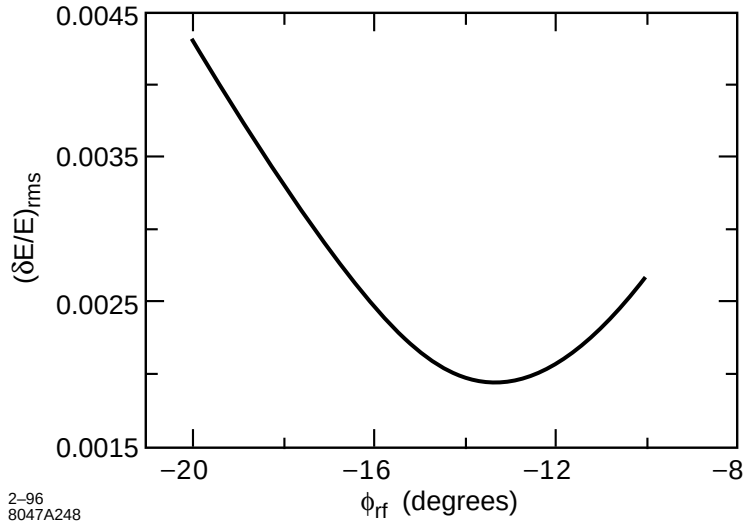


Figure 7-32. *The rms energy spread as a function of rf phase at the bunch centroid.*

Bunch Length Variations

Taking the bunch to be -13° off-crest (*i.e.*, 13° ahead of the crest, so that the tail receives more energy from the rf than the head, to make up for beam loading), we consider the effect of varying the bunch length, while keeping all other parameters constant. The effect on the rms fractional energy spread is shown in Figure 7-33.

Bunch Charge Variations

Taking the bunch to be -13° off-crest, we consider the effect of varying the bunch charge, while keeping all other parameters constant. The effect on the rms fractional energy spread is shown in Figure 7-34(a). The effect on the average (over all slices in all bunches) energy is shown in Figure 7-34(b).

RF Phase Variations

We consider the effect of rf phase ripple superimposed on the incoming 11.424-GHz rf pulse. Suppose the ripple is a cosine with given period (in ns) and amplitude (in degrees at the rf frequency).

It is essential to include dispersion of the rf pulse as it transits the structure, in order to treat this problem correctly. (If dispersion is neglected, a given bunch sees the same phase throughout the structure, since the variation of phase velocity with frequency, away from its value c at the rf frequency, is neglected.) Since calculations with dispersion are quite time consuming, we have treated each bunch as a single macroparticle for the calculations of Figure 7-35. Thus the contribution of intrabunch energy spread to the overall beam energy spread is not included. Here the amplitude of the ripple was taken to be 2°. Figure 7-35(a) shows the rms fractional energy spread as a function of the period of the ripple. Figure 7-35(b) shows the average energy as a function of the period of the ripple.

The variation of the energy spread scales approximately linearly with the amplitude of the ripple.

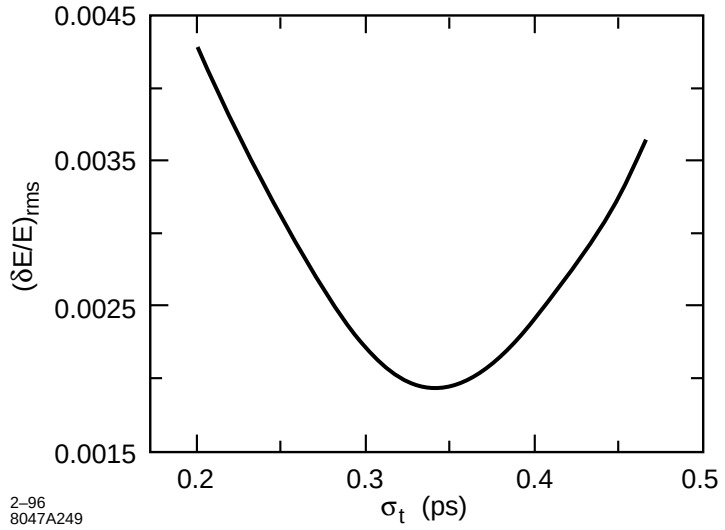


Figure 7-33. *The rms fractional energy spread as a function of bunch length, for bunches -13° off-crest.*

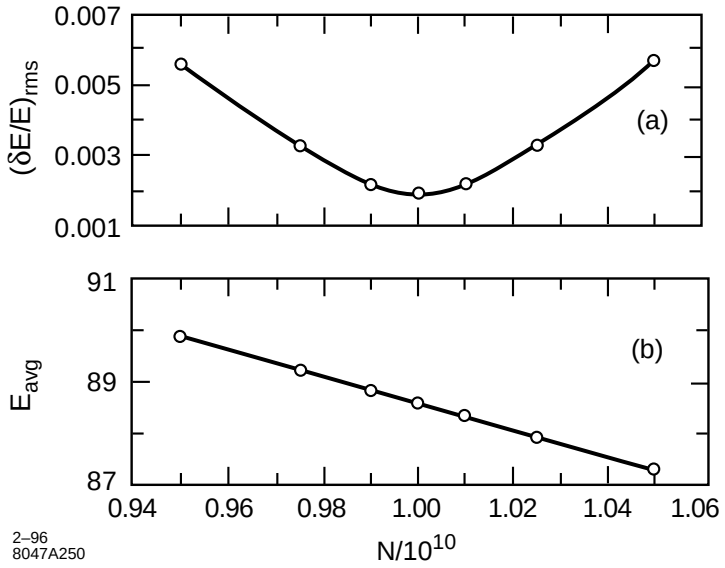


Figure 7-34. *(a) The rms fractional energy spread, and (b) the average (over all slices in all bunches) energy as a function of bunch charge, taking the bunch to be -13° off-crest.*

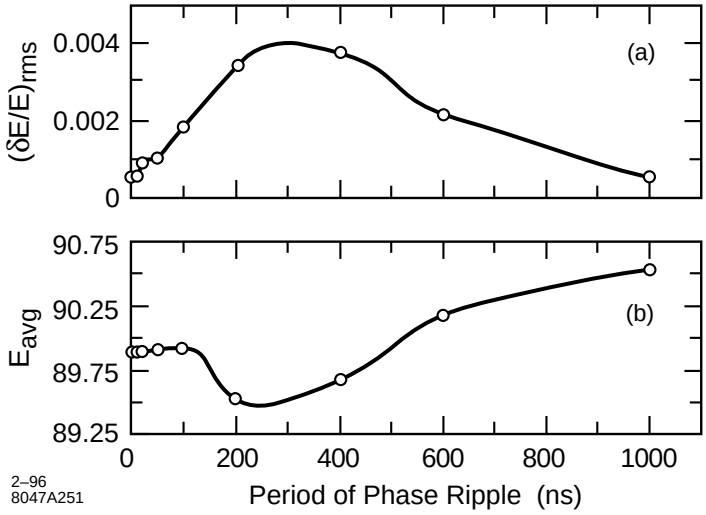


Figure 7-35. *(a) The rms fractional energy spread, and (b) the average energy, as a function of the period of phase ripple with an amplitude of 2°.*

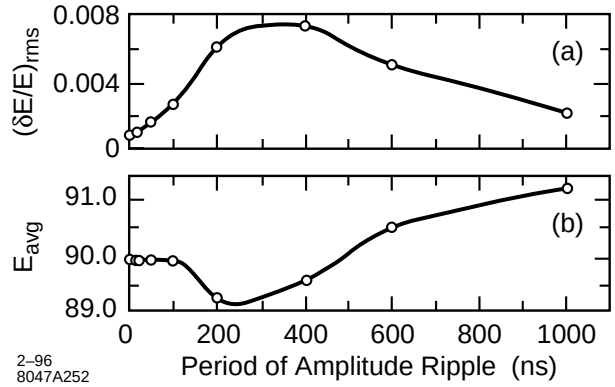


Figure 7-36. *(a) The rms fractional energy spread, and (b) the average energy, as a function of the period of ripple on incoming field amplitude. Amplitude of the ripple is 1 MV/m.*

Note that these are tolerances for variations that are the same from section to section. Thus they are pessimistic compared to tolerances for variations that are random from section to section.

RF Field-Amplitude Variations

We consider the effect of ripple in the amplitude of the incoming rf pulse. Suppose the ripple is a cosine with given period (in ns) and amplitude (in MV/m). We treat each bunch as a single macroparticle and include dispersion. Taking the amplitude of the ripple to be 1.0 MV/m, we obtain the result shown in Figure 7-36. Figure 7-36(a) shows the rms fractional energy spread (again ignoring the contribution from the intrabunch energy spread) as a function of the period of the ripple. Figure 7-36(b) shows the average energy as a function of the period of the ripple.

Again, note that for variations that are random from section to section, the tolerances should be looser than those computed here.

7.4.6 Multibunch Transverse Phase Space

In this subsection we discuss multibunch transverse phase space. Control of multibunch beam break-up has driven much of the work on accelerator structure design for the main linacs, since in the absence of a suitable cure (*e.g.*, detuning, damping of modes) it would be extremely severe. Some of the tolerances on the construction and alignment of these structures are very tight.

Structure Frequency Errors

The accelerating structures are designed to have a cell-to-cell variation in dimensions that detunes the dipole-mode frequencies and thus leads to a reduction in the amplitude of the long-range wakefield via decoherence of the modes. This frequency distribution is a truncated Gaussian (density in frequency space) and should be as smooth as possible to minimize partial recoherence of the wakefield over timescales less than or equal to the bunch train length. Small imperfections in the cell dimensions that lead to deviations in the mode frequencies by amounts comparable to or greater than the spacing between frequencies in the core of the Gaussian distribution produce a significant effect on the wakefield. The effect is worst if the frequency deviation is the same for a given mode in many or all of the structures. If, on the other hand, the frequency imperfections for the dipole modes are totally random from structure to structure, the effect is not as bad; in fact, randomizing the wakes at a given bunch spacing over a large number of sections can in some cases be beneficial.

Errors of 2.5 μm in the cavity diameter $2b$ produce fractional dipole frequency errors of about 1×10^{-4}. This is comparable to the fractional frequency spacing in the core of the Gaussian distribution used for detuning the structure.

The modification of the wakefield due to frequency errors has an effect on both the injection jitter tolerance and the misalignment tolerances. The effect can be quite significant for a detuned structure without any damping apart from the copper losses, and this was part of the motivation for going to the DDS.

Injection Jitter

There will be some offset of the injected trains with respect to the axis of the linac, and this offset will jitter from pulse to pulse. An off-axis bunch produces a transverse wake that drives bunches behind it. This wake must be kept small enough so that the projected multibunch emittance is not blown up significantly.

We have found that the emittance growth due to injection jitter is small if the envelope of the wake function is kept below about 0.8 V/pC/m/mm. For a structure with detuning alone, the wakefield envelope significantly exceeds this value even for a structure without fabrication errors (Figure 7-20).

As alluded to above, there are two extreme cases for the frequency errors: the case where the error in each frequency in the design distribution is the same in all sections (we denote this "systematic"), and the case where the error in each frequency is totally random from section to section. Note that our definition of "systematic" means that the errors are the same in corresponding cells of a given structure type, but they are still random from cell to cell in each structure type. In reality, the errors are not a simple superposition of these "systematic" and "totally random" errors, but for a first estimate, we shall treat them as such. Purely systematic errors of this sort would be a pessimistic extreme case. However, there will generally be some average error over all the cells of a given type, and this may be

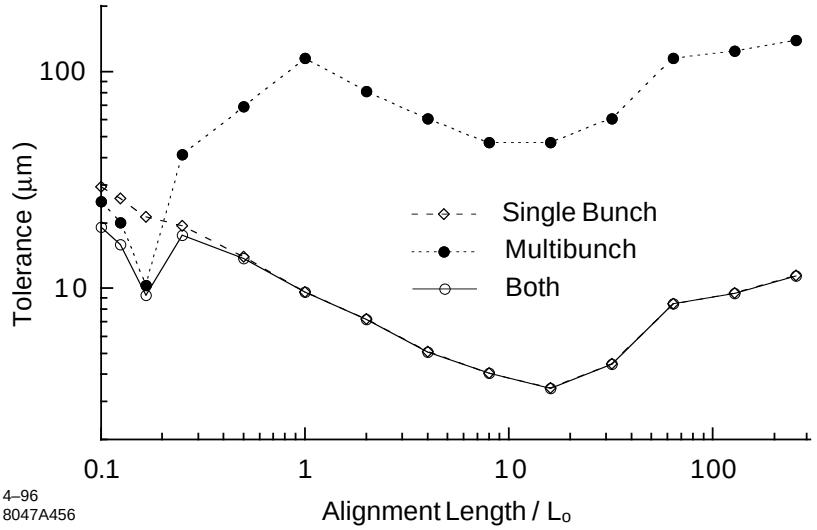


Figure 7-37. *Tolerance on the structure alignment for a 50% emittance growth in the NLC-IIb linacs as a function of alignment length, when considering only single-bunch effects, only multi-bunch effects, and both effects together.* L_0 *is the 1.8-m structure length.*

roughly approximated by the systematic component as defined here. Superimposed on this systematic distribution is the "totally random" component of the errors.

For a detuned structure with no damping, one finds that the blow-up of the betatron amplitudes of the bunches and the resulting multibunch emittance growth are unacceptably large. For the DDS, the situation is much better—even with systematic frequency errors with rms $\delta f/f = 10^{-4}$, the transverse blow-up and projected multibunch emittance growth are acceptable. This is quantified in Section 7.6.5 where we present multibunch simulation results.

Structure Misalignments

One of the tightest sets of tolerances for the NLC is that for the alignment of the DDS accelerator structures. We computed these tolerances by simulating beam transport in the linac, with the accelerator structures randomly misaligned in a piecewise manner. We considered alignment lengths (*i.e.,* piece sizes) ranging from multiples to sub-multiples of the 1.8-m structure lengths. For the sub-multiple lengths, the long-range wakefield was recalculated to account for the relative offsets of the cells.

In computing the tolerance for a given alignment length, a "seed" of random offsets was first generated for the structure pieces. The simulation program then "launched" the beam on axis and "steered" it to the centers of the quadrupoles, which were assumed to be perfectly aligned. Using the resulting emittance growth at the end of the linac and the rms of the piece offsets, the tolerance was computed for a 50% emittance growth.

For the NLC-IIb parameters, the average tolerance obtained from 100 seeds is shown in Figure 7-37 as a function of the alignment length. Three cases are plotted: one where only single-bunch effects are included in the simulation, one where only multibunch effects are included, and one where both are included.

In the figure, a relative minimum occurs in both the single bunch and multibunch curves near the alignment length of $20L_0 \approx 40$ m, which is the average value of the beta function at the end of the linacs; this is expected for a lattice for which $\beta \sim \sqrt{E}$. Also, for short lengths, the single bunch curve increases as the $-1/2$ power of the length: this

is because the wake kicks average out over distances shorter than the betatron wavelength. However, the multibunch curve drops for alignment lengths less than L_0 because the detuning is less effective in the shorter pieces. Finally, note that the combined tolerance is dominated by the single bunch effects except for alignment lengths less than about $0.2L_0$.

As with the quad alignment, we do not expect to align the structures with conventional techniques. Instead we plan to use the signals from the structure manifolds to align them relative to the beam. In Section 7.5.2 we discuss this procedure and the emittance growth that is expected based on the tolerances in Figure 7-37.

Ion Effects

In the NLC linacs, ions are created by collisional ionization of the residual gas. Light ions, such as hydrogen, are over-focused and lost between bunches, while in the first part of the linacs, heavier ions are trapped within a bunch train. Trapped ions affect the beam dynamics in three different ways. First, the additional focusing of the trailing bunches due to the ions will lead to an increased filamentation which is insignificant only for a vacuum pressure below 10^{-8} Torr [Raubenheimer 1992]. Second, the ions cause a nonlinear octupole-like coupling of horizontal and vertical betatron motion, whose effect is greatly reduced when the horizontal and vertical phase advances are separated by about 5% [Raubenheimer 1992]. Third, the coupled motion of beam and ions may result in a fast transverse multi-bunch instability of the electron bunch train in the linacs. If the pressure is 10^{-8} Torr, the expected instability rise time at the start of the Linac is about 160 ns [Raubenheimer 1995], assuming 90 bunches of 6.6×10^9 particles each (NLC-I) and carbon monoxide (CO) ions. The beam-ion instability disappears when the ions are no longer trapped within the train. The distance at which this happens depends on the beam current and on the ion mass. As an example, for a bunch train of 90 bunches and 6.6×10^9 particles per bunch, CO ions are trapped up to a beam energy of about 38 GeV, which corresponds to a distance of about 800 m in the NLC linacs. For an average CO pressure of 2×10^{-8} Torr and an initial bunch-to-bunch offset of $0.01\sigma_y$, the expected total dilution of the vertical emittance due to the beam-ion instability is then about 4% (or $\Delta\epsilon_{yN} \approx 2 \times 10^{-9}$ m). At 5×10^{-8} Torr, the dilution would exceed 100%. For higher beam intensities, less ions are trapped over a shorter distance, and the pressure tolerance is looser.

7.5 Alignment, Stability, and Feedback

In this section, we review the methods that will be used to align and stabilize the quadrupoles, BPMs and structures, and to control the bunch orbits and energies. These methods will be essential for the successful operation of the linacs. Without them, many of the bare alignment and stability tolerances could not be met.

For control purposes it is useful to consider separately short-term (seconds or less) and long-term (hours) variations in the linacs. While feedback systems will certainly be needed to control short-term variations, they will stabilize long-term changes as well. However, it is expensive to employ them everywhere, so one generally uses more invasive beam-based techniques, such as the "autosteering" and "phase-golding" procedures in the SLC, when dealing with long-term variations of large portions of the linacs. Thus we begin in the next section with a review of the error sources and their timescales. In particular, we examine the effect of ground motion on luminosity via the quadrupole vibration it generates.

We then turn to a discussion of the alignment of the quadrupoles and accelerator structures. The quadrupoles will be initially aligned using conventional methods. Achieving the required accuracy for long length scales ($\lambda > 200$ m) will be particularly important since beam-based methods are likely to be limited by systematic errors in this regime. As a prerequisite to the beam-based quadrupole alignment, the offsets of the BPMs with respect to the quadrupoles will be determined by changing quadrupole strengths and measuring the resulting orbit changes. The BPM electronics

will be designed to be stable over tens of hours to minimize the frequency of this procedure. Using the measured offsets, a beam-based quadrupole alignment algorithm will be run as a slow feedback loop ($\sim$30-min timescale) to limit dispersive emittance growth. As a fall-back position, beam-based correction schemes that are more invasive to machine operation will be used (*e.g.*, "dispersion-free" and "wake-free" steering).

For the accelerator structures, the alignment approach is to center them about the beam trajectory using as a guide the signals induced in their manifolds during the passage of the beam. This procedure, which we expect will be accurate to about $15\ \mu$m, will be iterated with the quadrupole alignment algorithm. Further discussion of the use of the DDS manifolds for measuring position of the structure relative to the beam is given in Section 7.10.3. In addition, or as a fall-back position, orbit bumps may be used to control the single-bunch emittance growth. Also, as discussed below, we will likely use multibunch feedback based on a fast kicker to control the projected emittance of the multibunch beam.

After the discussion of alignment procedures, we next consider how the beams will be stabilized using various feedback systems. We first discuss "fast transverse feedback" loops which will be used to stabilize the linac beam trajectory to ground motion and other disturbances that occur on timescales of a second and longer. These loops will measure the average properties of the bunch train with about a 100-nm resolution, and will reside in the five linac diagnostic stations. Information about corrections implemented at one station will be "cascaded" to subsequent stations, as is done in the SLC. As a fallback scheme to deal with ground motion, an expensive but realizable procedure is to implement feedback on each quadrupole based on information from a seismometer at the quadrupole.

Next we discuss bunch-by-bunch transverse feedback. Fast BPMs will be used to resolve transverse offsets at the sub-train level, and a fast kicker will be used to realign the bunches within a train. The main error source in this case is believed to be misalignments within and between accelerating structures, so the feedback loop will likely be operated at a relatively slow rate.

Some of the issues related to energy and energy spread have already been discussed, *e.g.*, manipulating the energy spread by the choice of rf phase profile along the linacs (Subsection 7.4.3) and the compensation of the heavy multibunch beam loading (Subsection 7.4.5). We conclude this section with a discussion of energy and energy spread feedback correction at the diagnostic stations. This last subsection also includes a discussion of linac energy management (LEM), which refers to the rescaling of the quadrupole magnet strengths when changes in the energy profile along the linacs are detected.

7.5.1 Error Sources and Timescales

For the NLC, it is important to know both the size and timescale of the sources of beam motion so that an efficient partition of feedback and alignment-type methods can be used. Data from the SLC linac operation gives us some measure of the quantities of interest. For example, attributing the Sector 18 position feedback corrections to the motions of quadrupoles between Sector 18 and Sector 11, where the next upstream feedback loop is located, yields a 0.3-μm to 0.6-μm-rms position variation per quadrupole on an hour timescale, and a 1.0-μm to 2.5-μm-rms position variation on a day timescale. These values are several times larger than can be tolerated in the NLC linacs if feedback is used only in the instrumentation regions. However, much of the apparent quadrupole motion in the SLC is likely to be the result of dispersion. The measured variations in the betatron phase advance between Sectors 11 and 18 correspond roughly to a 0.2% (0.6%) rms energy variation on an hour (day) time scale. Given the size of the quadrupole misalignments and the orbit bumps, these energy changes could produce orbit variations of the sizes observed. Thus we would need to suppress the dispersion in the SLC to be able to observe the level of quadrupole motion of interest. Until we get a better handle on the temporal characteristics of such motion, we will design correction systems that are as fast as possible.

As for fast beam motion ($f > 1$ Hz), the vertical beam jitter at the end of the SLC linac is about 40 times larger than can be tolerated at the ends of the NLC linacs. Part of the SLC beam motion is due to injection jitter; however, if one subtracts the component of jitter that is correlated with incoming beam motion, generally more than 70% remains. In studies of the jitter, we have observed that both the incoming and additive components have a white-noise-like spectrum above 1 Hz, so neither ground motion nor magnet power supply fluctuations are likely to be the cause. Also, the jitter is not strongly correlated with bunch intensity or energy variations. Understanding the sources of this jitter remains an important goal, for if they are present in the NLC, trajectory feedback will be of little help in this frequency regime, as is the case in the SLC.

Vibration and Ground Motion

At a minimum, there will be a trajectory feedback system in each of the five instrumentation regions of the linacs (see Section 7.5.3 for more detail). In each region, BPMs with resolutions of about 100 nm will measure the average bunch position in the train, and the information will be used to control the beam trajectory on subsequent pulses via air-core dipole magnets. How well these loops will suppress betatron motion will depend on the feedback algorithms chosen. Figure 7-38 shows the response functions of two example algorithms that sample the beam orbit at 120 Hz. The "Exp(-i/6) Weighted" routine applies a correction to the next pulse based on an exponential weighting (6 pulse time constant) of the current and previously measured incident beam trajectories. The algorithm is similar to that used in the SLC except that the correction is not delayed by two pulses. (Figure 7-42 shows the SLC response curve including the effect of systematic errors). The "Next Pulse" algorithm applies a correction to the subsequent pulse to remove any orbit deviation measured on the current pulse. It provides more suppression at low frequencies at a cost of a larger amplification ($\times$ 2) at high frequency. While the SLC loops sometimes oscillate, we believe that this problem will be solved to the degree that will allow us to achieve at least a factor of 10 suppression at 0.1 Hz in the NLC, even when including the complications associated with cascading the loops. We will also measure the stability of other algorithms to see if they would be practical to use as well.

If the ground motion in the NLC linac tunnel is comparable to that observed under "quiet" conditions in the SLAC linac tunnel, then its impact on the NLC luminosity should be small when the effect of trajectory feedback is included. To see this, we compare the measured ground motion spectrum with the tolerances. Figures 7-39 and 7-40 show the measurements of the differential and integrated ground motion spectra, respectively, that were made at SLAC as part of the ground motion study discussed in Appendix C. The power spectrum is actually the average of 10 spectra measured in Sector 4 of the SLAC linac on different days at 2 AM. The average power curve was used to compute the integrated spectrum which is the vertical rms ground motion that would be measured if only frequencies above the frequency plotted contributed to the motion (the plateau in the spectrum near 10 Hz results from the resonant vibration of the accelerator structure support system). Also shown in this figure are the ground-motion tolerances described in Section 7.4.4, "Single-Bunch Transverse Phase Space", for pure wave-like motion ("Waves"), and for wave-like plus uncorrelated motion ("Waves + Noise"), where as a worst case, an uncorrelated ground motion component equal to the seismometer electronic noise contribution to the measured power spectra is included. In addition, tolerances that include the effect of trajectory feedback, "Waves + Noise + FB", are plotted. They were generated by dividing the "Waves + Noise" result by the suppression obtained with the SLC-like "Exp(-i/6) Weighted" feedback algorithm. At frequencies where the integrated spectra are rapidly decreasing, one can compare the integrated spectra directly to the tolerance curves since most of the rms motion at any particular frequency is dominated the motion near that frequency.

The net effect on the luminosity is obtained by integrating the power spectrum weighted by the inverse square of the tolerance. For $f > 0.01$ Hz, this yields a luminosity reduction of 0.12% for the "Waves" case, 20% for the "Waves + Noise" case, and 0.13% for the "Waves + Noise + FB" case. Thus, the addition of feedback will still yield a small luminosity loss in the worst case scenario where uncorrelated ground motion is "hiding" just below our seismometer noise limits. Feedback will also be important for suppressing the non-wave-like components of ground motion that are significant at frequencies below 0.01 Hz.

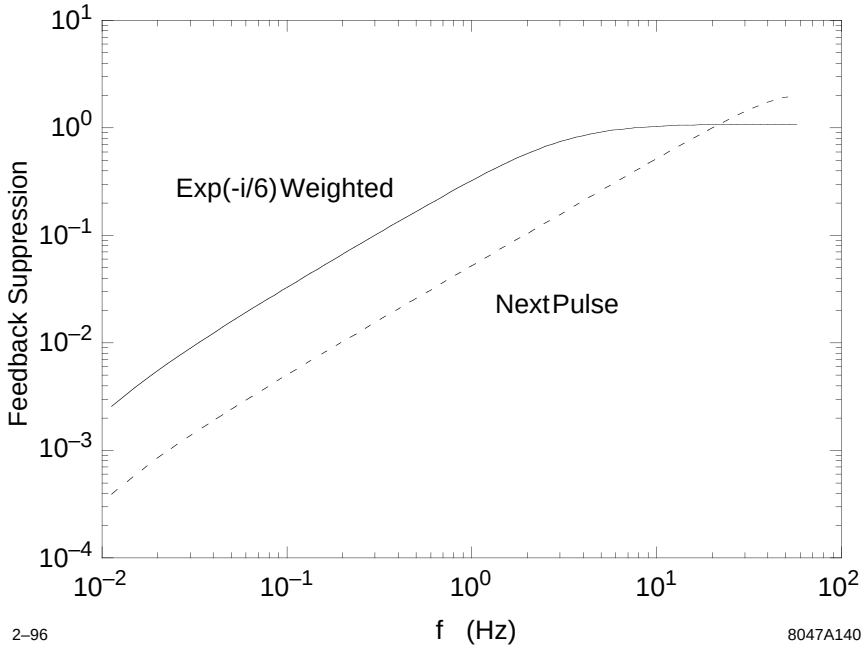


Figure 7-38. *Feedback response curves for an algorithm that exponentially weights the previous pulses (solid) and for an algorithm that makes a 100% correction on the next pulse (dashed).*

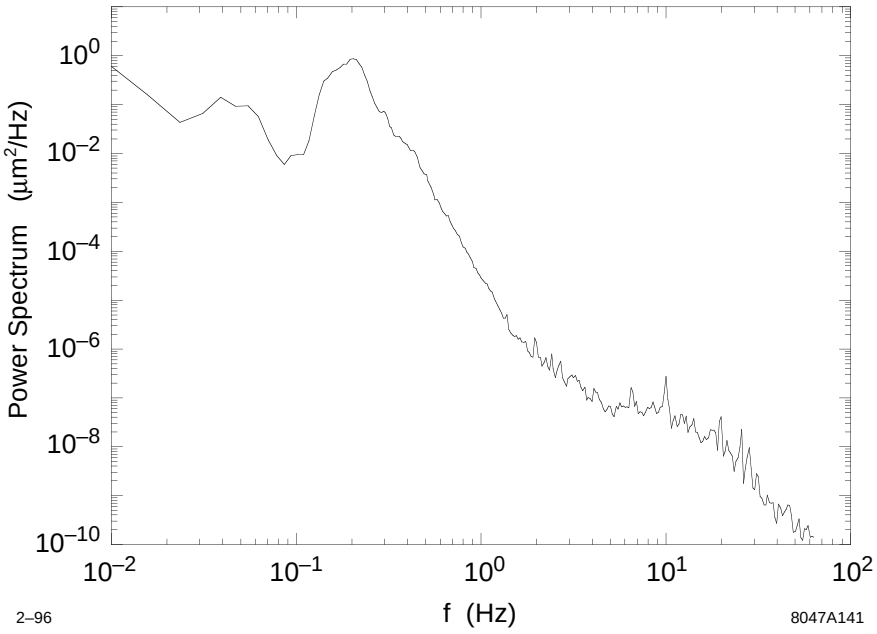


Figure 7-39. *Average of 10 vertical ground motion power spectra that were measured on different days in Sector 4 of the SLC linac at 2 AM.*

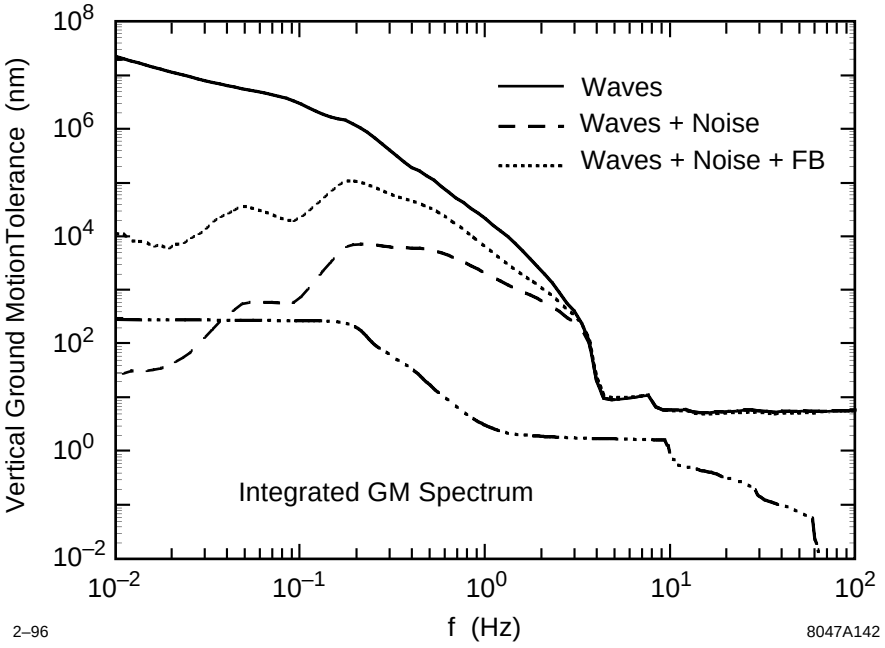


Figure 7-40. *Integrated vertical ground motion spectrum (dot-dashed) and tolerances on the vertical rms ground motion for a 1.5% luminosity reduction for three cases: pure wave-like motion (solid), wave-like plus uncorrelated motion equal to the seismometer electronic noise (dashed), and feedback suppressed wave-like plus uncorrelated motion (dotted).*

The contribution to quadrupole vibration from "cultural" sources is hard to estimate since a site for the NLC has not been chosen. However, if the daytime activities at SLAC are typical of what will occur around the NLC, then the effect on quadrupole motion should not be large since the daytime tunnel measurements at SLAC generally do not exceed the tolerances in Figure 7-40. Also, vibration due to quadrupole cooling and support appear not to be a big problem. Seismometer measurements that were made on top of the quadrupoles and on the floor in the FFTB show that there is only a few nm of relative motion ($f > 0.1$ Hz) between them when the quadrupoles are powered.

As part of evaluating an NLC site, ground motion measurements will be made to obtain the power and correlation spectra. If the ground motion is much larger than that presented above, or if the motion is less correlated over long distances, active types of damping could be used to suppress the effect of the motion on the beam (passive damping would probably be difficult to implement below 10 Hz). For example, the vertical quadrupole vibrations could be measured with seismometers and fed back to fast corrector magnets that would offset the quadrupole kicks. A feed-forward scheme could also be used at the end of the linacs where the measured position of the first bunch in the train would be used to correct the motion of later bunches in the same train. However, limitations due to bunch filamentation would only loosen the quadrupole vibration tolerances by about a factor of three in this scheme.

7.5.2 Alignment Procedures

Quad and BPM Alignment

The initial placement of the quadrupoles and structures will be done with conventional surveying techniques. Most likely, triangulation and leveling methods will be used in combination with Global Positioning System (GPS) data from

satellites. The long-range alignment accuracy will be within the dispersion related tolerances shown in Figure 7-24. However, at wavelengths less than a few hundred meters, the accuracy will level off to something that approaches 100 μm on a 20-m scale. At these wavelength scales, a beam-based approach will be used to control the quadrupole alignment. Beam-based methods tend to be sensitive to systematic errors at longer length scales, which is why we will rely on the initial placement accuracy to achieve the long-wavelength tolerances. As noted earlier, if the displacements due to ground settling significantly exceed the long-wavelength tolerances, a global realignment of the beam line will be done. Hydrostatic monitors may be installed along the beam line to monitor these displacements.

The beam-based quadrupole alignment will use as input the measurements of the average bunch positions from the BPMs that will be located in the quadrupoles. Various beam-based algorithms have been shown to work in theory, although there are many subtleties involved because the derived quadrupole offsets have correlations which vary with length scale. Thus one does not speak of the level of alignment achieved, but instead judges the method in terms of the amount of dispersion remaining on average after alignment. Given that there is some uncertainty as to how often the dispersion correction will have to be made, the correction procedure should be designed to be as fast as possible. Also, the correction will be made by moving the quadrupoles with mechanical movers as opposed to using dipole corrector magnets to offset the quadrupole magnetic centers. This will be done in part because a large investment in mechanical movers will be made for the structures, so adding additional movers (about 14% more) for the quadrupoles will be straightforward. Also, having the beams nearly centered in the quadrupoles alleviates the need for a tight regulation of the quadrupole power supplies, and a tight limit on higher harmonics of the fields.

Given these considerations, we plan to use a quadrupole alignment algorithm that requires only beam trajectory data taken with the nominal linac lattice. Having readings from N BPMs in N quadrupoles will allow us to determine $N-1$ quadrupole offsets, with the alignment end-points being the beam position in the first quadrupole and the BPM zero of the last quadrupole: the outgoing beam angle is unconstrained. Using this algorithm, we will align N-1 quadrupoles at a time, one group after the next, similar to the way that "autosteering" is done in the SLC linac. However, some global limitations in the quadrupole moves would have to be made to prevent the mechanical movers from "ranging-out."

This method will require BPMs with resolutions of about a micron. Such BPMs already exist, however they must also be aligned to the quadrupole magnetic centers to this same level (see reference [Adolphsen 1993] for more detail). This alignment can also be computed in a beam-based manner, but may be disruptive to normal operation. For example, the offsets could be derived from the BPM data taken in which the strengths of individual or strings of quadrupoles are changed. However, the changes in the quadrupole strength would have to be fairly large—turning off one quadrupole yields a betatron oscillation with an amplitude that is only a few times the size of the beam offset in the quadrupole. Thus we want to minimize how often we have to perform this procedure. A calibration scheme that injects signals into the BPMs, and not just into the readout electronics, should help to stabilize the offsets. This is discussed in Section 7.10. Tests of the BPM stability in the SLC and FFTB will be done to study this problem.

Due to the effects of wakefields, it is hard to accurately estimate by analytical means the emittance growth that remains after applying the quadrupole alignment algorithm. Instead we simulate the application of this algorithm to the linac; the methodology and results are presented in Section 7.6.

Structure Alignment

As described in Section 7.3.3, the accelerator structures are supported in pairs on common strongbacks, which themselves are mounted on structure movers. The alignment of the structures will be based on measurements of the beam position relative to the structure axes that will be obtained by analyzing the dipole-mode signals from the structure damping manifolds. Although the original motivation for adding the structure manifolds was to damp the dipole modes, their capacity to be used as a BPM with micron-level resolution will be essential as well.

In the readout electronics, the signal from two modes near the ends of the structure will be selected with filters and measured so that both the position and orientation of the structure relative to the beam can be determined. The difference in the measurements from each structure pair will be used to check the relative structure alignment on the strongback, while the average of the measurements will be used to remotely adjust the mover positions at the micron level to better center the structure pair about the beam trajectory.

The accuracy of the structure alignment system will depend on the degree to which we can internally align the structures. Since we do not have much experience aligning single structures, let alone preserving the alignment of a pair of structures in an operating environment, we have tried to be conservative in our estimates. We believe that structure pairs can be made straight to better than 15 microns rms, and that this alignment can be maintained over long periods of time. Thus far, we have achieved straightnesses of a few microns for the 40-cell stacks that are built before being combined into 206-cell structures. As we assemble whole structures and monitor their straightness, we will be able to better characterize the misalignments.

For now, we assume 15-micron-rms misalignment. To estimate its effect on the beam emittance, we consider two cases whose contributions sum to the general case. In the first case, we assume that the beam travels along the mean axis of the structure pairs, which suppresses the short-range wakefield effects. This leaves the growth due to the long-range wakefields which depends on the spatial correlations of the misalignments. This dependence is shown in Figure 7-37 where the tolerances for piece-wise misalignments of the structures are plotted as a function of the length of the pieces. From these tolerances, a 20% vertical emittance growth is expected if the offsets are uniformly distributed in piece length. This growth will be larger if the offsets occur mainly on a length scale equal to 1/6 of the structure length. To be conservative, we have budgeted 40% emittance growth for these long-range wakefield effects.

In the second case, we assume that the structure pairs are perfectly aligned internally, and consider how well the beam can be aligned to the mean axis of the pair, which was assumed to be done exactly in the first case. Since we expect that the resolution on the beam position derived from the structure signals will be a few microns or better, the alignment accuracy will be dominated by the 15-micron-rms structure misalignments (note that we consider the internal misalignments for this purpose even though they are ignored in this case for the wakefields that they generate). Having four measures of the beam position relative to the axis of a structure pair yields a 7.5-micron-rms beam-to-structure alignment. The resulting vertical emittance growth, which is dominated by short-range wakefield effects, is about 50% based on the Figure 7-37 tolerance for an alignment length of two.

In this analysis, we have ignored systematic distortions of the structure pairs. However, they will likely occur on length scales where the internal alignment tolerances are large. For example, if the pairs bow with a 100-micron sagitta, the emittance growth due to long-range wakefield effects will only be a few percent. The short-range wakefield effects, which depend on the average bunch offset in the structures, would be suppressed in this case since the strongback would be aligned based on the beam offset measurements at the ends and middle, which would center it about the beam.

During the commissioning of the NLC, the structure alignment will be done with a low-intensity beam so as to minimize the wakefield steering effects (*i.e.*, one does not want large orbit changes after moving one pair of structures). During normal operation, it is likely that the structure alignment procedure will have to be iterated with the quadrupole alignment algorithm since moving the structures will kick the beam. If only the power in the wake signal is measured, a scan of the structure position versus signal would be needed to determine the direction to move the structure, which could significantly slow the alignment process. Prototype tests of a structure alignment system will be done in ASSET and the NLCTA to help resolve such issues.

7.5.3 Fast Transverse Feedback

In this section we discuss the feedback systems in the main linacs which operate on relatively fast timescales. The primary focus is on the transverse feedback loops which correct the average trajectory of the beam at the diagnostic stations along each linac, but the general principles are applicable to other loops as well.

A generalized feedback system has been developed and commissioned in the SLC, and has been essential to high-luminosity operation. Operational experience has provided valuable information on feedback-response characteristics, especially with regard to machine and feedback-system imperfections. The proposed NLC feedback is based on the SLC design, with some changes based on lessons learned from the SLC and performance requirements of the NLC. Appendix D provides more information on the design and characteristics of the feedback system. This chapter concentrates on design and performance issues for the NLC linacs.

Linac Feedbacks

In each linac, there are five diagnostic sections, each with a feedback loop. Each loop includes measurements of 10 BPMs for launch control, in addition to one or two BPMs which provide energy information. Each feedback loop calculates and controls five states: the X and Y positions and angles, and the beam energy. To provide feedback control of the launch parameters, two X correctors and two Y correctors are needed. For the energy control, klystron phase or amplitude controls may be used. The feedback design assumption is that the actuators can change to a new setting within one 180-Hz pulse, or 5.55 milliseconds. An additional beam pulse period is allowed for BPM digitization, feedback calculations, and communication. The result is that, after a beam perturbation is detected by the feedback system, the feedback begins to correct it two 180-Hz pulses later. For the initial NLC design, the correction begins two 120-Hz pulses after a perturbation.

Feedback System Design

The feedback design is based on the state space formalism of digital control theory. It uses a Linear Quadratic Gaussian (LQG) controller, which has a built-in model of the beam transport and the expected beam and BPM noise characteristics. Matrices are designed and calculated offline and downloaded for use by the real-time feedback system. The mathematical design goal for the feedback matrices is to minimize the rms of the beam states over time, given the noise inputs. By modifying the input beam noise design assumptions, it is possible to tune the feedback performance response characteristics. The initially proposed feedback algorithm does not adapt to modeling or noise spectrum changes, consistent with the current SLC design. It is hoped that in the future, adaptive methods may provide improved performance, but there is insufficient experience to make design assumptions with them at this point. The beam transport model within a single loop (*i.e.*, R12s which take beam position monitor readings and corrector settings to beam positions and angles) may either be derived from the accelerator model, or may be measured by an invasive online beam-based calibration method.

In the linacs, a series of feedback loops are controlling the same beam parameters (positions, angles, and energy) at different points along the beam line. If the loops do not communicate with each other, this would result in overshooting of the feedback corrections and excessive amplification of beam noise. To avoid this problem, a "cascade" system is developed. The goal of the cascade system is that each feedback loop responds only to perturbations which are generated immediately upstream of it. If the beam transport is approximately linear, this can be accomplished by having each loop send its calculated states to the next downstream loop. The downstream loop mathematically transports the upstream states to its own fit point, then subtracts them from the positions and angles calculated from its own BPMs. This scheme is depicted in Figure 7-41. With this design, each upstream feedback sends information

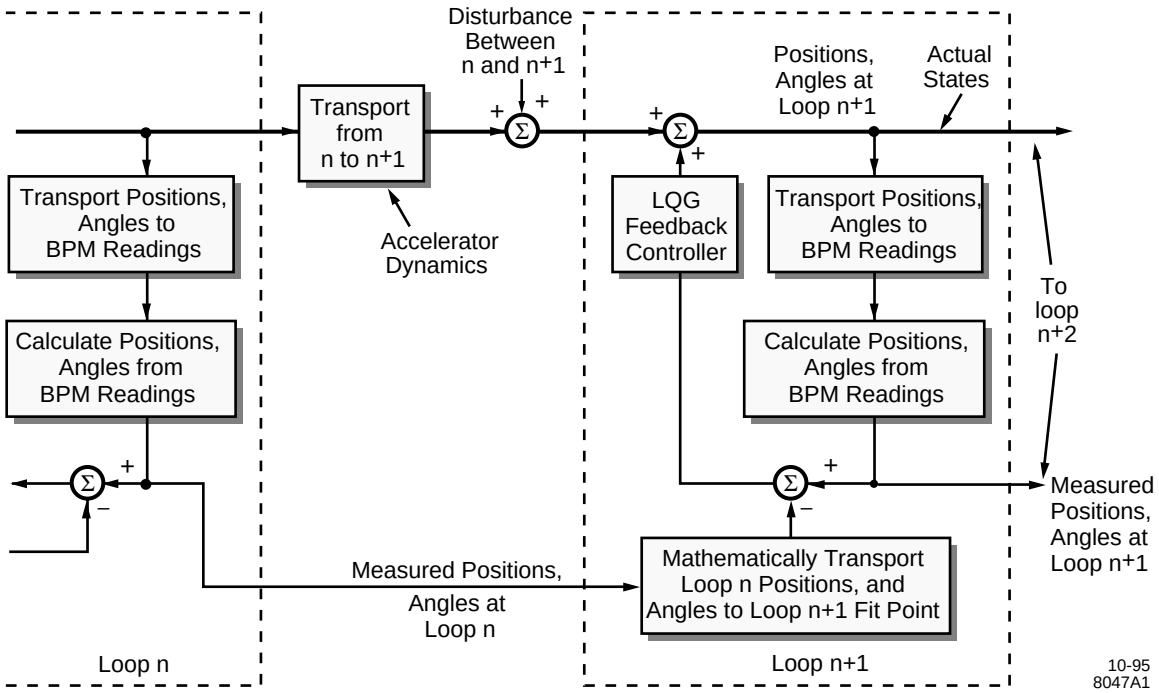


Figure 7-41. *Schematic of cascaded feedback system.*

only to the nearest downstream loop. The key is that each loop is sending the next loop information about what it measured in the beam, not what it is correcting. For example, if there is a perturbation upstream of Loop 1, this perturbation should be seen by all of the feedback loops. Loop 4, for example, will see it, and send this information to Loop 5, which will subtract the perturbation from what it has, so Loop 5 will correct nothing. Loop 5 doesn't know or care whether Loop 4 is correcting the perturbation, it just knows that someone upstream is taking care of it, and actually Loop 1 will correct it. This scheme assumes that the beam transport is linear and independent of the source of a perturbation. (If the NLC transport is excessively nonlinear, it may be advisable to consider a more complicated communications algorithm, where downstream loops receive information from all upstream loops and have sufficient information about the source of a perturbation.)

An important issue in the cascade design is how to mathematically transport the beam from one feedback loop's fit point to the next loop. The distance between feedback loops is so great that the accelerator model is not good enough, and SLC experience has shown that the interloop beam transport changes with time. In the SLC, adaptive techniques are used for the mathematical transport calculations, but the adaptively calculated transport matrices have had some performance problems which degrade the feedback system performance. These problems are described fully later in this chapter and in the feedback appendix. As a result, the proposed method for interloop mathematical beam transport for the NLC is to steal a small number of beam pulses each hour to measure the beam transport by perturbing the beam. It should be possible to provide an adequate measurement by producing a single set of perturbations upstream of the series of feedback loops, and measuring the interloop transport elements for all downstream loops at once.

Operational Performance Issues

SLC feedback performance studies have included five main areas of interest: slow correctors, downsampled feedbacks, poor modeling of beam transport within a feedback loop, poor cascade modeling of interloop beam transport, and optimization of the feedback design response. Slow corrector response has had a detrimental effect on SLC feedback performance, but a new design now allows response within a single beam pulse. In the SLC, due to CPU and BPM constraints, many feedback loops operate at lower than the full 120-Hz beam rate. This degrades the feedback response, and increases aliasing of higher frequency noise to lower frequencies. For the NLC, sufficient CPU and BPMs will be available so that the feedback loops may run at the full machine rate. The remaining performance issues are more challenging, and are interrelated.

While it should be possible to develop a nearly perfect model of the accelerator beam transport within a single feedback loop, SLC experience has shown that the model has imperfections in some cases. Performing the calibration procedure for several SLC feedback loops has provided data to quantify the effects of typical imperfections (these involve errors in the R matrices between correctors and states of 10–20%). NLC design simulations include consideration of comparable imperfections.

If the long-range beam transport in the NLC linacs is reasonably linear and not changing quickly with time, then measuring the interloop transport elements every hour should provide perfect cascade correction, and each loop should perform corrections only on perturbations generated immediately upstream. If this were true, the global system feedback response would be identical to the response of a single feedback loop. But to be conservative and account for small nonlinearities and changing beam transport properties, NLC feedback simulations assume that part of a given beam perturbation leaks through to the downstream loop, so that each loop incorrectly compensates for a portion of an upstream perturbation. In the simulations, it is assumed that each loop compensates for 15% of perturbations which are not immediately upstream.

The SLC design's noise response results in correction of a beam perturbation exponentially with a time constant of about six pulses. In other words, after six pulses, most of an incoming perturbation is fixed by the feedback. This design provides a good response to a step function and provides a stable feedback which does not over-react to white noise. For the NLC, it is preferred to have a feedback response which is tuned to damp strongly at low frequencies. In the NLC, high-frequency noise is expected to be damped by passive systems and the expected input noise spectrum would be dominated by low-frequency ground motion. Note that while it is possible to tune the feedback response curve, any design will damp at some frequencies while amplifying at others. Several alternative noise designs have been simulated and evaluated. While the more aggressive designs, which damp strongly at low frequencies, provide an excellent response curve under perfect conditions, these designs are more sensitive to imperfections in both intraloop modeling and interloop cascade transport. To provide a stable, robust feedback system with better response at lower frequencies, more work must be done. Figure 7-42 shows the feedback response curve for the SLC feedback design, which includes the effects of imperfections. Note that this curve is for a 120-Hz beam pulse and feedback rate. If the rate is 180 Hz, the response curve would be similar but shifted in frequency. At this point, the SLC design is proposed as a conservative initial approach, which provides adequate response and is not unduly sensitive to machine and modeling imperfections.

7.5.4 Bunch-by-bunch Transverse Feedback

The fast-feedback systems described in the preceding section operate on averages of the bunches in a train. The repetition rate of these loops is relatively fast. We turn next to loops which distinguish variations within a bunch train, but which perform corrections less frequently than the “fast” feedback loops previously described. It is probably not

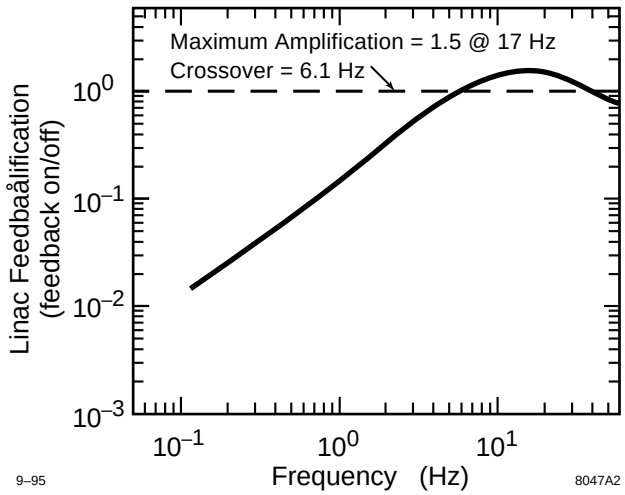


Figure 7-42. *Feedback response curve for the SLC feedback design with a 120-Hz beam pulse and feedback rate, including the effects of imperfections.*

necessary to resolve all the bunches individually, but it is necessary to have some resolution within the train in order to control the multibunch emittance (and also bunch energies and energy spread, as is discussed later in this section).

These bunch-by-bunch feedbacks will also be implemented in the five diagnostic/control sections along each main linac. Measurement of the individual transverse bunch positions will be at these stations using special fast BPMs. Trajectory correction will be done using high-bandwidth kickers, able to tailor the kick delivered to individual bunches assuming the variation is fairly smooth along the train. Simulations yielding the requirements on BPM resolution, kicker strength, and kicker speed are given in the next subsection, and a design for a high-speed kicker suitable for use in such feedback loops is given in 7.9.4.

Multibunch Kicker Simulations

In the multibunch kicker simulations, bunch-by-bunch trajectories were assumed to be measured by two fast BPMs located at consecutive vertically-focusing quadrupole magnets (90° phase difference) at five locations in the NLC-I linacs: 30 GeV, 60 GeV, 100 GeV, 150 GeV, and 250 GeV; these positions closely correspond to the locations of the diagnostic stations. Two fast kickers were also located at consecutive vertically-focusing quadrupoles just upstream of the BPMs and the bunches were kicked to line all bunches to the trajectory of the first bunch. One-to-one trajectory correction was also performed to ensure that the beam centroid passes through the center of every focusing quadrupole magnet.

The tracking was performed with 90 bunches including both single- and multi-bunch effects. Each structure was divided into seven pieces and the pieces were aligned independently with rms errors of 25 μm. This misalignment was chosen so that the multi-bunch effects are much stronger than single-bunch effects because we are concentrating on multi-bunch correction. But, because it is possible that single-bunch effects become significant due to a large bunch offset caused by the multi-bunch effects, the single-bunch effects were also considered in this simulation. The expected emittance growth without corrections was simulated to be $\Delta\epsilon_y/\epsilon_y$ = 133%.

Figure 7-43(a), (b), and (c) show emittance growth as a function of the kicker strength, BPM resolution, and kicker speed, respectively. The kicker strength of 56 kV corresponds to a deflection of roughly six times the beam size. The residual emittance that remains after the correction is primarily due to the single bunch emittance dilution that arises

from the misaligned structures and the large bunch offsets between correction stations. Finally, Figure 7-44 shows the effect of the kickers in one simulation with 25-μm-rms accelerator structure piece misalignments. The uncorrected emittance dilution was greater than 150% in this case. The residual dilution after correction is roughly 15%.

Given these results, we will design a kicker with a speed of 250 MHz and a strength of 56 kV, and a bunch-by-bunch BPM with a resolution better than 0.2 μm. This is discussed further in Section 7.9.4.

7.5.5 Linac Energy Management and Energy Feedbacks

The energy management of the main linacs includes the measurement of the energy, energy feedback and feedforward, the control of energy and magnet strength settings along the linacs, the handling of vetoes due to klystron faults, control of individual amplitudes and phases of the klystrons, and the confirmation of these parameters with beam-derived data. The energy spread within a bunch, which can be controlled by phase adjustments, is also discussed in this section.

Energy Measurement

The energy of the beam will be measured in high-dispersion regions that provide resolutions of about 0.1% or better. Five diagnostic stations are envisioned for this purpose, as discussed in Section 7.3.4. These stations will also be able to measure the energy variation along the bunch train.

Measurements of other quantities will also be important since they will influence the energy. Among these are the bunch train current and distribution, the beam phase, the state of the upcoming klystron pulses (*e.g.,* OK, veto, or in transition), cascaded information from upstream feedbacks, and any other information (such as history of slow ramps in current and energy) that is relevant to controlling the energy of the bunches.

Energy Correction

The energy correction consists of a system of phase shifters, which will adjust the phase on a pulse-to-pulse basis like the SLC energy feedback loops. The slope and higher orders of the energy profile of the bunch train will also need to be controlled (this is discussed in Section 7.4.5 and at the end of this section).

Depending on the information source, the correction will be a feedback or feed-forward system. These systems should have a very general design with a self-calibrating mechanism. They should also allow auxiliary inputs and outputs (for instance, dumping the beam instead of correcting it).

A feed-forward system like that in the SLC will be necessary for different beam currents. By measuring the charge of each bunch in the damping ring and correcting the rf amplitude by opposite phase settings of two subboosters, good beam loading compensation can be achieved. Due to the high beam loading of about 25%, small current changes of 1% already cause a 0.25% energy change. The feed-forward information needs to be forwarded to the five energy stabilization stations (upstream from the chicanes) and combined with the local feedback.

Linac Energy Management

The management of the energy along the linacs is as important as controlling the energy at the ends, especially if a change in the energy profile might change a tail cancelation into an addition. Switching a klystron is probably the

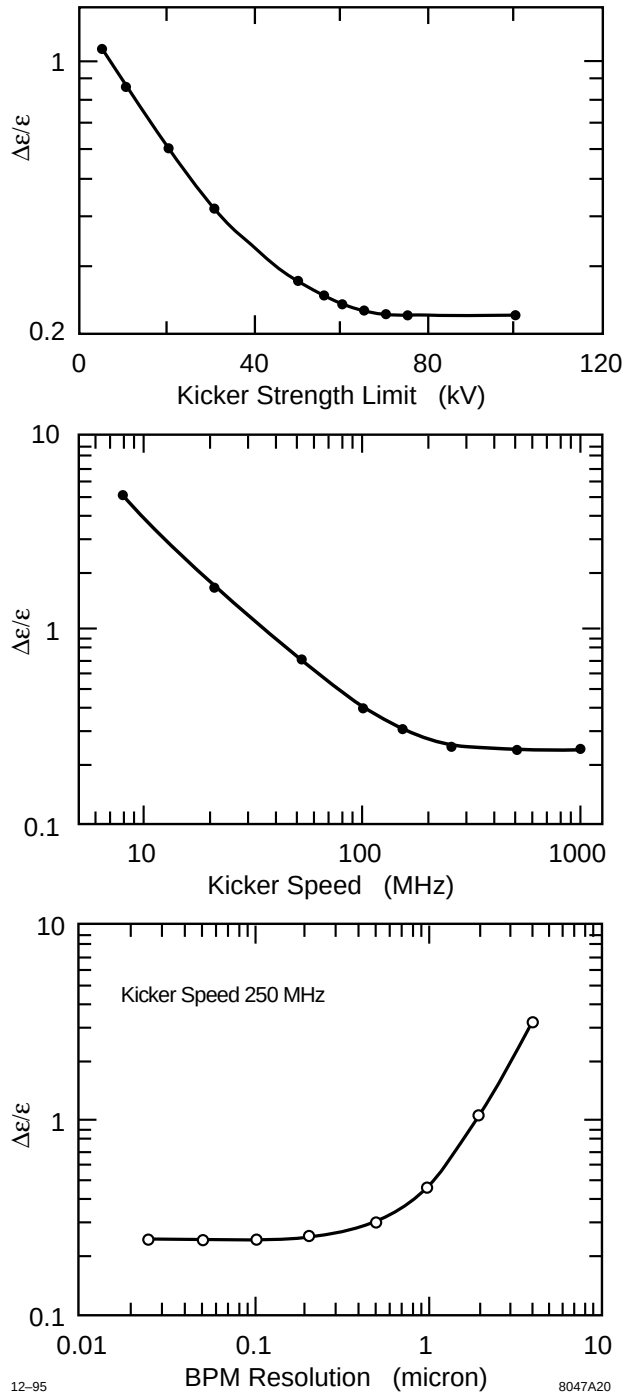


Figure 7-43. *Emittance dilution versus: (a) kicker strength for a kicker speed of 250 MHz and a BPM resolution of 0.2 μm; (b) kicker speed for a kicker strength of 56 kV and a BPM resolution of 0.2 μm; (c) BPM resolution for a kicker strength of 56 kV and a speed of 250 MHz. Each structure was divided into seven pieces with a random misalignment of 25 μm rms. Each symbol represents the average of tracking with 25 random seeds. The residual emittance dilution is due to the single-bunch dilutions arising from the trajectory and structure offsets.*

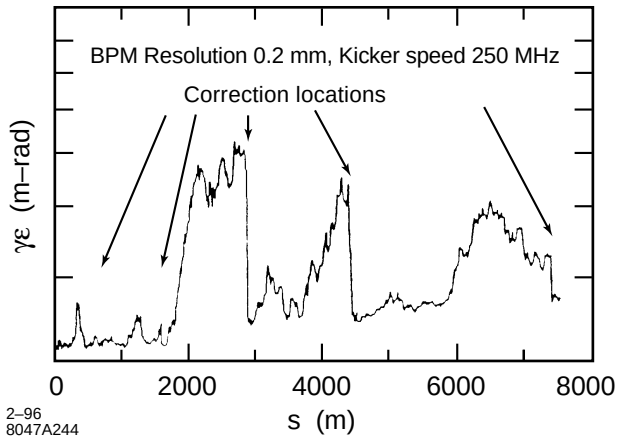


Figure 7-44. *Example of a kicker simulation with 25-μm rms accelerator structure piece misalignments.*

most important single event which requires a linac energy management (LEM) operation, that is, an adjustment of the magnets to reestablish the matching of the beam to the lattice. For this procedure we have to know the local beam energy and the energy gain (amplitude) and phase of the klystrons. The rf parameters should be checked with beam measurements. Also, the lattice should be checked with diagnostic pulses—special beam pulses which are transversely kicked and therefore oscillate down the linacs giving the required lattice information.

Another energy-related problem in the SLC that makes LEM necessary is changes in temperature which can generate energy variations of up to 2%. The linearized LEM adjustments help but are not sufficient so that the above mentioned additional tools will be helpful, especially in the front part of the main linacs. The NLC at 500 GeV is about 10 times less sensitive to energy changes of one klystron (230 MeV) than the SLC. But it has also to maintain 100-times-smaller emittances. One klystron powers two structures (NLC-I) producing 100 MeV, which is 0.1% of the energy at 100 GeV, but 1% at injection where the energy is 10 GeV.

Bunch-by-bunch Energy and Energy Spread Feedback

It will be necessary to control the effects on the energy and energy spread of the multibunch train due to variations in the energy profile delivered by the rf system and to variations in the bunch charges. As discussed previously in Section 7.4.5, multibunch energy compensation can be done by shaping the rf pulse to simulate the steady-state beam loading profile before the beam is injected. Nearly ideal compensation of the beam loading will be obtained by a programmed series of phase shifts during the rf pulse (see Chapter 8 for details). However, there is no reason that the phase variations cannot be modified to obtain a slightly different shape to the rf pulse. A feedback system could thus be designed to fix small variations in energy along the bunch train, assuming the variation is smooth enough that the klystron variability has sufficient bandwidth.

Bunches belonging to a given train in the damping rings ride at different synchronous phases due to the differing amount of beam loading from bunch to bunch. The beam loading in turn depends on the bunch charges, which may vary on a pulse-to-pulse basis. We may have to compensate for such pulse-to-pulse variations by feedforward of bunch charge information from the damping rings to the main linacs.

7.6 Simulation Studies

The previous sections describe the basic design of the NLC main linacs. The results of analytical estimates and numerical simulations were used throughout the text, mainly to discuss and justify specific design choices. In this section, we describe numerical calculations that were performed to study the NLC main linacs in greater detail. By assuming realistic errors in all major accelerator components, we can study the complex interactions between different mechanisms of emittance growth and the proposed correction algorithms. Since stability problems are of major concern for the NLC, we also discuss simulation results for the alignment stability in the NLC main linacs.

In the following, we will briefly describe the simulation program LIAR ("LInear Accelerator Research code") which was recently developed [Assmann 1995]. Then, we discuss the overall objectives for the NLC linac simulations and present the simulation results. A final outlook will present the conclusions and will describe directions for future work.

The goal of the simulation studies done so far is to show the feasibility of the acceleration and emittance preservation in the NLC linacs. Several important problems still do not have their final solutions. However, we will show that we have working solutions to the major problems that already fulfill the requirements.

7.6.1 Simulation Program

The simulation and tracking program LIAR ("LInear Accelerator Research code") is designed as a general tool to study the beam dynamics of linacs. LIAR's major design goals are to provide a simulation platform that includes all the basic accelerator physics for linacs, that allows to add easily advanced features and that is not bound to a specific linac design. Presently LIAR is used both for NLC and SLC simulations. We briefly describe its main features and observables.

All magnets are simulated using the thick lens representation. The beam line description includes "realistic" support points with a number of elements mounted on a common girder. Thus we can simulate the effects of ground motion and we can also study correction algorithms using movers at the support points. Rf structures can be divided into pieces and have an rf-BPM assigned at both ends.

The beam is described as a train of bunches. Each bunch is divided into slices along the longitudinal direction and each slice is described by a number of mono-energetic beam ellipses. The tracking part of the program performs a 6D coupled-beam transport including the beam-induced transverse and longitudinal wakefields in the rf structures.

The basic features of the simulation program are complemented by the ability to set random and systematic errors on most accelerator parameters:

1. Transverse misalignments and roll angles for most element types.
2. BPM misalignments with respect to the quadrupoles. Finite BPM resolution.
3. Strength errors for quadrupoles, bending magnets and correctors.
4. Phase and gradient errors of the rf structures.
5. Random or "ATL"-like misalignments of the accelerator support.
6. Misalignments of individual rf structure pieces ("bowing", etc.).

Given those capabilities we can simulate basic and advanced optimization schemes for the main linacs of NLC. So far the following correction methods have been implemented:

1. One-to-one steering using dipole correctors.
2. Trajectory feedbacks.
3. Beam-based quadrupole alignment.
4. Beam-based alignment of the rf structures.
5. SLC-type emittance bumps with deterministic optimization.

Other optimization schemes like dispersion-free steering or multibunch kickers will be added in the future. All those correction schemes can be applied and tested in the presence of multiple interacting error sources. Furthermore we can study the convergence of these techniques when they are iterated many times.

We use a number of different measures to describe the performance of the linacs. Many observables are available at each BPM so that they can easily be studied as a function of longitudinal position s:

1. Horizontal and vertical beam and bunch positions with respect to the design plane or the BPM centers (by beam we refer to the average of the bunches).
2. Horizontal and vertical emittances for all bunches.
3. Horizontal and vertical luminosity reduction for all bunches.
4. Energy spread of the beam.
5. Beam energy.
6. Horizontal and vertical B_{mag} (see Equation 7.3) for a selected bunch.

At every marker point the beam distribution in phase space is available for a selected bunch and the whole bunch train. Twiss parameters are available for every beam line element. All observables, together with other parameters like misalignments, can be saved into external files which then are further analyzed. LIAR also provides screen summary output about single and multibunch emittance growth, rms trajectory offsets and the other observables.

7.6.2 Objectives

The NLC main linac simulations are undertaken with the general objectives to study the complex interactions between different parameters and to get a complete and coherent picture of the relevant emittance growth and optimization processes. By studying multiple error sources we try to evaluate the main linac performance under "realistic" conditions. Here we give a list of relevant simulation studies.

- Static quadrupole and BPM misalignments and errors (compare Sections 7.4.4 and 7.4.6):
  - Dispersive emittance growth as a function of transverse quadrupole and BPM misalignments.
  - Sensitivity to lattice mismatch from quadrupole strength errors.

  - Sensitivity to betatron coupling from quadrupole rolls.

- Static structure misalignments and errors (compare Sections 7.4.4 and 7.4.6):
  - Emittance growth due to random structure misalignments. Sensitivity to different wavelengths of misalignments.
  - Emittance growth due to systematic structure misalignments ("bowing", etc.).
  - Lattice mismatch from gradient and phase errors.
  - Sensitivity of dispersive emittance growth to changes in the BNS phases.

- Beam tails:
  - Determination of the average beam power that needs to be collimated.

- Beam line stability (compare Section 7.5):
  - Emittance growth from ATL-like drifts of the quadrupole and structure alignment.
  - Emittance growth from jitter and vibrations of the accelerator support.
  - Feedback requirements for effective trajectory stabilization.
  - Effects from bunch-to-bunch variations of charge or initial position and angle.

- Emittance control techniques (compare Section 7.5.2):
  - Efficiency of interleaved quadrupole steering and rf structure alignment in the presence of errors (compare Section 7.5.2).
  - Dispersion-free and wakefield-free steering.
  - Emittance optimization with trajectory bumps.

Several of these studies are interdependent and some need to be done for both single-bunch and multibunch cases. We cannot yet present all the final results. However, most of the problems mentioned above are addressed in the simulations that we are going to present.

7.6.3 Simulation Parameters

The NLC tolerances are tighter for the higher bunch currents and the longer bunch lengths at the higher beam energies. Therefore we restrict the simulations to the 500-GeV version of NLC-II. This is the case "NLC-IIb" in Tables 7-1 and 7-2. The beam consists of 90 bunches, where each bunch contains 1.1×10^{10} particles, is 150-μm long and has an initial uncorrelated energy spread of 1.5%. We describe a single bunch at 10 longitudinal positions each with 3 monoenergetic beam slices. The initial bunch shape is Gaussian. The initial horizontal and vertical beam emittances are $\gamma\epsilon_x = 4.0 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 4.0 \times 10^{-8}$ m-rad. The beam is accelerated to a final beam energy of 500 GeV.

The full NLC-IIb lattice is used for the simulations and we assume that the chicanes in the diagnostics stations are switched off. Beyond that we assume that multibunch beam loading is perfectly compensated. Parametrizations for short-range and long-range wakefields were provided from detailed calculations as mentioned in Sections 7.4.1 and 7.4.2. The simulations that are presented here focus on the vertical plane where the small initial emittance makes it much harder to avoid dilutions.

Config	ϕ_1 (°)	E_1 (GeV)	ϕ_2 (°)	E_2 (GeV)	ϕ_3 (°)
1	4	30	-7	485	-30
2	8	30	-5	455	-30
3	10	30	-3	425	-30
4	12	30	-1	400	-30
5	14	30	1	380	-30
6	16	30	3	360	-30
7	18	30	5	335	-30
8	20	30	7	320	-30
9	22	30	9	300	-30

Table 7-8. *BNS configurations: ϕ_1, ϕ_2 and ϕ_3 are the three rf phases in the linac and E_1 and E_2 are the switching points.*

7.6.4 BNS Configuration

As explained in Section 7.3.2, BNS damping significantly reduces the emittance growth in a linac. If the BNS autophasing condition is fulfilled then an incoming betatron oscillation will propagate downstream without perturbation. In this case, there will be minimal emittance growth. Figure 7-4 shows the autophasing energy spread for a simple two-particle model (compare Equation 7.2). The figure suggests an average autophasing energy spread of about 1.1%.

In order to identify an optimal BNS configuration we studied a number of energy spread profiles along the linac, each of which is generated by adjusting the rf phases in three groups. The BNS configurations number 1 to 9 are defined in Table 7-8. All BNS cases were calculated to give the required final beam energy spread of about 0.3% rms. In Figure 7-45 we show the beam energy spread along the linac for the different cases. Figure 7-46 characterizes the different BNS configurations in terms of average energy spread and the energy overhead required to generate them. The energy overhead is defined as the energy difference between operating with the nominal average rf phase and operating with the BNS phases. In the following we refer to the BNS cases by their BNS energy overhead.

To study the BNS configurations in terms of emittance preservation, we simulated the emittance growth for an initial 1σ vertical beam offset (2.2 μm). The initial uncorrelated beam energy spread was set to zero. As is shown in Figure 7-47 we find the smallest emittance growth for a BNS energy overhead of 1.3%. For lower energy overheads, wakefields cause large emittance growth while dispersive emittance growth dominates for higher BNS energy spreads. Since the BNS overhead of 1.3% is well inside the specifications (up to 3%), we chose the BNS configuration number 3 for all further studies. However, we always made sure that this BNS case is a good working point. Later we will discuss our choice of BNS energy spread for two other emittance growth problems and we will see that the BNS configuration number 3 is still the optimal working point within the cases that are considered. We note that the rms energy spread for this BNS configuration is about half of what is expected for autophasing from Figure 7-4.

If the autophasing condition is fulfilled, then we would expect that the betatron oscillations from the initial 1σ beam offset, normalized with the beta function and the beam energy, propagate downstream without perturbations. Figures 7-48 and 7-49 show the normalized betatron oscillations and the vertical single bunch emittance growth along the linac. The first figure does not include any initial uncorrelated energy spread, while the second case includes the nominal 1.5% initial energy spread. It is easily seen that the amplitude of the normalized betatron oscillation is not maintained in either case. We note that the initial uncorrelated beam energy spread significantly changes the dynamics in the beginning of the linac. It causes strong filamentation and modifies tolerances for incoming jitter.

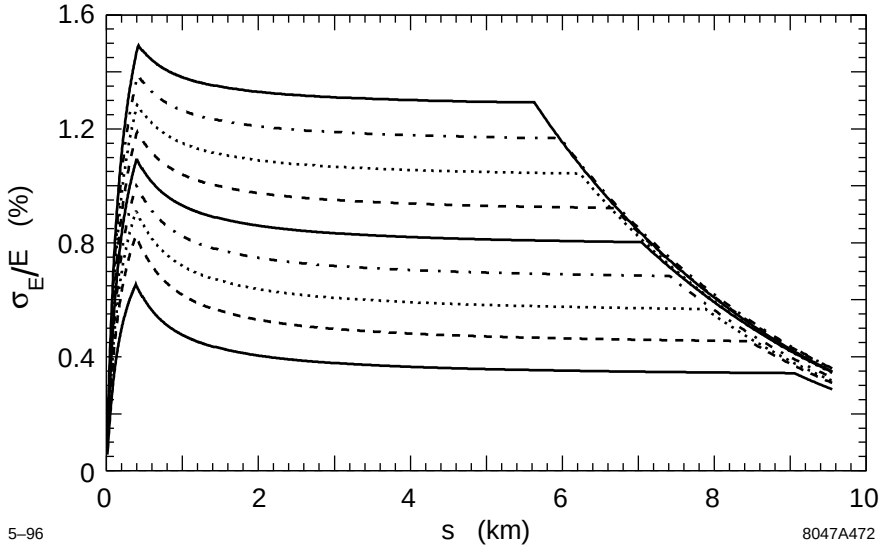


Figure 7-45. *Beam rms energy spread along the linac for different BNS configurations. Note that the initial uncorrelated energy spread was set to zero. The final extracted rms energy spread is fixed to about 0.3% as required by the final focus energy bandwidth. The different BNS configurations are defined in Table 7-8 and are referred to as number 1 to 9, where 1 is the lowest curve.*

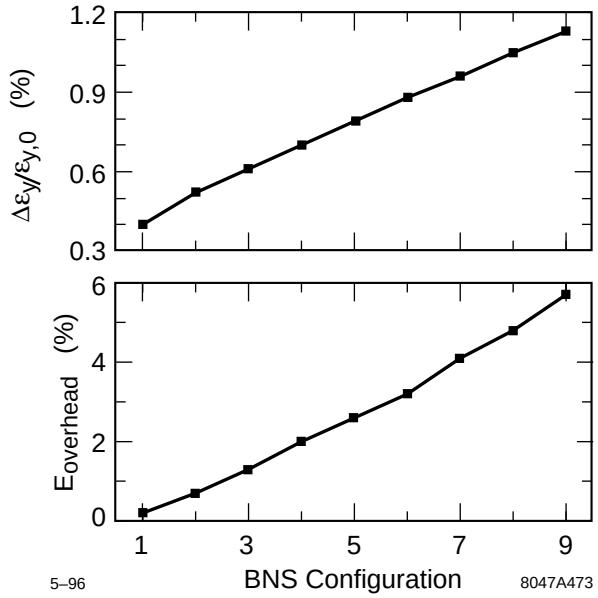


Figure 7-46. *The different BNS configurations from Figure 7-45 are characterized by the average energy spread σ_E / E along the linac and the BNS energy overhead E_{overhead}.*

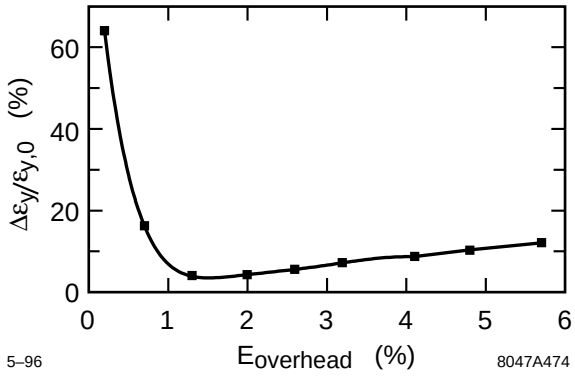


Figure 7-47. *Average vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ for the different BNS configurations from Figures 7-45 and 7-46 and a 1σ (2.2 μm) initial beam offset. The initial uncorrelated beam energy spread is set to zero.*

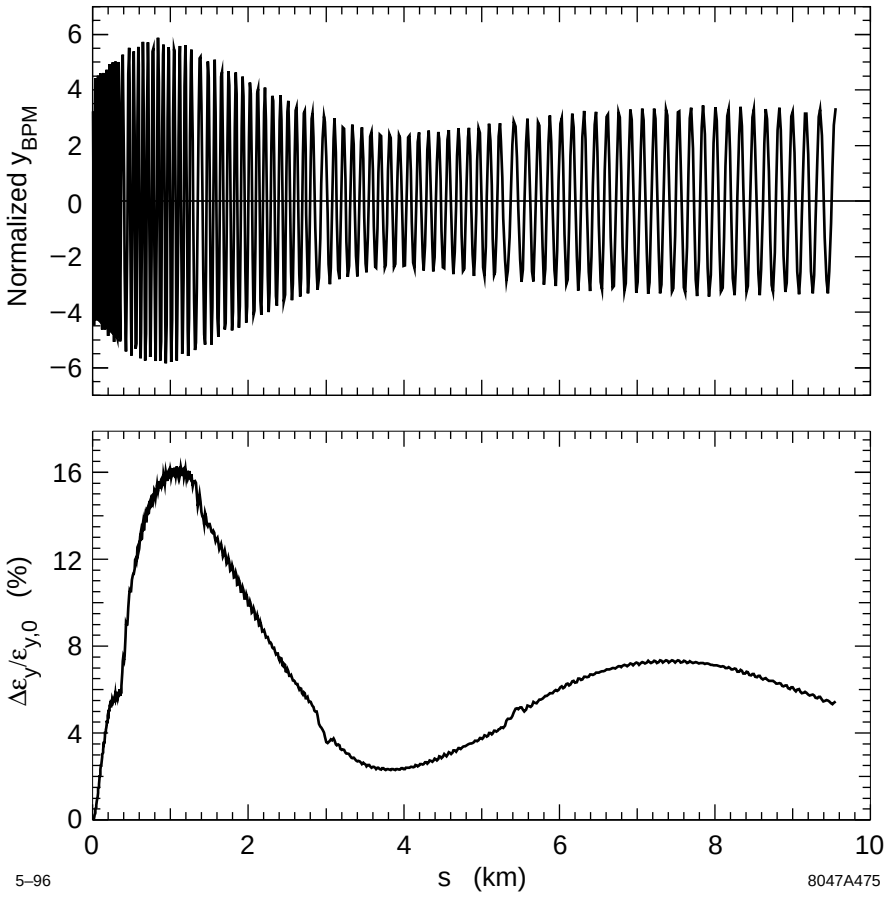


Figure 7-48. *Beam offsets y_{BPM}, normalized with the beta function and the beam energy, and emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ for a 1σ (2.2 μm) initial beam offset and no initial uncorrelated beam energy spread. The BNS configuration number 3 with a BNS energy overhead of 1.3% is used.*

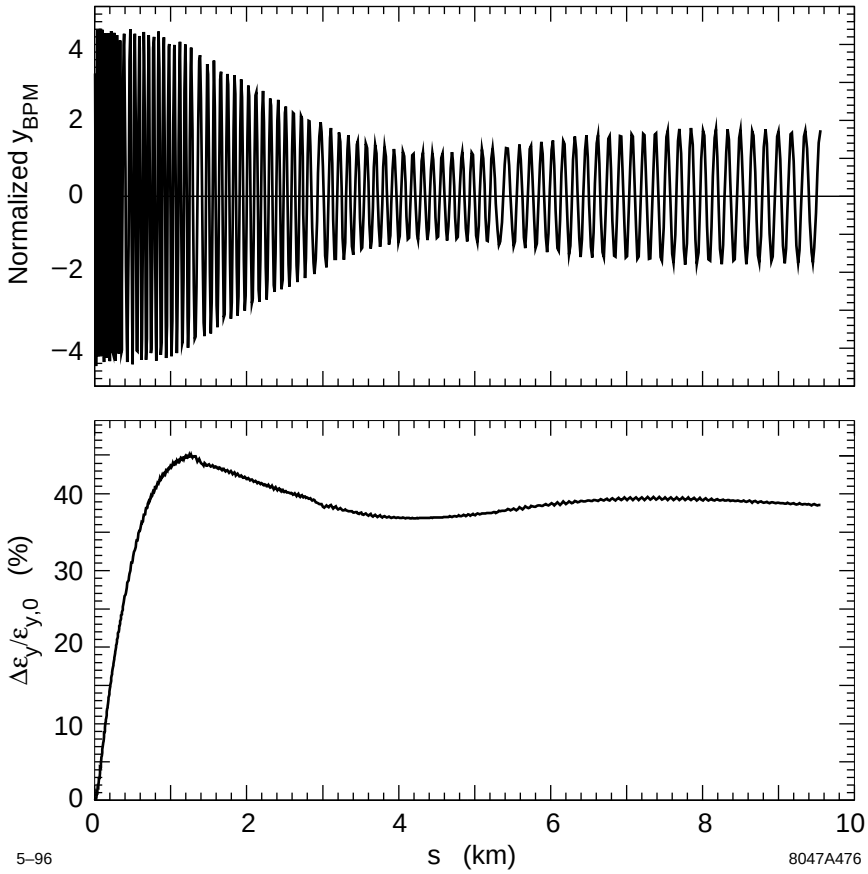


Figure 7-49. *Beam offsets y_{BPM}, normalized with the beta function and the beam energy, and emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ for a 1σ (2.2 μm) initial beam offset. Here the initial uncorrelated beam energy spread is set to 1.5%. Again we use the BNS configuration number 3 with a BNS energy overhead of 1.3%.*

7.6.5 Static Imperfections and Their Correction

Imperfections are called static if they do not change during the typical timescale of beam-based alignment and correction procedures. Initially static imperfections in the NLC linacs are allowed to be large compared to the final tolerances. In order to achieve the required tolerances the "conventional" (not beam-based) alignment must be complemented by beam-based alignment techniques. In this section we explain the algorithm proposed for beam-based alignment and its implementation in the simulation program. We discuss detailed simulations that show that the required precision levels can indeed be achieved, even when many additional error sources are included.

The Beam-based Alignment Algorithm for Quadrupoles and RF-structures

The emittance growth is driven by transverse offsets between the beam and the centers of quadrupoles and structures. Those offsets must be effectively minimized in order to maintain the normalized emittances.

The basic concept for the NLC beam-based quadrupole alignment is to use the available BPM information to solve for the beam-to-quadrupole offsets and the initial beam position and angle. Assuming that all beam deflections are caused by the quadrupoles, we can use N BPM measurements to solve for $N-2$ quadrupole offsets and the initial y, y' of the beam. The positions of the first and last quadrupole in the corrected section are fixed. The endpoint BPMs define the reference line for the alignment. If the alignment is done in many sections then the beam is launched from one section into the other with a single dipole corrector at the border between the two. Therefore, for all but the first section we adjust only the initial y'. In the first section we adjust both y and y' of the incoming beam ("launch feedback"). The quadrupoles are aligned with magnet movers at each quadrupole support. Since the optics model is perturbed by wakefields and to account for other imperfections, the alignment is implemented as an iterative process. Furthermore, the quadrupole alignment is interleaved with the alignment of the rf-structures.

The rf accelerator structures have an rf-BPM at each end. Two rf-structures are always mounted on a single support structure. Every support structure can be moved by stepping motors at either end. Thus, for each support structure a total of four rf-BPMs measure the beam positions in the structures and two movers align the girder. The movers are adjusted such that the average rf-BPM reading on a girder is minimized. The rf-structure alignment is performed after each iteration of quadrupole beam-based alignment.

Here, we assume that the step resolution of the magnet and girder movers is infinitely small. The typical step size of 0.25 μm is indeed small compared to the resolution of the rf-BPMs of about 15-μm rms and can therefore be neglected for the rf-structures. However, for quadrupoles the step size of the movers is an important limitation. The problem is avoided by having dipole correctors at each quadrupole that shift its effective magnetic center. Small quadrupole misalignments are therefore "cured" with dipole correctors. If the dipole strengths get large enough they are "exchanged" for a step of the quadrupole mover.

The alignment algorithm that was sketched above was implemented into the simulation program LIAR. First, we consider a simple case where we start with a random quadrupole misalignment of 100-μm rms and assume perfect BPMs and rf-BPMs. Both kinds of BPMs have no offsets and zero resolution. The quadrupole alignment is done in 14 regions to allow for best convergence. Each region contains about 52 quadrupoles and is iterated 15 times. The number of iterations is chosen higher than necessary in order to explore the optimal solution.

The misalignment of quadrupoles, BPMs and rf structures, after the application of the interleaved alignment procedure, is shown in the upper part of Figure 7-50. The dipole kicks at the boundaries between correction regions are shown in the lower part of the same figure. A very smooth alignment between the endpoints of each section is indeed achieved. At the endpoints the beam is deflected into the next section, causing sharp kinks. The endpoints are not moved and reflect the initial random quadrupole misalignment of 100 μm. Between the endpoints, the alignment is bowed towards zero. The absolute zero line is known to the system only because the initial misalignment was random about it. Constraints on the rms size of magnet movements bias the solution towards the initial average misalignment between endpoints.

The solution shown in Figure 7-50 is largely determined by the different weights on the "measured" BPM readings, the rms of the calculated magnet movements and the strength of the dipole kick that launches the beam into the alignment section. Changes in the relative weights will result in solutions that are not equivalent in terms of emittance growth. We have chosen to constrain the rms magnet movements and the rms of the "measured" BPM readings relatively strongly while allowing for large dipole kicks. This choice of weights results in a small final emittance growth.

The vertical single-bunch emittance growth and the trajectory along the linac are shown in Figures 7-51 and 7-52. The trajectory in Figures 7-52 shows residual sub-micron beam offsets in the quadrupoles that result from the "bowing" of the alignment trajectory. Figure 7-51 shows that these small beam offsets cause no significant emittance growth. The residual emittance growth occurs mainly at the transitions from one sector to the next. The large dispersive deflections cause a stepwise emittance growth behavior. One can imagine smoothing the transitions by distributing the beam

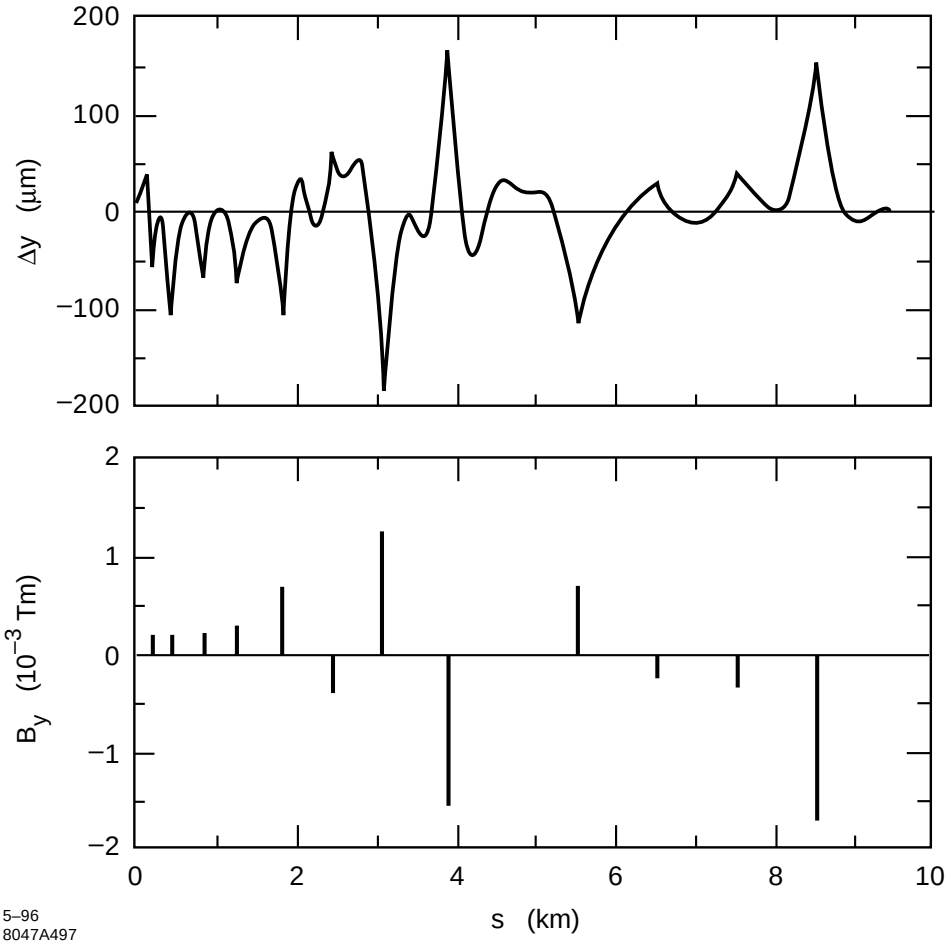


Figure 7-50. *Example of the beam-based alignment algorithm with perfect BPMs and rf-BPMs. The initial random quadrupole misalignment was 100-μm rms. The alignment is done in 14 sections and 15 iterations. At the end of each section a dipole corrector is used to launch the beam into the next section. The upper plot shows the misalignment Δy of quadrupoles, rf-structures and BPMs after alignment. The lower plot shows the integrated fields of the dipole correctors.*

deflection. A better method would be to also move the sector endpoint quadrupoles such that the deflection kicks are minimized.

We should stress that the present algorithm already works very well. The quadrupole and structure alignment is minimized over short wavelengths and beam offsets with respect to the BPMs are effectively reduced from many cm's to the sub-micron level. The smooth alignment of quadrupoles and structures minimizes the vertical emittance growth to about 20%. The convergence of the method is illustrated in Figure 7-53 where the vertical single-bunch emittance growth and the vertical rms beam offset in the BPMs are shown as a function of iteration number. It is seen that after about 5 iterations a very reasonable situation is achieved with almost no further improvements after about 15 iterations.

The emittance growth for the case of perfect BPMs is a function of the initial quadrupole misalignment which determines the magnitude of the dispersive deflections between sections. This is shown in Figure 7-54. The vertical single-bunch emittance growth and the rms trajectory offset at the BPMs are shown as a function of the initial rms

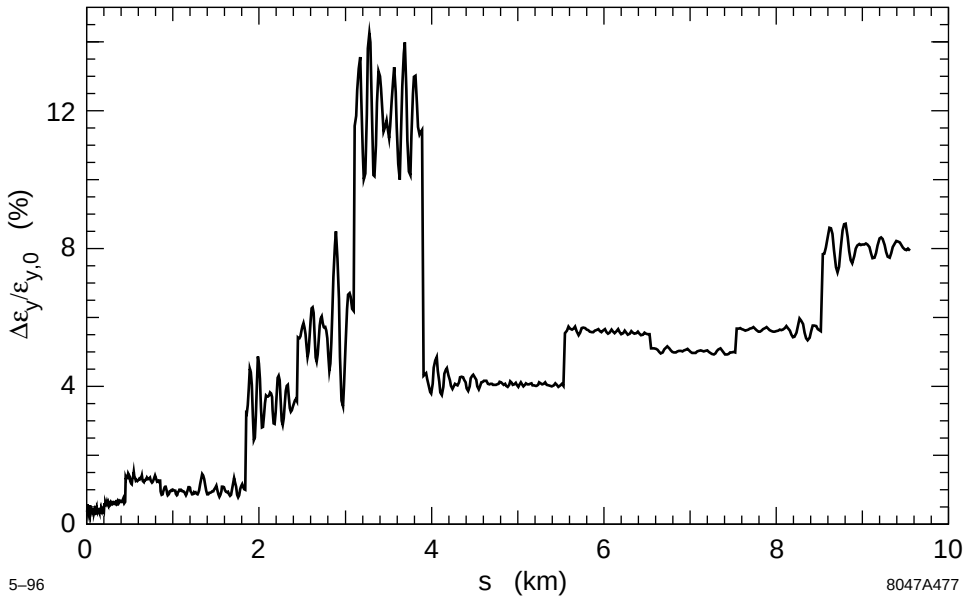


Figure 7-51. *Emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ along the linac for the example from Figure 7-50. Note that the steps occur at the end of the alignment sections where dipole kicks launch the beam into the next section.*

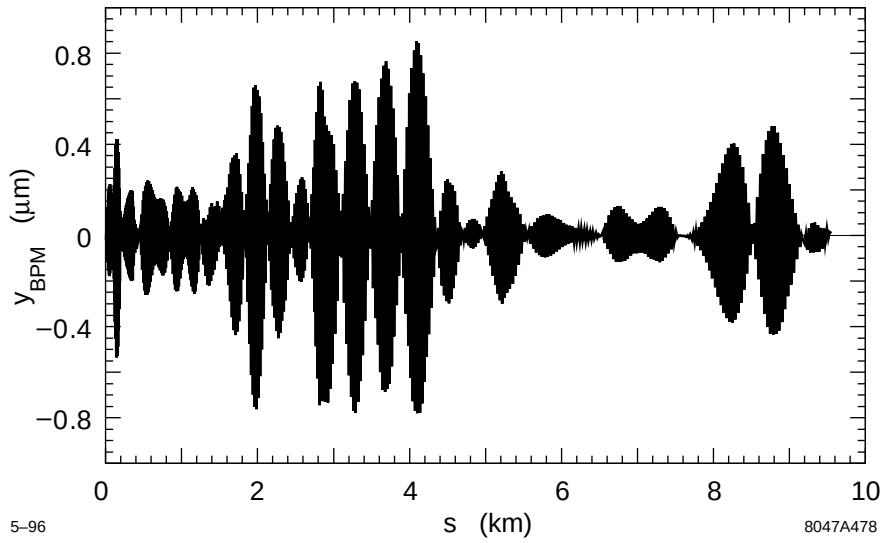


Figure 7-52. *Vertical trajectory offsets at the quadrupole BPMs along the linac for the example from Figure 7-50. The residual beam offsets in the quadrupoles bend the beam along the "bowed" beam line after alignment. They are not betatron oscillations and cause only negligible emittance growth (compare to Figure 7-51).*

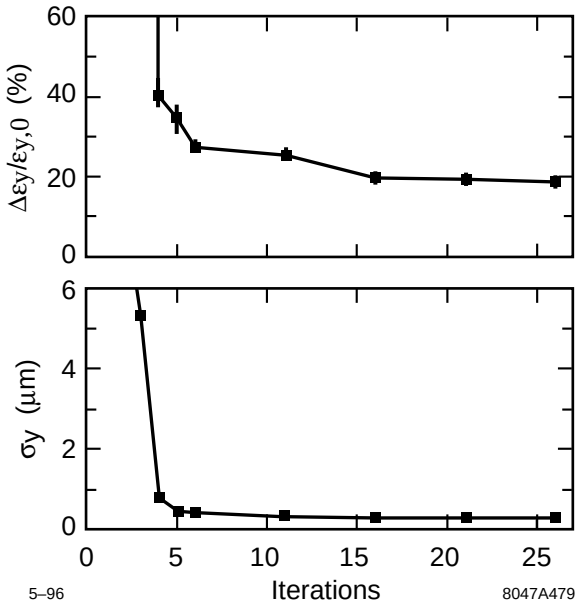


Figure 7-53. *Convergence of the alignment algorithm for an initial random quadrupole misalignment of 100-μm rms and perfect BPMs. The upper plot shows the average emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ as a function of the iteration number. The lower plot shows the corresponding rms trajectory offset with respect to the BPMs.*

quadrupole offset. As expected, the rms trajectory offset in Figure 7-54 increases linearly and the vertical emittance growth increases quadratically with the initial rms quadrupole misalignment.

Now we consider a more realistic case where the BPMs have a 2-μm-rms static offset with respect to the quadrupoles and a 1-μm-rms resolution. The initial quadrupole misalignment is 50-μm rms. We still assume perfect rf-BPMs. The alignment is done in 14 sections and 5 iterations. Figure 7-55 shows the average vertical single-bunch emittance growth for the different BNS cases and the imperfections defined above. The BNS configuration number 3 still yields the smallest emittance growth. However, due to the additional BPM imperfections, the emittance growth for the nominal BNS increased from about 7% to about 28%. This is well below the emittance growth budget of 175% for NLC-IIb.

The Effects of Transverse Long-range Wakefields

Thus far, we have only considered single-bunch effects. However, for the bunch train of 90 bunches, transverse long-range wakefields become important. In order to set the scale of those effects, we show the multibunch emittance blowup due to a 1-sigma initial vertical beam offset. Figure 7-56 shows the special case where each bunch is described as a single beam ellipse (zero bunch length). Single-bunch wakefields do not apply and the phase space distribution of the bunch train only reflects long-range wakefields. The phase space distribution of the bunches is compared to the beam ellipse that is obtained from the average single-bunch emittance. The bunches spread all over the beam ellipse but do not reach far beyond that. Multibunch emittance growth is therefore small and is found to be 25.6%.

Figure 7-57 shows the same case for a more realistic beam distribution. The bunch has a finite length and single bunch wakefields apply. Filamentation and BNS damping cause the phase space distribution of the bunches to be smaller. In this case the emittance of the first bunch grows by 38.7%, the average single-bunch emittance growth is 49.0% and the

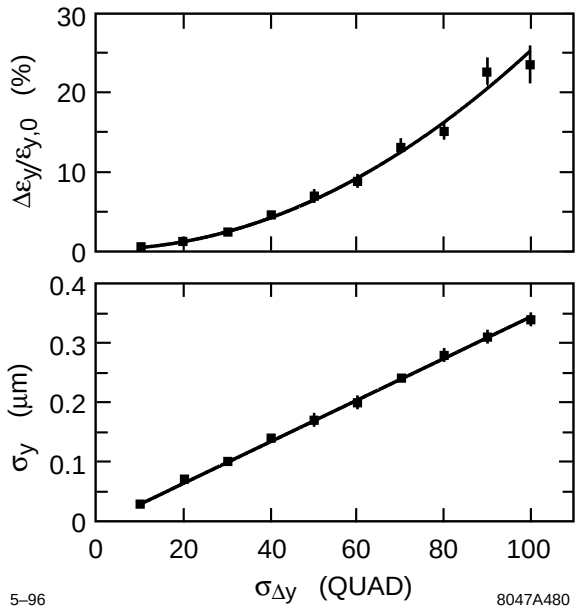


Figure 7-54. *The average emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ and rms trajectory offset σ_y after alignment versus the initial rms quadrupole misalignment $\sigma_{\Delta y}$.*

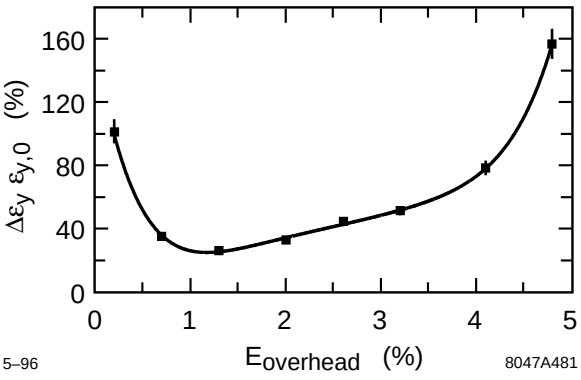


Figure 7-55. *Average emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ after alignment for different BNS configurations. Here we assume an initial quadrupole alignment of 50-μm rms, a static BPM to quadrupole offset of 2-μm rms, a BPM-resolution of 1-μm rms and perfect rf-BPMs. The alignment is done in 14 sections and 5 iterations.*

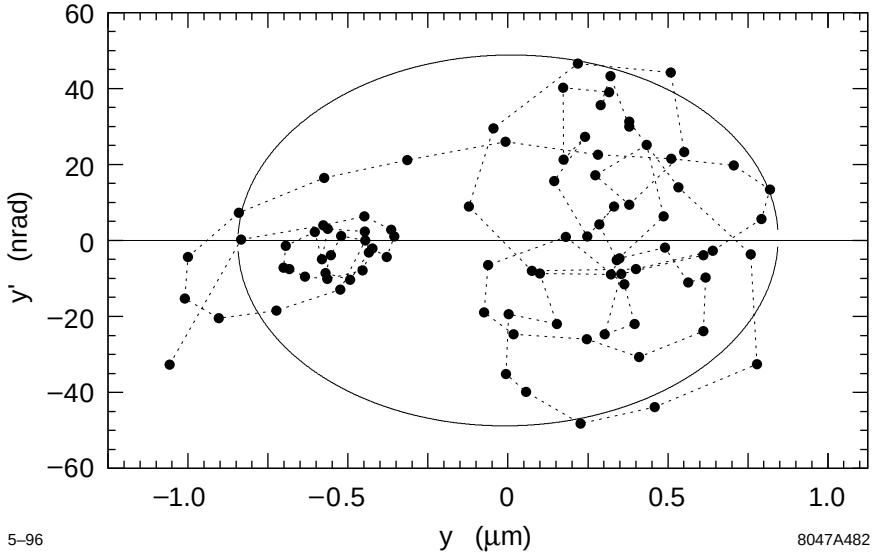


Figure 7-56. *Phase space locations y and y' of the bunches at the end of the linac for a 1-sigma initial vertical beam offset. y and y' are referenced to their average values. The bunch positions are compared to the beam ellipse from the average single-bunch emittance. Every bunch is tracked as a single beam ellipse so that no single-bunch wakefield effects apply. For this case a multibunch emittance growth of 25.6% was calculated.*

total emittance growth is 54.4%. The difference between the last two numbers reflects the impact of the long-range transverse wakefields. Here emittance growth is dominated by single-bunch wakefield effects.

Betatron Coupling

In the NLC, the horizontal emittance is about 100 times larger than the vertical emittance. Betatron coupling can couple part of the large horizontal emittance into the vertical plane. This effect is illustrated in Figure 7-58. Since the horizontal trajectory was kept flat for this study, no vertical betatron oscillations are induced from the coupling. If there were any, they would be absorbed into the alignment and trajectory correction. The emittance growth shows a quadratic dependence on the rms quadrupole roll angle. Assuming the tolerance value of the roll angle of about 300 μrad rms, the emittance growth is well below 10% and is not important.

7.6.6 Full Simulation

For a full simulation, we put the most important imperfections together, apply the correction algorithms and observe the emittance growth. In order to illustrate the importance of the several effects, we proceed in steps. For each case we quote the emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ at the end of the linac and the rms beam offset σ_y at the BPMs.

1. Initial random quadrupole misalignment of 100-μm rms. Rf structures are aligned to the beam.

$$\frac{\Delta\epsilon_y}{\epsilon_{y,0}} = (24.4 \pm 2.3)\% \tag{7.8}$$

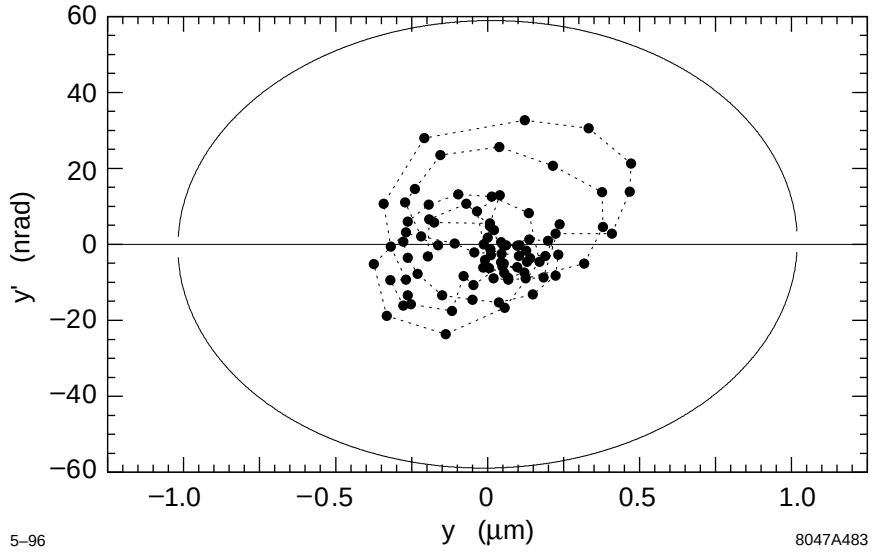


Figure 7-57. *Phase space locations y and y' of the bunches at the end of the linac for a 1-sigma initial vertical beam offset. y and y' are referenced to their average values. Single-bunch wakefield effects are included. The amplitude of the incoming betatron oscillation and the resulting long-range wakefield effects are reduced by filamentation. The bunch positions are compared to the beam ellipse from the average single-bunch emittance. Here we obtain a total emittance growth of 54.4%. The emittance growth of the first bunch is 38.7% and the average single-bunch emittance growth is 49.0%.*

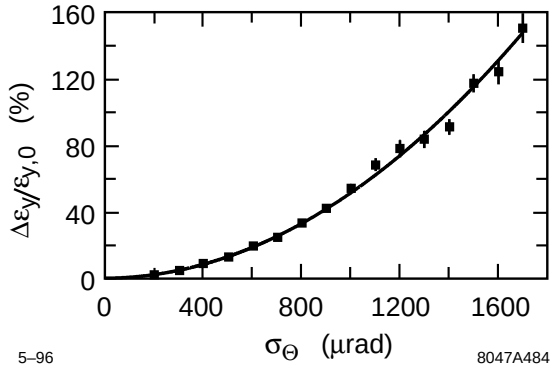


Figure 7-58. *Average vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ as a function of the rms quadrupole roll θ around the longitudinal direction. There are no horizontal beam offsets in the quadrupoles. The data only shows "emittance coupling".*

$$\sigma_y = (0.35 \pm 0.01)\ \mu\text{m} \tag{7.9}$$

2. Add: BPM resolution of 1-μm rms. Static BPM-to-quadrupole offsets of 2-μm rms.

$$\frac{\Delta\epsilon_y}{\epsilon_{y,0}} = (41.1 \pm 2.4)\% \tag{7.10}$$

$$\sigma_y = (1.08 \pm 0.01)\ \mu\text{m} \tag{7.11}$$

3. Add: rf-BPM resolution of 15-μm rms.

$$\frac{\Delta\epsilon_y}{\epsilon_{y,0}} = (90.2 \pm 6.0)\% \tag{7.12}$$

$$\sigma_y = (1.21 \pm 0.01)\ \mu\text{m} \tag{7.13}$$

4. Add: Rf-phase errors of 1° rms. Rf amplitude errors of 0.2% rms. Quadrupole roll errors of 300-μr rms. Quadrupole gradient errors of 0.3% rms.

$$\frac{\Delta\epsilon_y}{\epsilon_{y,0}} = (97.8 \pm 3.6)\% \tag{7.14}$$

$$\sigma_y = (1.22 \pm 0.01)\ \mu\text{m} \tag{7.15}$$

5. Add: Multibunch long-range wakefield effects.

$$\frac{\Delta\epsilon_y}{\epsilon_{y,0}} = (106.6 \pm 3.9)\% \tag{7.16}$$

$$\sigma_y = (1.23 \pm 0.01)\ \mu\text{m} \tag{7.17}$$

All emittance growth numbers, apart from the last one, refer to the single-bunch emittance growth. The total multi-bunch emittance growth of about 110% is well below the allowed emittance dilution of 175% for NLC-IIb (compare Table 7-2). Internal structure misalignments, special multibunch imperfections and the effects of missing BPMs will be added to the simulations in future studies.

The most important imperfections considered so far are BPM and rf-BPM errors. They determine the quality of the correction and the residual emittance growth. In all cases, the correction and alignment is done on the first bunch, assuming that all other bunches behave similarly. The small additional multibunch emittance growth shows that this is a valid assumption, although we have not yet included the effects of internal structure misalignments. The distribution of emittance growth for different error distributions is shown in Figure 7-59 for the full simulation (last case). The exponential tail for large emittance dilutions tends to bias the average emittance growth towards larger values. It results from error distributions that have a large component at the betatron frequency. Fortunately, these errors are easily corrected using bump (global) correction methods.

Figure 7-60 shows an alternate way to quantify the effect of imperfections. The so-called "luminosity reduction" (L_y) is calculated from a cross-correlation of the bunch distributions with themselves. Beam tails are naturally de-weighted in this process. The result is quoted as the fraction of the maximum luminosity that would be achieved for the particular case. The luminosity reduction comes out to be 71.1% for the full NLC simulation. That means that on average about 70% of the maximum luminosity would be obtained with the given errors.

Figure 7-61 shows the average emittance growth along the linac for 100 seeds of the full simulation. The growth shows a square-root dependence on the longitudinal position s. No particularly bad region can be identified and the emittance

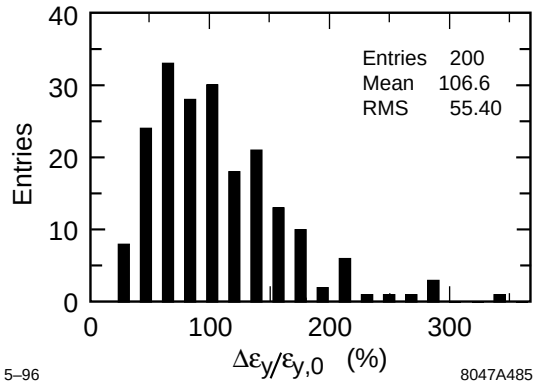


Figure 7-59. *Histogram of vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ for 200 different error distributions. The average emittance growth is 106.6% ± 3.9%. Note the exponential distribution for large emittance dilutions.*

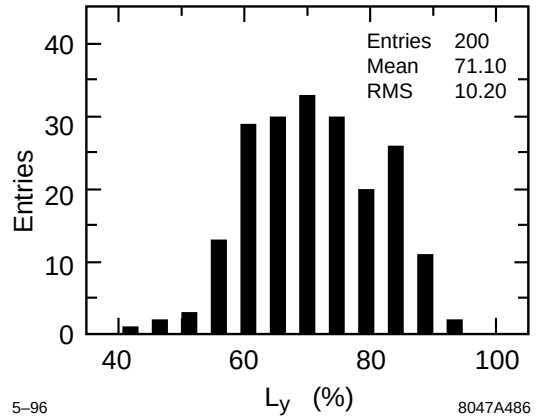


Figure 7-60. *Histogram of the luminosity reduction L_y for 200 different error distributions. The average luminosity factor is 71.1% ± 0.7%.*

growth behaves rather smoothly. Small residual step increases of the emittance can be seen. They are caused by the transitions between alignment sections. However, they are of little importance to the overall emittance growth.

Figures 7-62, 7-63 and 7-64 show an example of the long-range wakefield effects for the full set of imperfections. Figure 7-64 is of particular interest. The single-bunch emittance is shown as a function of the bunch number. It shows an oscillatory behavior and is even reduced below the value of the first bunch for some of the trailing bunches. Long-range wakefields induce bunch oscillations that can work as "emittance bumps" and reduce emittance. If one could measure the emittances and trajectories of all bunches one could pick the bunch with the smallest emittance and use multibunch kickers to correct the other bunches to this reference. The multibunch scheme can in principal be used to conveniently explore the phase space to find the smallest possible emittance dilution.

7.6.7 Stability

Stability questions determine how often the alignment and correction algorithms need to be applied. In the previous section we described a procedure that serves as both the alignment and trajectory correction algorithm. We have shown

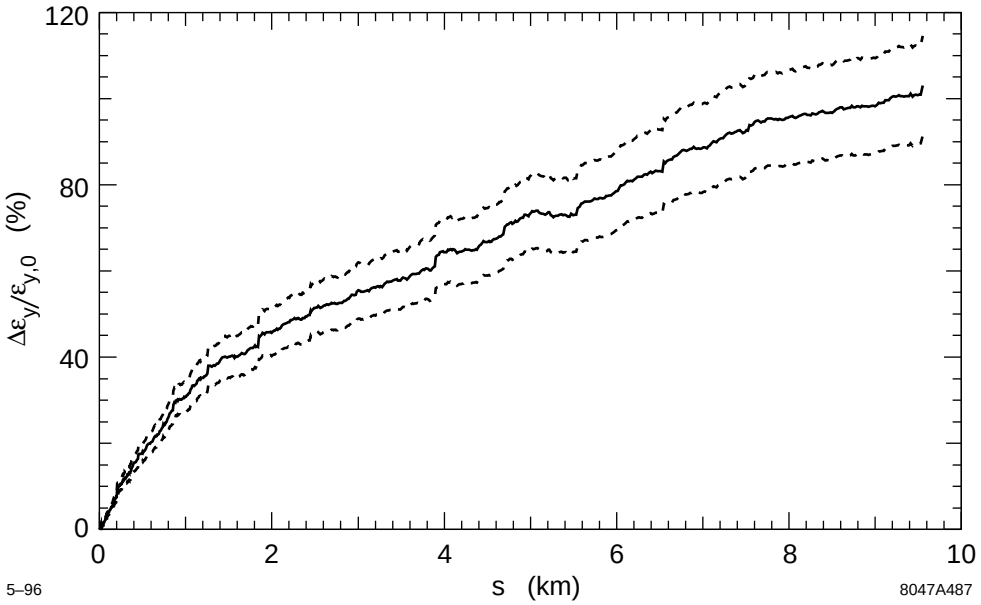


Figure 7-61. *Average emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ along the linac for the full simulation. The dashed curves specify the error bars around the average (solid curve).*

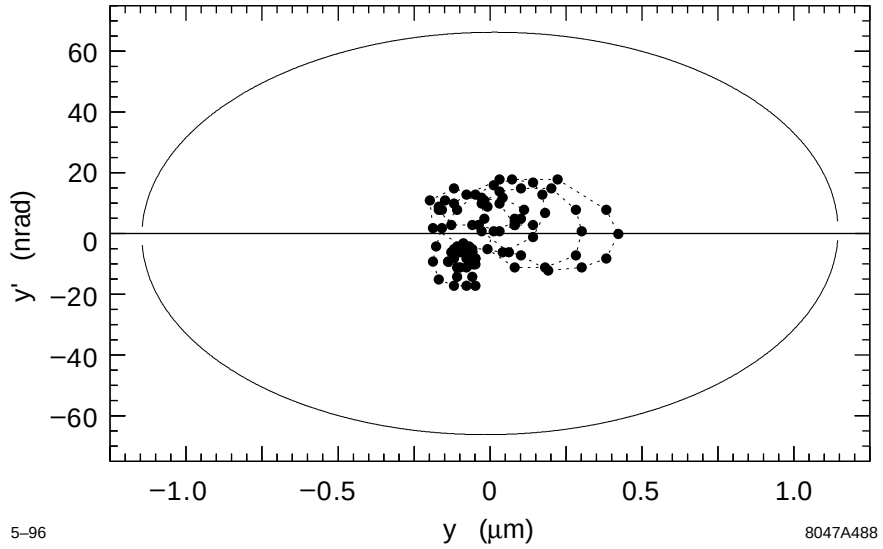


Figure 7-62. *Example of the vertical phase space distribution of the bunch train at the end of the linac. The bunch locations are compared to the beam ellipse as obtained from the average single-bunch emittance. The total vertical emittance growth is 89.0% for this case. This is to be compared to an emittance growth of the first bunch of 83.0%, and an average single-bunch emittance growth of 85.2%.*

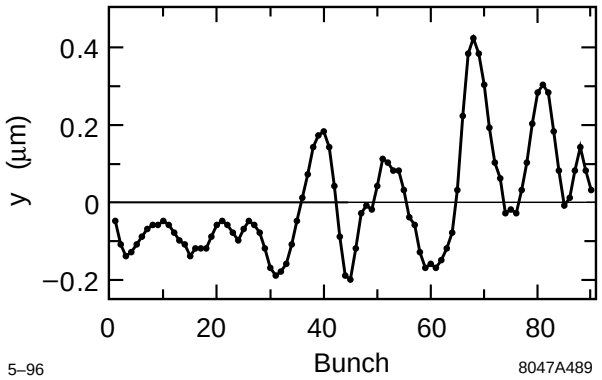


Figure 7-63. *Vertical bunch positions y as a function of bunch number at the end of the linac (same example as in Figure 7-62).*

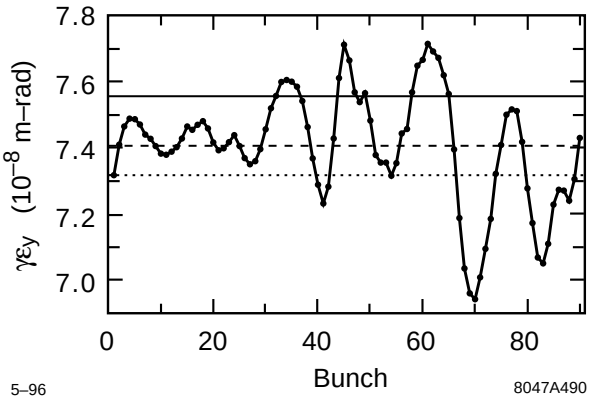


Figure 7-64. *Vertical normalized single-bunch emittance ϵ_y as a function of bunch number at the end of the linac (same example as in Figures 7-62 and 7-63). The three lines indicate the emittance growth of the first bunch (dotted), the average single-bunch emittance growth (dashed) and the total beam emittance growth (solid).*

that static imperfections can be corrected down to the required levels. Now we consider the question of how stable the optimized linac is and how fast it deteriorates. From the beam dynamics point of view, linac stability problems are dominated by quadrupole drifts; the quadrupoles generally have the tightest alignment tolerances.

Here we do not discuss BPM stability questions in detail. However, the requirements are tight. For the alignment algorithm we require a 2-μm static rms offset between the BPM and quadrupole centers. This tolerance can be achieved with a time-consuming beam-based alignment procedure and it must be stable over significant periods of time (days). The question of BPM stability is discussed elsewhere. Since quadrupoles and BPMs are mechanically mounted together, the BPM stability that can be achieved is mainly determined by the BPM electronics, cable lengths and similar factors.

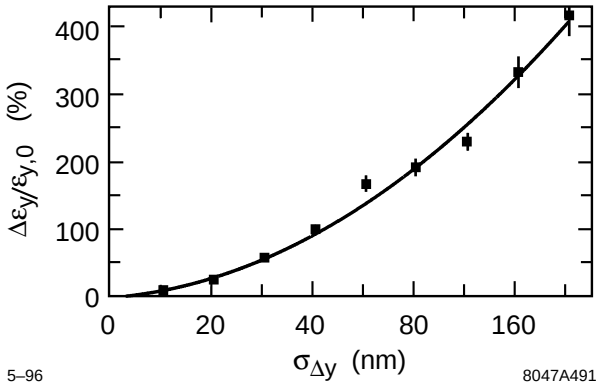


Figure 7-65. *Average single-bunch emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ as a function of the rms quadrupole offset Δy.*

Quadrupole Vibrations

Vibrations are fast random motions of the quadrupole alignment around its average position. Here we assume that the vibrations are white noise and that they are not damped by trajectory feedbacks. In reality the quadrupole vibrations due to ground motion, cooling water turbulence, etc will depend on frequency and will be partly suppressed by the trajectory feedbacks.

The average vertical emittance growth is shown in Figure 7-65 as a function of the vertical rms quadrupole offset $\sigma_{\Delta y}$. Here we only consider single-bunch emittance growth with respect to the beam centroid. Offsets that make the beams miss at the IP are assumed to be suppressed by fast feedbacks. Those effects are discussed in more detail in Section 7.4 and in the final focus chapter. From Figure 7-65 we see that the quadrupole white noise jitter should be kept below 20 nm in order to avoid single-bunch emittance growth in excess of 10%; this tolerance is well above the measured quadrupole jitter. Note that a tighter tolerance results if one considers the emittance growth with respect to the beam axis.

The emittance growth for quadrupole jitter is mainly caused by the uncorrected betatron oscillation that builds up along the linac. This will be discussed in the next section.

ATL-like Alignment Drifts

Next, we consider the question of how often the alignment and correction algorithm needs to be applied. Assuming a perfect starting point (flat trajectory, no emittance growth) we let the alignment drift and observe the deterioration of the trajectory and the emittance. In order to model alignment drifts we use the ATL-model. The ATL-model predicts that the rms vertical misalignment $\sigma_{\Delta y}$ (in μm) deteriorates with time T (in seconds) and over the length L (in m) as follows:

$$\sigma_{\Delta y}^2 = A \cdot T \cdot L \tag{7.18}$$

We use an A coefficient of 5×10^{-7} μm^2/s/m. This value was inferred from data taken on the SLAC site for times T over 30 hours. We should caution that the validity of the ATL-model (diffusive drifts) is not well established. Recent observations indicate a linear increase of the rms misalignment with time for longer periods (systematic drifts). Assuming diffusive alignment drifts from the ATL-model for time periods of the order of hours might well be overly pessimistic.

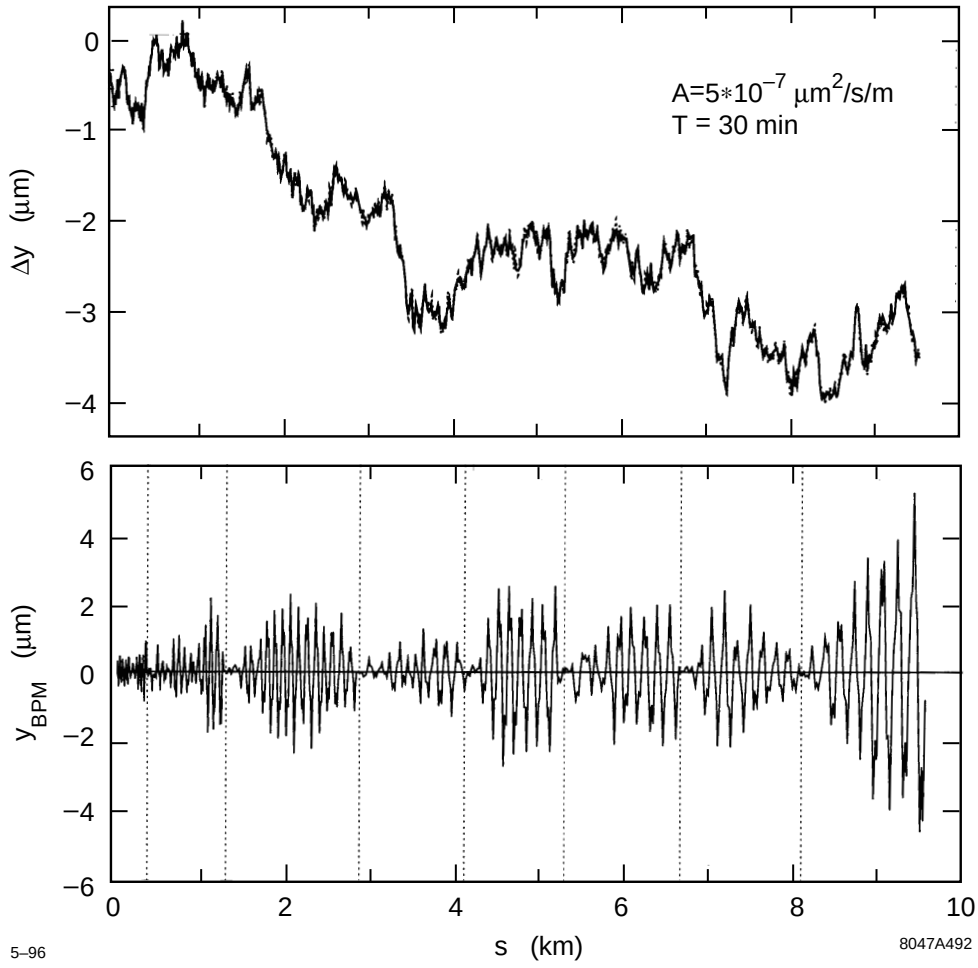


Figure 7-66. *Example of ATL-like alignment drifts. The upper plot shows the displacements of quadrupoles, rf-structures and BPMs after 30 minutes with an A-coefficient of* 5×10^{-7} μm^2*/s/m. The alignment was flat initially. The lower plot shows the corresponding trajectory offsets* y_{BPM} *at the BPMs. The dotted lines indicate the locations of trajectory feedbacks where* y *and* y' *are corrected back to zero. Thus the size of coherent betatron oscillations is constrained.*

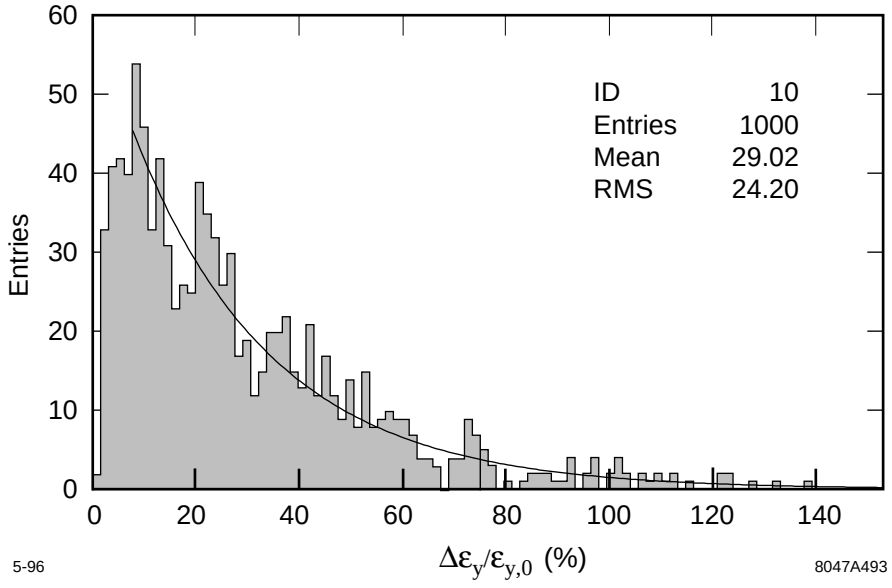


Figure 7-67. *Histogram of vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ for 1000 different error distributions. The average emittance growth is 29.0% ± 0.8%. The solid curve shows an exponential fit for large emittance dilutions.*

Figure 7-66 illustrates ATL-like alignment drifts. It shows the displacements and corresponding trajectory offsets at the BPMs after 30 minutes of deterioration. The offsets of quadrupoles, BPMs and rf-structures are overlaid in the plot and are essentially indistinguishable. The trajectory offsets at the BPMs show the coherent betatron oscillations that build up. The dotted lines indicate the locations of seven trajectory feedbacks that constrain y and y' to zero. The coherent betatron oscillations are thus broken up into eight smaller oscillations. The oscillation amplitude is a few μm and is large enough to be detected with a BPM resolution of 1 μm.

A histogram of the vertical single-bunch emittance growth from alignment drifts after 30 minutes and an A of 5×10^{-7} μm^2/s/m is shown in Figure 7-67. The average emittance growth is found to be 29.0% ± 0.8%. This size of emittance growth will prompt a beam-based alignment of the quadrupoles, that serves as an effective trajectory correction at the same time. Note, however, the exponential distribution for large emittance dilutions. The most probable emittance growth is only about 10%. Figure 7-68 shows a histogram for the luminosity reduction in the same case. The distribution shows an approximately Gaussian distribution for smaller luminosity, with an average luminosity of about 90% of its ideal value.

The average emittance growth along the linac is shown in Figure 7-69. The locations of the trajectory feedbacks are clearly seen. As a coherent betatron oscillation builds up the emittance starts to grow exponentially. The feedbacks stop this exponential growth. More effective feedbacks can be imagined if the average y and y' is minimized up to the next feedback instead of correcting y and y' locally.

Figure 7-70 shows the average vertical single-bunch emittance growth for different BNS configurations. We consider the same case as above. All previous results were obtained using the standard BNS configuration with an energy overhead of 1.3%. Figure 7-70 shows that BNS configurations with higher energy overheads reduce the emittance growth from 29% to about 16%. This result is in better agreement with the simple two-particle model autophasing condition than the other results that were discussed before. One can therefore imagine to trade alignment performance against better stability. However, this is not necessary.

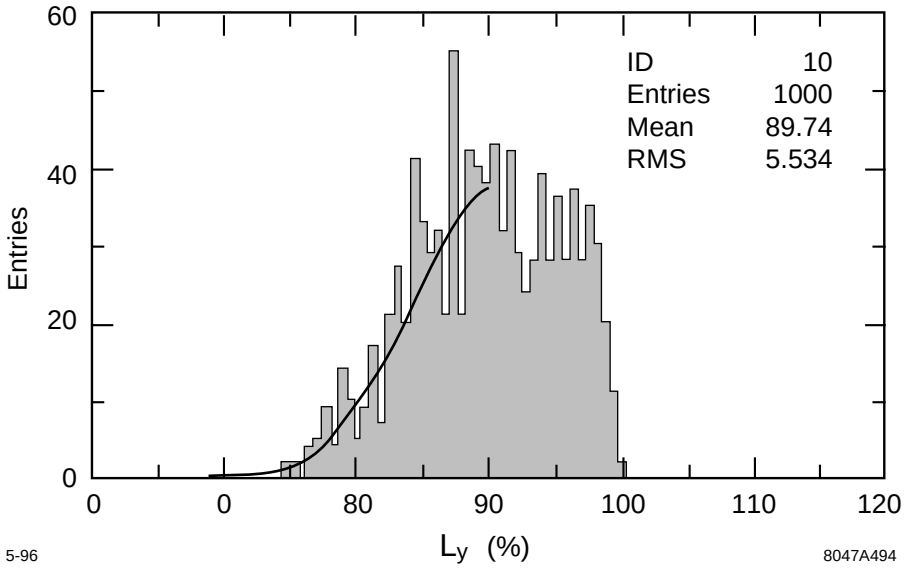


Figure 7-68. *Histogram of the luminosity factor L_y for 1000 different error distributions. The average luminosity is 89.7% $\pm$ 0.2%. The solid curve shows a Gaussian fit to the lower half of the distribution.*

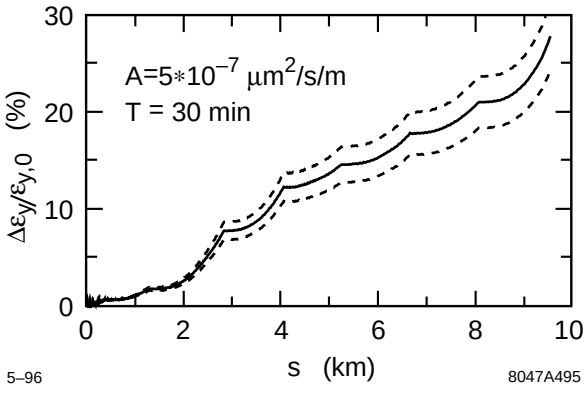


Figure 7-69. *Average vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ along the linac for ATL-like drifts after 30 minutes. We assume an A-coefficient of 5×10^{-7} μm^2/s/m. The dashed curves specify the error bars around the average (solid curve).*

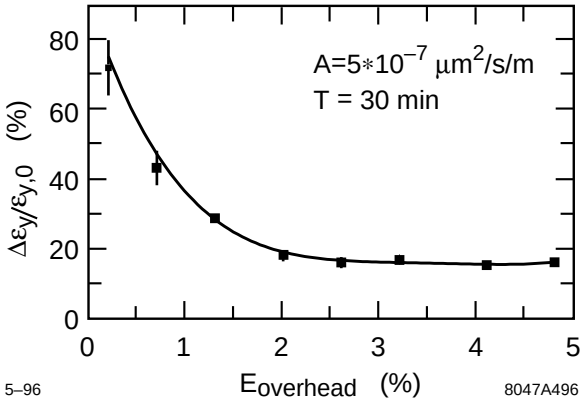


Figure 7-70. *Average vertical emittance growth $\Delta\epsilon_y/\epsilon_{y,0}$ from ATL-like alignment drifts for different BNS configurations.*

Let us relate the ATL-like alignment drifts to the other results. Since the emittance growth is linear in time we would get an additional average emittance growth of about 15% when we assume a beam-based alignment every 30 minutes. This is about a factor of six smaller than the emittance growth expected after beam-based alignment of quadrupoles and rf-structures. It is small enough not to be an important limitation of the NLC linac performance as long as the linacs are corrected regularly every 30 minutes. Since the alignment and correction algorithm does not interfere with the standard operation, its frequent application should be no major obstacle.

7.6.8 Summary and Outlook

A beam-based alignment algorithm for quadrupoles and rf-structures was simulated with a realistic BNS configuration. It was shown that the large initial misalignments from conventional alignment procedures can be corrected to acceptable levels. The emittance growth that finally can be achieved depends on the initial misalignment and most importantly on the performance of the BPMs and rf-BPMs. Assuming an initial misalignment of 100-μm rms, static BPM to quadrupole offsets of 2-μm rms, a BPM resolution of 1-μm rms and an rf-BPM resolution of 15-μm rms, we find an average emittance growth of 90.2% $\pm$ 6.0% for a single bunch. Adding errors of the rf-phases, rf-amplitudes, quadrupole roll angles and quadrupole gradient errors increases the emittance growth to 97.8% $\pm$ 3.6%. Adding the effects of transverse long-range wakefields on the bunch train yields the final multibunch emittance growth of 106.6% $\pm$ 3.9%. This emittance growth is smaller than the allowed total emittance growth of 175%. As the emittances roughly add in quadrature the impact of additional imperfections gets smaller with larger emittances. It is anticipated that the alignment algorithm can be further optimized by smoothing the transitions between alignment sections. Further details can be found in the text.

The dominant stability problem is caused by drifts of the quadrupole alignment. We simulated this effect by using the ATL-model with a coefficient $A = 5 \times 10^{-7}$ μm^2/s/m as measured at SLAC. The use of this model for times of about 1 hour might be overly pessimistic. We showed that the alignment drifts drive coherent betatron oscillations that lead to exponential emittance growth. The addition of seven trajectory feedbacks breaks the coherent betatron oscillation down into eight smaller oscillations. Assuming a beam-based quadrupole alignment every 30 minutes, we get an additional average emittance growth contribution of 15%. Since the emittances roughly add up in quadrature, this is small compared to the 107% average emittance growth after alignment. The alignment algorithm does not interfere with the normal linac operations, so that it can be applied very frequently. We conclude that we can handle alignment drifts safely.

Source		$\Delta\epsilon_y/\epsilon_y$	Dynamics
Quad alignment:		40%	Incoherent dispersion
BPM resolution	1 μm		
BPM-to-quad alignment	2 μm		
Quad drift between alignment:		10%	Coherent dispersion and wakes
Steering period	30 min		
Structure alignment:			
Internal alignment	15 μm	25%	Long-range wakes
Beam measurement accuracy	15 μm	50%	Short-range wakes
Other (*e.g.,* quad roll, ion effects, and rf deflections):		50%	
Total:		**175%**	

Table 7-9. *NLC-IIb linac emittance growth budget.*

Future simulation studies will include internal structure misalignments, multibunch imperfections (bunch-to-bunch charge, energy, etc. variations) and the effects of missing BPMs. In addition, we further want to apply emittance bumps in order to compensate the emittance growth below what has already been achieved. Finally, we need to study the impact of different bunch shapes on the linac emittance transport and we need to simulate the collimation requirements at the end of the linacs.

7.7 Tolerances

In this section we give a brief summary of the tolerances needed to control emittance growth in the main linacs. Simulation studies leading to these tolerance numbers have been discussed in previous sections, especially Section 7.6.

Unless otherwise noted, we assume the NLC-IIb parameters, *i.e.,* a charge per bunch of 1.10×10^{10} with a 150-μm bunch length and an injected normalized vertical emittance $\gamma\epsilon_y = 4 \times 10^{-8}$ m-rad. The tolerances are specified to deliver a beam with a vertical emittance of $\gamma\epsilon_y = 11 \times 10^{-8}$ m-rad. As noted earlier, the primary source of emittance dilution arises from the accelerator structure misalignments with respect to the beam. In the NLC design, the other sources are relatively insignificant, primarily because established beam-based alignment techniques can be used to attain accurate alignment of the BPMs and quadrupoles.

The total emittance budget is given in Table 7-9 for the NLC-IIb linac design. At the present time, we have budgeted roughly half of the allowed 175% total emittance growth to the misaligned accelerator structures. As stated, this partitioning arose because we have more confidence in the quadrupole alignment techniques. Regardless, in the future, we will probably redistribute the tolerances to allow a greater margin for the other effects; this work is awaiting more simulations that incorporate the interactions of these many disparate sources and account for stability errors.

Finally, we should note two points: First, we are calculating the *expected* average emittance growth. In any given case, the growth can be larger or smaller. The larger growths occur when there is a significant correlation of the errors at the betatron frequency, thereby driving the dilution resonantly. Fortunately, these cases are usually very straightforward to correct with simple global correction schemes and thus we are not very concerned about getting a bad "seed" in the collider.

Second, when considering large emittance dilutions, the rms emittance growth is not a very accurate measure of the luminosity degradation. This occurs because the emittance growth mechanisms, such as wakefields, are nonlinear. Thus, they populate the beam tails without degrading the core emittance as greatly as suggested by the rms emittance growth. A more representative quantity is the luminosity reduction which describes the decrease in the overlap of the beams at the IP; we will adopt this measure in the future.

dispersion ($\delta_{\rm eff}$=.2%)

7.8 Operations and Machine Protection System

The operation of the linacs is constrained by the requirements of the Machine Protection System (MPS); the MPS is described in greater detail in Chapter 16. Because the beams have extremely high charge densities, a single bunch train, or even a single bunch, incident on an accelerator structure could cause damage. The philosophy of the MPS is that all preliminary beam-based alignment, tuning, and diagnostics be performed with a full current high emittance single-bunch beam which cannot cause damage in the linac; these operations can be performed at low repetition rates.

As described, the accelerator structures are protected from this diagnostic single-bunch beam with spoilers which are located after every accelerator structure. The spoilers have an aperture that is 2 mm in radius, this is roughly half that of the accelerator structure irises, and they are 0.2 radiation lengths in length. If the beam is steered sufficiently far off-axis to intercept an accelerator structure, it must first pass through one of more spoilers. These will increase the beam angular divergence so that, by the time the single bunch beam strikes a structure, it will not cause any damage. To prevent damage to the spoilers themselves, the emittance of the diagnostic beam must be increased by a factor of ten.

Once the initial beam-based alignment, the basic beam trajectory, and the energy profile are established, a high repetition rate of 120–180 Hz can be established. At this point, the beam emittance can be decreased and additional bunches can be added to the bunch trains. All subsequent tuning must be performed at the high repetition rate.

To verify the beam loading compensation when going to the longer bunch trains, the train length will be increased in steps. Presently, we believe that the loading could be checked using only one intermediate train length, namely, going from one bunch per train to 10 bunches per train and then to a full 90 bunches per train; this needs further verification and increasing the steps has minimal implications for the rest of the systems. In addition, to prevent thermal changes of the accelerator structures when changing from low repetition rate to high rate and from the single- to the multibunch modes, the klystron phases are varied so that the addition power which normally would accelerate the missing bunches is dumped into loads rather than into the accelerator structures; this is discussed further in Chapter 8.

To prevent the high rate beams from striking accelerator components, a $\pm 200\,\mu$mtrajectory window and 10% energy window are established about the nominal values. If the beam deviates beyond these limits in any single pulse, the collider is returned to the high-emittance single-bunch diagnostic mode while the source of the problem is diagnosed from data that was taken during the errant pulse. Extensive logic will be used to prevent erroneous MPS faults due to bad BPMs readings.

This system relies on the fact that there are no transverse deflecting fields that can change sufficiently, within a single interpulse period (roughly 8 ms), to deflect the beams from the operational trajectory window into the accelerator structures. This is attained by limiting the strength of all fast correctors and limiting the decay time of the quadrupole and bending magnet fields by using solid core magnets and thick conducting vacuum chambers.

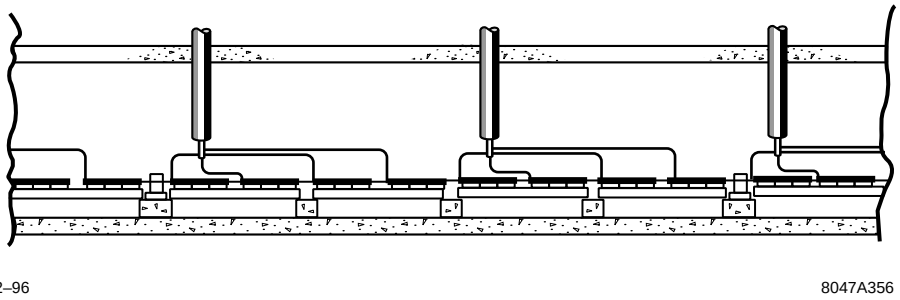


Figure 7-71. *Tunnel cross section and accelerator supports.*

In addition, the MPS must verify that the rf systems are operational before the beams are launched into the linacs. However, fairly large energy deviations can be tolerated. For example, a 20% energy deviation in combination with 100-μm random quadrupole misalignments, which are well in excess of what we expect, would only cause 1-mm orbit offsets.

Thus, the MPS system must only verify that 85% of the rf systems are operational. To this end, all modulators will be polled roughly 100-μs before beam time. At the same time, the klystron timing and phase information will be checked. If there are a sufficient number of failures, the beams will be aborted downstream of the damping rings; this verification procedure is described in greater detail in Chapter 8. Finally, there will be sacrificial spoilers and dumps located in the diagnostic station chicanes to prevent energy errors larger than 25% from propagating further down the linac.

7.9 Linac Components

In this section we give more detailed information on various components of the linacs. In the case of supports and movers, we present our most recent design ideas, and discuss the issues that impact the design.

7.9.1 Supports and Movers

Support Layout

The beam line components will be supported 0.5-m above the tunnel floor at roughly four-m intervals (Figure 7-71). This low height minimizes amplification of ground vibrations but still leaves access below the accelerator. Between supports, accelerator modules will consist of two 1.8-m-long sections of disc loaded wave guide which will be assembled and aligned on a common support girder. Focusing quadrupole magnets will be mounted to the supports at intervals along the accelerator. The spacing between quadrupoles will increase in five discrete steps, scaling roughly as the square root of the beam energy along the accelerator. Figure 7-71 shows the support layout at approximately the 200-GeV point.

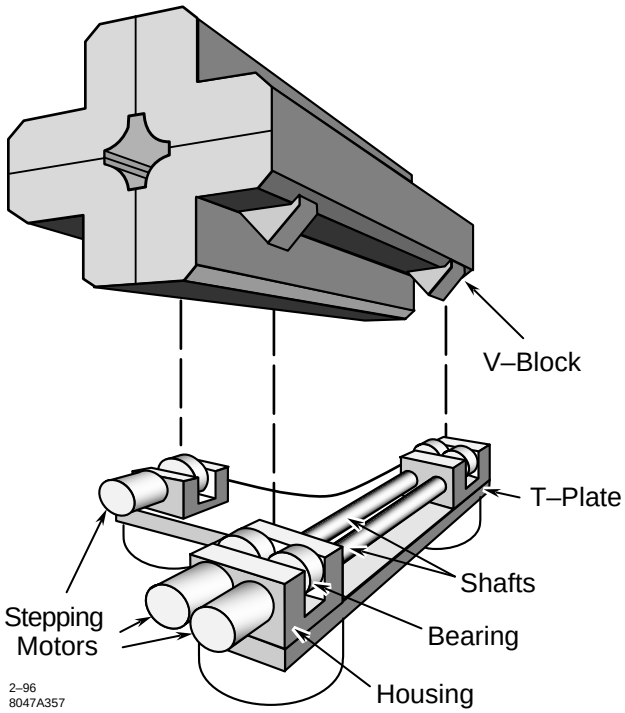


Figure 7-72. *Quadrupole roller cam support.*

Support Design

Both the quadrupole and accelerator structure positions will be remotely controllable. All accelerator and magnet supports will have stepper motor positioning mechanisms with a $\pm$1.5-mm range and a 0.3-μm resolution. Both the accelerator structures and quadrupoles will be kinematically supported in cradles formed by motor-driven roller cams as illustrated for a quadrupole in Figure 7-72.

These mover systems will be similar to the Final Focus Test Beam (FFTB) magnet movers which provide a $\pm$1.5-mm range with 0.3-μm position resolution. The type of kinematic support which will be used, where the number of contact support points equals the number of spatial degrees of freedom, is similar to the kinematic support designs for laboratory optics. It insures reproducible zero-play fixturing of components and will be free of all bending moments which might distort magnets or accelerator modules when they are remotely repositioned. The accelerator modules will rest on 5 cams: cams (1,2,3) on one support will fix roll and the x, y coordinates of that end. Cams (4,5) on the next support will fix the x, y position of the module's other end. Thus each module will be adjustable in x, y, roll, pitch and yaw as illustrated in Figure 7-73.

During operation only the inner eccentric shaft of a support cam will rotate under motor control. The outer bearing race remains fixed in contact with the magnet or module contact foot as shaft rotation lifts the module. Motion is strictly bounded by the design geometry. Limit switches and electronic override protection are not needed. Control system failures can only cause a cam to continue rotating. This is an important advantage for any system with as many as 30 thousand channels of motion. All support cams are arranged so that gravity applies a load torque to each cam shaft drive removing all backlash. All parts move by pure rolling motion and are free of hysteresis from intermittent and reversing sliding friction. These features make sub-micron positioning of heavy objects practical using inexpensive mass produced ball or roller bearings. Details of this mechanism are fully covered in [Bowden, 1996].

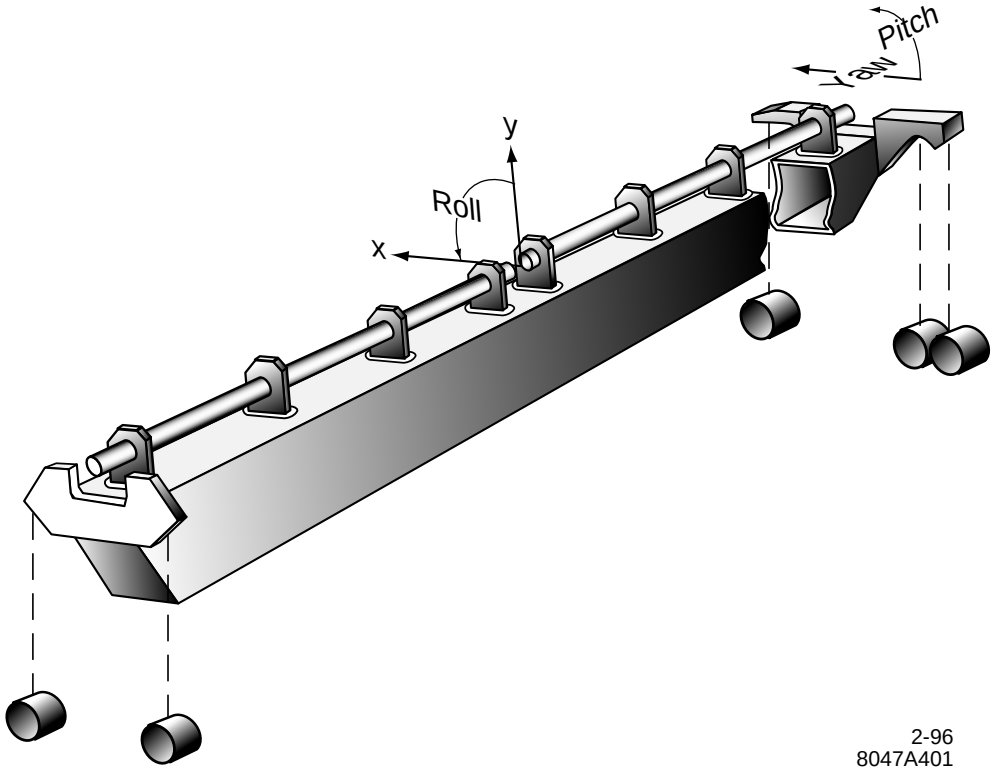


Figure 7-73. *Accelerator girder module support geometry.*

Initial Alignment

All supports will incorporate standard screw and shim manual adjustments which will be used during initial installation to align each support. After the support is manually aligned, accelerator modules and quadrupole magnets will be lowered into their cam cradles as illustrated in Figure 7-74. No further manual alignment is possible or needed. Each module or magnet has had its cam contact feet aligned to their nominal position during assembly. This guarantees the relative alignment between module and quadrupole on installation, reducing, by a factor of two, the number of objects to be manually aligned during installation in the tunnel.

Tunnel Heat Transfer

During operation, thermal distortion can cause serious misalignment of the accelerator. Any successful support system design must avoid or accommodate temperature gradients which could alter alignment. Thermal distortion is the result of thermal gradients which develop when heat is lost by the accelerator to its immediate surroundings. The accelerator cooling system temperature must be set at a temperature sufficiently above the outside air temperature that heat can be expelled even on the warmest days. For example, the existing SLAC accelerator runs with 45°C cooling water because summertime temperatures sometimes reach 40°C in that part of California. This cooling system heats the buried accelerator tunnel to about 40° which is 20° warmer than the surrounding earth. The temperature field for heat flow up from the warm tunnel to the cooler surface is illustrated in Figure 7-75.

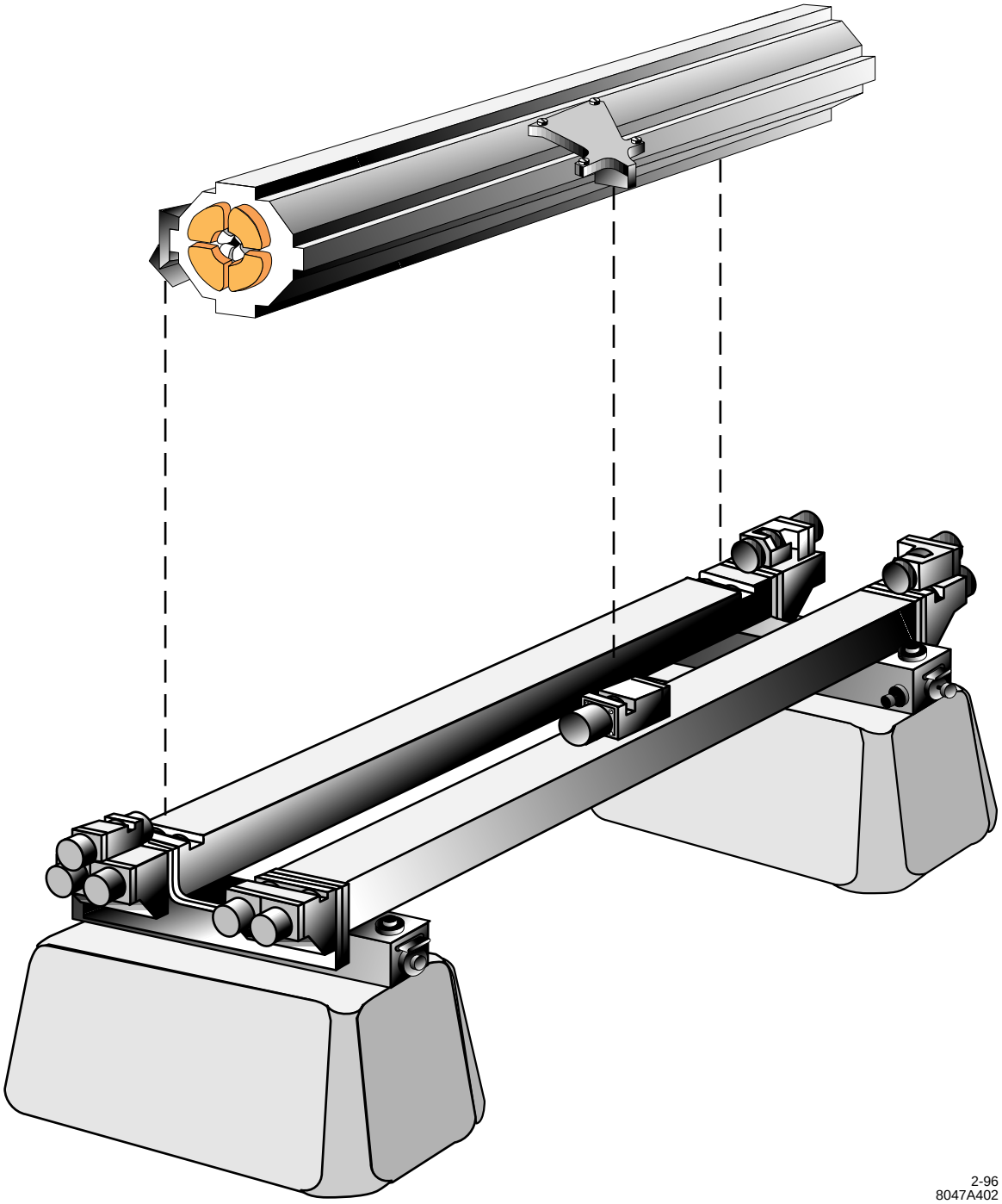


Figure 7-74. *Floor-mounted accelerator remote positioning support assembly.*

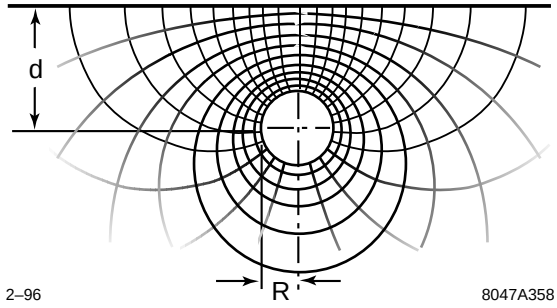


Figure 7-75. *Thermal field of a buried heat source*

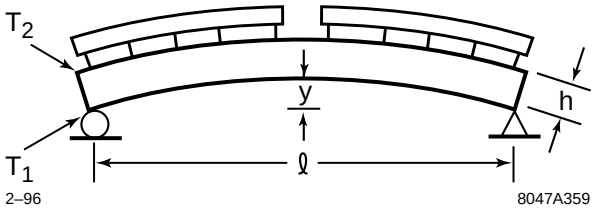


Figure 7-76. *Thermal distortion of a girder*

The thermal conductance between a circular tunnel and the surface is;

$$\kappa = 2\pi\lambda \left(\cosh^{-1}\left(\frac{d}{R}\right)\right)^{-1} \quad (\mathrm{W/\,m^{\circ}C}) \tag{7.19}$$

Heat loss depends on thermal conductivity of the earth λ and the ratio of tunnel depth to radius d/R. For the SLAC accelerator ($\lambda_{\mathrm{sandstone}} = 1.85\,\mathrm{W/\,m^{\circ}C}, d = 10\,\mathrm{m}, R = 1.5\,\mathrm{m}$), approximately 100 W is lost by the tunnel every meter of its length. Although this heat is small compared to the several kW/m carried away from the accelerator structure by cooling water, it is the source of thermal distortions which disturb accelerator alignment.

Girder Distortion

Accelerator modules consist of two 1.8-m-long accelerator structures mounted to a support girder. If the accelerator is mounted on top of its support girder, heat flow will create a thermal gradient across the girder section. The top surface of the girder will be warmer than the bottom which is further away from the source and cooled by natural convection. Differential expansion will cause the girder to arch up toward the warm accelerator, misaligning it. If the girder temperature gradient is linear across its depth, then the distortion is a circular arc of sagitta:

$$y = \frac{\alpha l^2}{8}\left(\frac{T}{h}\right) \qquad \alpha = \text{expansion coefficient } \frac{\Delta l}{l}/^{\circ}\mathrm{C} \tag{7.20}$$

This is illustrated in Figure 7-76. To limit the sagitta to 100 μm, for example, for a 4-m-long, 150-mm-tall, aluminum girder, the thermal gradient can not exceed $\Delta T = 8(100 \times 10^{-6}\,\mathrm{m})(.150\,\mathrm{m})/[(4\,\mathrm{m})^2(23 \times 10^{-6\circ}\mathrm{C}^{-1})] = 0.33^{\circ}\mathrm{C}$. For comparison, measurements across a similar support girder on the SLAC accelerator during operation show a ΔT of .25°C.

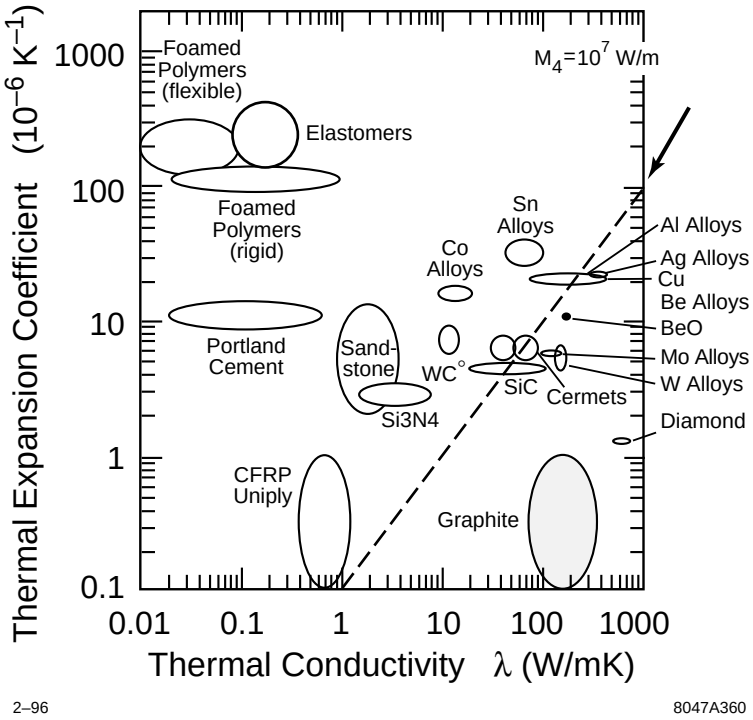


Figure 7-77. *Thermal properties of materials.*

Material Choice

Aluminum has a moderately high expansion coefficient which might appear to make it a poor choice for girder construction. Since thermal gradients are inversely proportional to thermal conductivity, thermal distortion depends on the ratio of conductivity to expansion coefficient λ/α. Judged by this criterion, aluminum is a good choice because of its high thermal conductivity. In a plot of conductivity versus expansion coefficient (Figure 7-77) aluminum proves superior to such low expansion coefficient materials as carbon fiber reinforced plastic (CFRP). Furthermore aluminum has an expansion coefficient $(23 \times 10^{-6\circ}\mathrm{C}^{-1})$ similar to copper $(17 \times 10^{-6\circ}\mathrm{C}^{-1})$ which is used for the accelerator structure and wave guides.

Low Distortion Girders

One method of removing the thermal gradient which distorts girder alignment is to heat the cooler side of the girder, perhaps with a cooling water tube. It may also be possible to enclose the girder in a common insulation jacket together with the temperature stabilized accelerator. From the standpoint of thermal distortion, the girder, ideally, should be an insulated shell enclosing the accelerator.

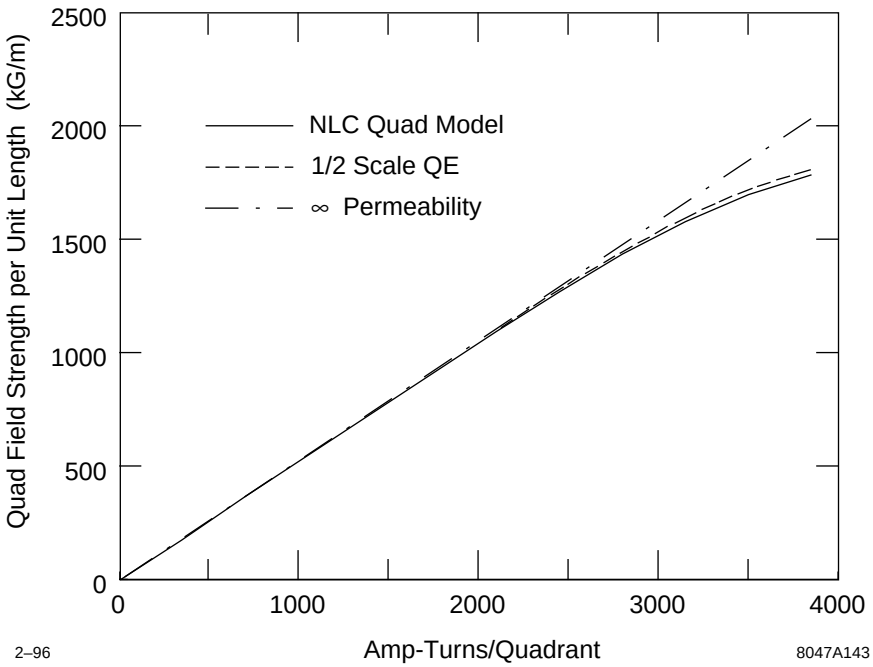


Figure 7-78. *Excitation curve for the NLC model quadrupole magnet and the half scale SLC linac QE magnet. Also shown is the ∞ permeability result, which is essentially the same for both magnets.*

7.9.2 Quadrupole Magnets

Based on the optics requirements, a design study for these quadrupoles was done by scaling the "QE" quadrupole design that is used in the SLC linac. As a starting point, the transverse quadrupole size was halved so the resulting gap width of 13.8 mm will comfortably accommodate a stripline BPM and still not be a limiting aperture compared to the structure irises. To allow the use of standard-sized conductor (0.1875″ square with a 0.100″ diameter cooling hole), and to reduce the magnet power consumption, the coil pocket size was increased while taking care not to significantly reduce the field strength to amp-turn ratio near saturation. A reasonable solution was found even after roughly doubling the coil pocket size. Figure 7-78 shows the resulting excitation curve together with the QE result with a 50% scaling, and the result if the magnet material had infinite permeability (1010 steel was assumed nominally). The NLC quadrupole model magnet would accommodate 17 turns of the standard sized conductor per quadrant compared to 35 turns in the nominal size QE design. This yields a conductor current density of 8.3 kA/in^2 at the maximum current setting (226 A) in Figure 7-78. The field harmonics of this design are also reasonable: 2.6% 12 pole and 2.9% 20 pole at the pole-tip radius, which easily meet the NLC field-quality requirements.

Based on this quadrupole design, quadrupole lengths were chosen to match the linac optics requirements plus allow some overhead (generally more than 30%) at the highest energy operation. Five lengths were chosen which range from 0.2 m to 0.8 m. The operating properties of these magnets can be estimated from Table 7-10 which lists the properties of a somewhat shorter (0.1 m) and longer (1.0 m) magnet that is operated at high current.

Property at I = 225 A	0.1-m magnet	1.0-m magnet
Power Dissipation	1.1 kW	6.9 kW
Water Flow	0.5 gpm	1.5 gpm
Water Δ T	8.7°C	17°C
# of Water Circuits	2	8
Total Coil Resistance	26 mΩ	160 mΩ
Magnet Inductance	0.96 mH	9.6 mH
Time Constant	37 ms	59 ms

Table 7-10. *NLC linac quadrupole magnet properties for two magnet lengths.*

7.9.3 Quadrupole Power Supplies

Since the maximum power dissipation levels in the quadrupoles will be under 10 kW, the quadrupoles can be powered individually at a cost that is competitive with powering them in strings and using shunts. This is because much of the cost is in the control electronics, and because the individual power supplies can be of the simple rack-mounted variety. Having individual control of the quadrupoles is very desirable for beam-based alignment measurements, so we will likely power them individually. This also eases the supply stability tolerance—a regulation of 6×10^{-5} will be required, which should not be difficult to meet.

7.9.4 Multibunch Kicker

The kicker for the bunch-by-bunch trajectory correction is assumed to be a terminated TEM-mode parallel plate vacuum transmission line. The electric and magnetic kicks are equal in magnitude, but cancel unless the TEM wave is traveling in the direction opposite to the beam, so it is not possible to synchronize the wave with the beam. If the stripline length is shorter than half the bunch spacing, different bunches are deflected by entirely independent parts of the driving waveform.

The Fourier spectrum of the kicker was assumed to cut off abruptly at some frequency for the beam dynamics studies. There was no improvement for cutoffs above 250 MHz, which is lower than the bunch frequency of 714 MHz. This indicates that it is not necessary to have independent control of each bunch. A stripline of length 60 cm would have a zero-crossing in its response at 250 MHz, but the response is not flat. A 30-cm stripline length would have a flatter response out to 250 MHz, and would have its first zero-crossing at 500 MHz.

The kicker strength available was assumed to be 56 keV/c for the beam dynamics studies. For $E = 100$ GeV, $\epsilon = 3 \times 10^{-8}$ mr, and $\beta = 10$ m, we have $\sigma_\theta \approx 0.1\,\mu$r. This is equivalent to $\sigma p_\perp \approx 10$ keV/c. Thus the kicker strength is about 6σ in the beam dynamics simulation.

A deflection of 10 keV/$c \approx 1\sigma =$ requires about 30 Gauss-cm. For a deflection of 10σ, the magnetic field integral is $BL \approx 300$ Gauss-cm $= 3 \times 10^{-4}$ T m. Generating a magnetic field B between strips of width w separated by gap g requires a current $I = Bwf/\mu_0$, where f is a geometrical form factor that depends on g/w and is of order 2. In terms of BL, this can be written $I = (BL)(w/L)f/\mu_0$. For $L = 30$ cmfor independent bunch control, and $w = 1$ cm, a 10-σ deflection requires 16 A of current. The impedance of the stripline structure is $Z \approx (g/w) \times 377\Omega$. Assuming $Z = 200\Omega$, 3.2 kV are required for 16 A. These parameters are within the range of solid-state components. The instantaneous power is 50 kW, but with a very low duty factor.

Chicane	Max. Energy (GeV)	Num. Bends	Max. B (kG)	$L_{\rm eff}$ (m)
1st	32	4	5.25	1.60
2nd	80	4	2.26	3.50
3rd	165	8	1.63	2.65
4th	290	8	1.27	3.50

Table 7-11. *Dipole magnets for the linac diagnostic chicanes.*

There are several conservative assumptions in the above estimates. The kicker strength is 100 keV/c, where 56 keV/c was adequate. The electric kick was neglected, though it is equal to the magnetic kick. Doubling the strip length would allow the same kick strength at half the voltage and current, although it would cost some strength near the 250-MHz cutoff.

The striplines would see the full beam image current, which would be coupled into the input and output ports. The coupling should be very strong, to avoid trapped modes on the strips. The drivers and terminations would be exposed to the coupled image currents. Since the bunches are very short compared to the drive bandwidth, low-pass filters would protect the equipment. Reflections of the image currents would disturb subsequent bunches. A simple way to avoid this problem is to use cables or waveguides with delays longer than half the bunch-train length between the striplines and the filters, drivers, and terminators.

7.9.5 Chicanes

The chicanes are constructed from relatively low field rectangular bending magnets. They have 2-cm vertical gaps and a horizontal good field region $\Delta B/B \leq 10^{-3}$ of 4 cm; the magnet lengths and fields are listed in Table 7-11. All magnets in each chicane, *i.e.*, the four or eight magnets, are powered in series with a 20-kW power supply that is rated at 100 V and 200 A and has a stability of $\Delta I/I \leq 1 \times 10^{-4}$ over the period of one hour.

7.9.6 Beam Dumps

The present design does not have any beam dumps located along the linacs. There is a high power dump after the final diagnostic station at the end of each linac. These dumps will be able to absorb the full 16-MW beam power so that the linacs can be commissioned and studied without sending beams downstream. The design is the same as final beam dumps located after the IPs which is described in Chapter 11.

The beams will be deflected to the dumps using pulsed kickers that operate at the full 180-Hz repetition rate. This allows tuning of the linacs before sending full-rate beam downstream as required for the MPS. At this time we do not have designs for the dump line kickers, but they are straightforward magnets, similar to pulsed dumper magnets in the SLC.

Resolution (minimum variation meas.) at 10^{10} ppb	1-μm rms
Position stability during 24 hours timescale	1-μm rms
Position absolute accuracy	0.2 mm
Position dynamic range	0–2 mm
Number of bunches	1–90
Bunches separation	1.4 ns
Beam current dynamic range	$5 \times 10^8 - 1.4 \times 10^{10}$ ppb
Maximum cycle frequency	180 Hz
Maximum time for data processing	1 ms
Total number of monitors	3000

Table 7-12. *Requirements for "Q" BPMs.*

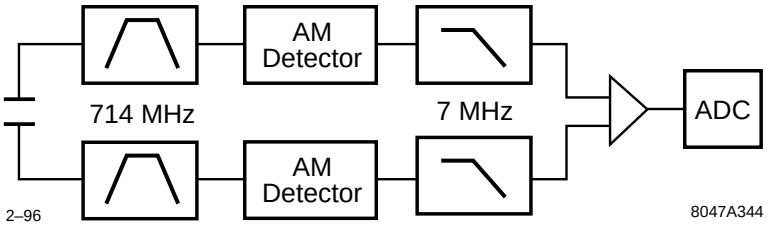


Figure 7-79. *"Q" BPMs signal processing block diagram.*

7.10 Instrumentation

The NLC Beam Position Monitor System (BPM) must measure the average beam position of a single pulse containing from 1 to 90 bunches, and must provide bunch-by-bunch position information to the feedback system. Different systems are used to accomplish these tasks: quadrupole BPMs ("Q"), feedback BPMs ("FB"), and structure BPMs ("S"). There are about 3000 "Q" BPMs, 200 "FB" BPMs, and 9700 "S' BPMs. The "Q" and "FB" BPM mechanical detectors are realized with four buttons or short striplines, two in each plane. The "S" BPMs are realized with rf couplers within the cavities. The BPM layout is not included in this study, and therefore no cable plan or signal attenuation is considered in the analysis.

General design considerations should be applied: keep as much electronics as possible out of the tunnel for low radiation and easy maintenance, keep as much preprocessing as possible at the BPMs electronics rack location to reduce the communication bandwidth and expedite the feedback system.

7.10.1 Quadrupole and Feedback Beam Position Monitors

The quadrupole BPMs ("Q" BPMs) must be able to detect the average position of the train of bunches, that can contain from 1 to 90 bunches in the single pass. The system's requirements are shown in Table 7-12.

The toughest requirement is the position stability, which asks for better than 1-μm systematic error. This means 0.005-dB stability for a 6-mm-radius circular beam pipe. In order to achieve such a good performance, a special calibration scheme is necessary.

A synchronous detection technique is used to process the signals from the striplines, which amplitude demodulates the signal and provides either linear or logarithmic functions, as shown in Figure 7-79.

The expressions to calculate the position are respectively:

$$x = \frac{b}{2}\frac{A-B}{A+B} \tag{7.21}$$

for linear demodulation, where b is the beam pipe half aperture, and:

$$x = \frac{1}{S_x}\log\left(\frac{A}{B}\right) \tag{7.22}$$

for logarithmic demodulation, where S_x is the sensitivity (in dB/mm), dependent on the electromagnetic sensor geometry.

The expected signal level for a gaussian beam with a bunch length much shorter than the bunch spacing, is given by:

$$V(\omega) = \sqrt{2}\frac{Ne}{T}gZ\sin\left(\frac{\omega l}{c}\right) \tag{7.23}$$

where N (=10^{10}) is the number of particles per bunch, T (=1.4 ns) is the bunch spacing, g (=0.2) is the electrode coverage, Z (=50) is the impedance, l (=20 mm) is the stripline length, and ω is the processing frequency (= $2\pi \times 714$ MHz). Under these assumption the rms voltage at the stripline is:

$$V = 1.3V \quad . \tag{7.24}$$

The system's resolution is given by the following formula:

$$x[mm] = \frac{b}{4}\left(\frac{N}{S}\right) \tag{7.25}$$

where b is the beam pipe radius (in mm) and N/S is the voltage signal-to-noise ratio. One of the resolution limits is the thermal noise, which is:

$$P_n[dBm] = -174 + 10\log(BW) + NF \tag{7.26}$$

where BW is the electronics bandwidth and NF is the electronics noise, while the signal strength must be computed for the different type of electromagnetic pickup and electronic processing.

The rms noise voltage, using 7-MHz bandwidth and 15-dB noise, is:

$$V = 7\ \mu\text{m}\ V \tag{7.27}$$

and the expected resolution due to the thermal noise is:

$$x = 8\,\text{nm} \quad . \tag{7.28}$$

Another resolution limit is given by the amplifier's Common Mode Rejection Ratio (CMRR) which is defined as the ratio of response for a differential change signal when both signals have the same amplitude. A hard but achievable value is 80 dB, which corresponds to:

$$x = 150\,\text{nm} \tag{7.29}$$

according to the expression in Eq. 7.21. If ADCs are used to digitize the values and the difference is digitally calculated, 14 bits are required for a comparable resolution. Things get worse when additional dynamic range must

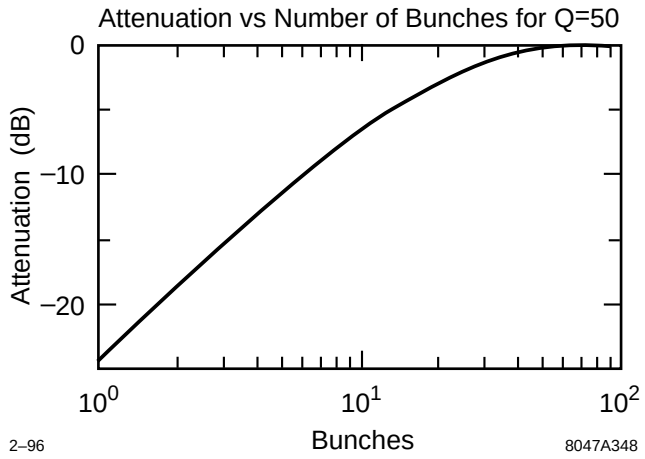


Figure 7-80. *Attenuation vs. number of bunches.*

Number of channels	4
Analog bandwidth	1 GHz
Sampling rate	10^9/s
Number of bits	10
Triggering	int/ext

Table 7-13. *Single-bunch BPM characteristics.*

be used for beam current fluctuation, because part of the CMRR must be used to take care of it. When the signal generated by the single bunch pass through the bandpass filter that has 714-MHz center frequency and $Q = 50$, to give 7-MHz bandwidth to the electronics, it produces a ringing whose amplitude is attenuated by 24 dB with respect to the 90 bunches case, as shown in Figure 7-80. This number was derived assuming the demodulation process bandwidth to be limited by the single-pole low-pass filter.

The expected resolution due to thermal noise in single-bunch s therefore:

$$x = 127\,\text{nm} \tag{7.30}$$

to which the CMRR contribution should be added as well.

7.10.2 Feedback BPMs

The “FB” BPMs must be able to detect the position of individual bunches within the pulse, for feedback correction. The same “Q” BPMs system requirements in Table 7-12 apply, the only difference being the resolution:

$$x = 0.25\,\text{nm} \quad . \tag{7.31}$$

The electronics processing is realized with four channels digitizers that sample the signal peaks from the striplines. An example of such a module is the Tektronix TVS545, a $20k VXIbus based module whose pertinent characteristics are shown in Table 7-13.

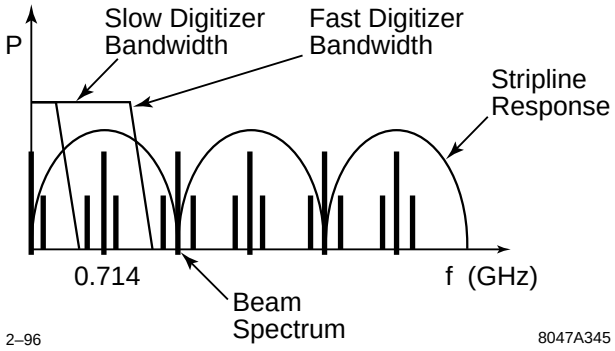


Figure 7-81. *Principle of "FB" BPMs*

The trigger is either generated on the signal's peak, or generated by the timing system. It is very important to minimize the trigger jitter, because it produces 1-μm position error each 5 ps jitter, given the 12-mm beam pipe full aperture. A possible solution to this problem is to oversample the signals and to reconstruct the waveform, but higher sampling rate than currently possible will be required. This technique will reduce the jitter performance requirements and will not require the clock to be beam synchronous. The expected signal level peak for a gaussian beam with a bunch length much shorter than the stripline length, is given by:

$$V = \frac{gZ}{2}\frac{Ne}{\sqrt{2\pi}\sigma}\left(1-\exp\left(-\frac{\tau^2}{2\sigma^2}\right)\right) \approx \frac{gZ}{2}\frac{Ne}{\sqrt{2\pi}\sigma} \quad . \tag{7.32}$$

Where N (=10^{10}) is the number of particles per bunch, g (= 0.2) is the electrode coverage, Z (= 50) is the impedance, $\tau = 2l/c$ (0.7 ns), l (=20 mm) is the stripline length, $\sigma = 1/(2\pi BW)$, and BW is the filter -3 dB point. Under these assumption the rms voltage at the stripline is:

$$V = 14V \tag{7.33}$$

and the rms noise voltage, using 1-GHz bandwidth and 15-dB noise, is:

$$V = 80\,\mu\mathrm{m}\,V \quad . \tag{7.34}$$

The expected resolution is:

$$x = 9\,\mathrm{nm} \quad . \tag{7.35}$$

Another resolution limit is given by the ADC number of bits. The TVS545 module, with its 10 bits, guarantees:

$$x = 1.5\,\mu\mathrm{m}\,\mathrm{m} \quad . \tag{7.36}$$

In order to achieve the required resolution performance, the ADC must provide at least 13 bits or the use of hybrids must be included to reduce the system's required CMRR.

If the technology becomes available and the cost is affordable it might be possible to include the "FB" BPMs in the "Q" BPMs using a wideband AM detector, working, for example, with 11.4-GHz carrier and 350-MHz bandwidth. An interesting possibility is to reduce the system bandwidth to 350 MHz and maintain the oversampling process. The equivalence of the two processes is shown in Figure 7-81 where the beam spectrum is relative to a 714-MHz pulses repetition, with sidebands indicating amplitude modulation due to current and position fluctuation, and the length of the stripline is adequate to peak the pulse repetition rate. The fast digitizer samples synchronized with the bunch spacing, and in this way amplitude demodulates the signal, providing the equivalent power of the first three sidebands. When the slow digitizer is used, there is still signal available, but the power is only given by the first sideband.

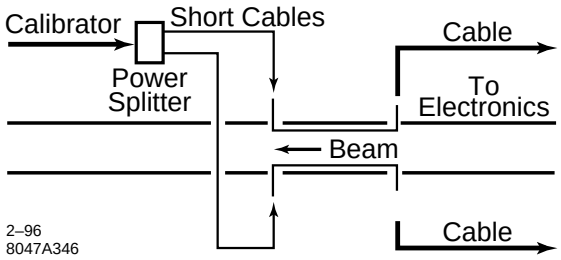


Figure 7-82. *Calibration scheme.*

The best condition for the slow digitizer option is when the length of the stripline is approximately 15 cm: in this case its response notch corresponds to the bunch repetition frequency.

This slow digitizer principle is different from the amplitude demodulation in the sense that there is no DC signal available because of the stripline transfer function. A way of defining this type of signal processing would be: AC coupled amplitude demodulation, where the modulation signal is the position variation versus time. It is important to notice that when the beam doesn't move there is no modulation, and therefore no signal. In a sense this type of signal processing provides the modes of the position variation: if the beam presents a bunch-by-bunch displacement, the output of the BPM will show a 357-MHz line in the frequency plot, if it presents a sinusoidal displacement with period 10 bunches, the output will be a 71.4-MHz line in the frequency plot, and so on.

It is possible to deduce from the picture in Figure 7-81 that the amplitude of the signal increases with the frequency of the modulation signal, that is the position variation. It is possible, in principle, to recover the signal information using the transient at the beginning of the pulse. The idea is to calibrate the system with a single bunch, measuring the impulse response. The position is then available by deconvolving the measured signal with the impulse response.

Calibration

In order to achieve 1-μm stability from the "Q" and "FB" BPMs, which corresponds to 0.005 dB for a 6-mm-radius circular beam pipe, it is necessary to include an efficient calibration scheme.

The calibration scheme in Figure 7-82 includes a signal to be injected at the other end of the striplines. The power splitter divides the signal between the two striplines simulating the beam on center. In this way both electronics and cables will be calibrated. This calibration scheme is more expensive than the conventional one, where the signal in injected at the electronics through a combination of three power splitter/combiner, because each stripline needs two feedthroughs instead than one and each BPM needs one more long cable for the calibration signal. The possible sources of instability are the power splitter unbalance and the short cables insertion loss. Off-the-shelf stripline power splitters have 0.2-dB typical amplitude unbalance, which corresponds to 40 μm in our case, and 0.005-dB drift over 10°C temperature change.

The short cables insertion loss variation can be kept under control by using good cables as short as possible. A typical temperature change for 12-ft coaxial cable is 0.05 dB/10°C: that means that the cables must track to better than 10% in order to maintain within the required tolerance.

7.10.3 Structure Beam Position Monitors

Overview

In this section, we outline the physical scalings and technique for use of accelerating structures as monitors of beam position and beam phase. We also discuss briefly their employment as monitors of higher-order moments.

The "S" or structure BPMs are the damped detuned accelerating structures, operated parasitically as beam position monitors. In this structure, as described in Chapter 8, four nearly rectangular vacuum manifolds are situated parallel to and mated in the structure in such a way as to couple out deleterious "dipole" modes which otherwise would disturb the beam. These manifolds effectively lower the Q of the dipole modes of the structure, reducing the multibunch wakefield and aiding in the control of emittance growth. At the same time these manifolds provide a conduit for microwave signals that can provide information about the beam. Our use of this structure as a BPM is entirely parasitic to this manifold-damping.

The S-BPM allows one to monitor the transverse position of the bunch train with respect to the *electrical center* of the front, rear, or other specific portions of the structure. This is distinguished from other BPMs that measure the transverse position of the bunch train (Q-BPMs) or a single bunch (F-BPMs) with respect to a locus (the BPM structure centroid) with no *a priori* dynamical significance, except that established by accurate alignment with respect to other beam line elements (*e.g.*, quadrupole centroids). Misalignments revealed by an S-BPM could (on a time-scale to be determined) be subject to feedback to zero them out, as such offsets give rise to deflecting wakefields that tend to produce emittance growth. The feedback might take the form of a pair of movers (X and Y) at the front and rear of each structure.

Physical Principles

Structure "dipole" modes are rather similar to the TM_{11} modes of cylindrical guide, insofar as they couple linearly (for sub-mm displacements) to the beam offset from the electrical center of the structure. However, each structure consists of some $N_{\rm cell} = 206$ coupled cavities or "cells," and thus there are some 206 different dipole modes, each of which tends to be synchronous with the beam over only a limited region of a few cells. For the DDS as presently designed, the mode frequency is correlated with position along the structure, such that the lower frequency range (14 GHz) modes tend to couple well to the beam near the entrance of the structure, while the higher frequency range modes (16 GHz) tend to couple well near the exit. Thus the filtered microwave signal emanating from the manifold ports provides information on the misalignment of the beam relative to different portions of the structure. The raw phase and amplitude of a dipole mode can only be interpreted as a signed displacement if we have knowledge of beam phase and amplitude. The phase information will come from the timing system and the amplitude will come from "Q" BPMs located in the nearest quadrupole. We expect to be able to measure the position of trains of 90 bunches of 5×10^9 particles per bunch to a resolution of better than 1 micron rms for at least several modes corresponding to both ends of each structure.

Some simple scalings can be employed to describe the rf structure as a pick-up. In one 1.8-m structure the $N_{\rm cell}$ modes will be distributed in frequency from 14–16 GHz (roughly a mode-spacing of 10 MHz) and a mode $Q_e \sim 1000$, corresponding to a decay time $2Q_e/\omega \sim 20$ ns. Modes will overlap since the typical FWHM = $\omega/4\pi Q_e \sim 15$ MHz. The energy deposited in one mode by a single bunch of charge Q_b is $U = k_l Q_b^2 (x/a)^2$, with x the beam offset, a ~0.3–0.7 cm the iris radius, and k_l ~1–4 V/pC/cell the loss factor for the mode in question. For a bunch train, the figure for stored energy is U multiplied by a train form factor F, which is a function of mode frequency and the number of bunches in a decay-time, $N_{bf} \sim 14$. This form factor can be computed by convolving a periodic sequence of delta-functions with a damped sinusoid. For modes with frequencies that are integer multiples of the bunch train frequency

(714 MHz), $F = N_{bf}^2 \sim 200$. These frequencies are 14.280 GHz (0.714 $\times$ 20), 14.994 GHz ($\times$ 21), 15.708 GHz ($\times$ 22), 16.422 GHz ($\times$ 23). For modes that are half-integer multiples of the bunch train frequency, $F = 1/4$. These frequencies are 14.637 GHz ($\times$20.5), 15.531 GHz ($\times$ 21.5), 16.065 GHz ($\times$ 22.5). For other modes, $F \sim N_{bf} \sim 14$. A rough estimate of peak power flowing through one manifold arm from a given mode is just $P \sim \omega U/2Q_e$. For the time being, we adopt this as an estimate for power coupled out of the manifold port, recognizing that the actual figure will depend on attenuation and propagation characteristics in the manifold.

To appreciate these scalings, consider the extreme cases. Maximum peak power from one mode, resonantly excited by a bunch train steered into an iris would be 1×10^3 W, assuming $k_l \sim 4$ V/pC/cell, and 5×10^9 particles per bunch (0.8 nC). Peak power in the full bandwidth (14.5–16.5 GHz) would be about 1×10^4 W. Average power deposited in this case would be less than 1 W. Minimum power would correspond to the case of a single bunch of 3×10^9 particles (0.5 nC) offset by 5 μm, $P \sim 6\mu\,\mathrm{W}(-22\,dB)$, assuming $k_l \sim 1$ V/pC/cell and $a \sim 0.7$ cm. For comparison note that thermal noise in any 15-MHz interval is $N \sim -102$ dB, roughly the power level corresponding to a 0.5-nm beam offset. In addition one wants to compare this figure with that for the $m \neq 1$ terms in the impedance (both broadband and resonant terms if there is any coincidental overlap) over the bandwidth of interest, a figure not yet calculated.

Technique

Each structure has four damping manifolds. The manifolds must be terminated on each end to dissipate the higher-order mode power. Two of the terminations, one on a horizontal manifold and one on a vertical manifold, will be provided by waveguide to coaxial vacuum feedthroughs which bring the higher-order mode power out onto external cable to provide the position signals. The other six terminations are probably internal to the manifold to minimize costs. Coaxial cable can be used since a run of no more than 80 ft will extend out of the accelerator housing proper to a front-end electronics rack. Cable losses are high in this frequency range, *e.g.*, ~30 dB/100′, but signal power is plentiful if we're only trying to achieve micron resolution. No special requirements for temperature-stability of the cable are envisioned.

Rf pre-processing modules are located at the top of the nearest tunnel penetrations. These integrated microwave modules bandpass filter, attenuate (through fixed and programmable attenuators), and selectively mix down any frequency in the dipole band to a fixed intermediate frequency (IF). Mode selection is accomplished by tuning the local oscillator provided at each crate of rf processing modules. The down-conversion mixers can be protected from burnout by a combination of fixed attenuation and PIN diode limiters even if programmable attenuators are inappropriately set. The IF signals are pre-amplified and sent to the nearest instrumentation alcove for further processing.

In the instrumentation alcove is a VME crate containing IF processors for the structure BPMs. Each IF processor measures the phase and amplitude of the X and Y IF signals from each of four structure BPMs. The input signals are quadrature demodulated, sampled at a time specified by the timing system, and digitized. Phase rotation and amplitude normalization is provided by an onboard signal processor. The processor will provide simultaneous X and Y position measurements for all four structure BPMs on the same beam pulse. Since the local oscillator is shared across BPMs, they all measure the same mode (frequency) on any one machine pulse.

Other Issues

We intend to study two additional signals as possible candidates for the list of raw signals to be processed. These are a second-moment coupled signal (TM_{21}-like, at f_2) and a third-moment signal (TM_{31}-like, f_3). Such signals derived from several structures can provide single-shot measurements of beam emittance and beam profile asymmetry, at least, in the horizontal. A rough estimate of energy deposited in such modes would be $U = k_l Q_b^2 (x/a)^{2m}$ where $m = 2$ or 3. Assuming $Q_e \sim 5000$, $k_l \sim 1$ V/pC/cell and $a \sim 0.7$ cm, and a single bunch of 3×10^9 particles (0.5 nC), this

corresponds to $P_2 \sim -40$ dB for a 100-μm beam feature and $P_3 \sim -77$ dB. These numbers indicate that pursuit of still-higher moments is probably not practical. The third moment is looking a bit problematic for a single bunch.

The contribution of the $m \neq 1$ terms in the impedance to the 14.5-16.5-GHz-band signal offset will be studied at ASSET.

Conclusion

The electronics alluded to here and the structure itself are still being designed and will be tested with beam during the next ASSET experiment and at the NLCTA. It is expected that modifications and improvements will derive from practical operating experience.

Primary issues will be cost, reliability, dynamic range, and circuit protection. For a simple nonsigned S-BPM, the primary cost is the GADC unit. For a signed S-BPM, the primary cost is likely to be in the microwave electronics (LOs, mixers, etc.). An assessment of reliability awaits experimental studies at ASSET. The need to protect the processing electronics in the worst case (bunch train steered into an iris) will probably limit the useful dynamic range.

To summarize, S-BPM resolution is favored by minimum insertion loss for microwave signals emerging from the manifold. The S-BPM will be capable of resolving 5-μm beam offsets shot-to-shot relative to specific portions of the structure, *e.g.,* near the entrance and the exit. The limits of S-BPM resolution and functionality, as well as tests of the electronics await experimental studies at ASSET.

7.10.4 Laser Wire Scanners

Laser-based beam profile monitors, to be used with damped beam, will be described in this section. The laser-based profile monitors will be similar to the one under construction for the SLC/SLD IP. This device uses a diffraction limited, 350-nm wavelength, laser. The laser beam will be scanned across the particle beam on a consecutive sequence of pulses. The system will have four components, a single synchronized laser, a transport line to deliver photons to several interaction chambers, the interaction chambers themselves and scattered radiation detectors that detect boosted photons and/or degraded e^+ or e^- from the interaction of the laser light with the particle beam. The amplitude of the light can be varied in order to accommodate a very wide dynamic range of beam size and intensity.

Laser

A laser will be used to provide a 10-mW peak power pulse to six interaction chambers. In these chambers the particle beam will be brought into collision with the focused laser beam. The light pulse will be 0.1-ns-long, short enough to interact with only one of the 90 bunches. The synchronization of the light pulse with the particle beam will done using a mode locked seed laser. The seed laser output can be amplified using a regenerative YAG 20-pass amplifier. Expected synchronization accuracy is at worst 10 ps. The laser phase space is defined by a high-power pinhole spatial filter. The laser will be located in an accessible area.

The laser can be operated in the lowest transverse spatial harmonic mode or in a higher mode. In the first mode, a null is generated where the center of the gaussian beam would ordinarily be. Two side lobes are arranged around the null and this structure can be scanned across the beam as an alternative to scanning the gaussian zero mode spot. The spacing between the lobes provides an estimate of the performance of the laser focusing and transport optics.

Transport

The transport is used to bring the laser light to up to one of six interaction chambers. It can be of roughly any length provided the surface figure of its optics is good enough. Typical surfaces must be l/40 or better. Starting from the spatial filter and progressing through the transport and the interaction chambers, the accumulated optical errors must be kept below a modest surface figure budget of about a half wave. The transport must be evacuated and free from significant contaminants. Steering elements are required for trajectory stabilization and optimization. Diagnostic imagers are needed to maintain transport alignment and monitor optical quality.

Interaction Chambers

The function of the interaction chambers is to focus the light to a spot small compared to the expected beam size. This can be done using an F2 focusing system with a 95% transmission aperture. In this system demagnification is traded off against diffraction-related tails and distortion. Absorbing materials are needed to keep the optics free from the contamination caused when metal material is removed by stray laser light.

The locations of the interaction chambers will be dictated by the local optics. For the measurement of an uncoupled beam matrix a minimum of three scanners, separated by 45 degrees of optical phase advance, are required. For the determination of a fully-coupled beam matrix, six scanners are needed. The placement of the scanners in the diagnostic stations is described in Section 7.3.4.

The laser beam will be scanned across the particle beam using piezo-electrically-driven mirror actuators. The laser beam can either be scanned slowly, so that it advances a fraction of a bunch size on each successive machine pulse, or scanned rapidly, so that it crosses the entire bunch train during a single passage. In the latter case, an estimate of the projected full train emittance can be derived from the response of a single high-bandwidth detector or group of detectors. The focused laser spot must move <4 sigma during the 100-ns passage time. Fine synchronization is required to time the collision between the light and the particles.

Detectors

Scattered radiation detectors are needed to detect the degraded particles and the boosted photons. In the latter case, the beam line geometry requires a bend to remove the electron beam from the path of the photons. A very large signal of about 1000 photons is expected for the single-bunch parameters and laser power listed above. In some cases the Compton scattered signal may be used as an indication of the degree of polarization of the beam.

7.11 Discussion

In this chapter, we have presented the design of the NLC main linacs and have described how the linacs will be operated so as to prevent large dilutions of the beam emittances. We have presented the emittance dilution budgets for the different operating modes of the linacs: NLC-Ia—NLC-IIc. The budgets are generous and have been chosen to allow for reasonable fabrication and construction tolerances.

As discussed, the approach to preserving the small beam emittances is two-fold. First, many problems have been explicitly `designed-out'. In particular, the damping and detuning of the accelerator structures will make the wakefield effects associated with multibunch operation manageable. Similarly, the thermal regulation in the tunnel and gallery

will prevent significant fluctuations in alignment and instrumentation response. Second, we will utilize beam measurements to suppress unavoidable imperfections that are not readily solvable by design. These include the alignment of the quadrupole magnets and accelerator structures and the stabilization of the beam orbits and energies.

To perform the alignment, each quadrupole and pair of accelerator structures will have remotely adjustable supports. With this control, the only tight alignment tolerance that will have to be achieved prior to operation is that for the straightness of the accelerator structure pairs. The beam-based alignment algorithms that will be used are fairly straightforward. The quadrupole alignment method is a variation of one that has been successfully applied in the SLC and FFTB, although it places higher demands on the BPM stability and the performance of the control system. The structure alignment technique is more novel in that it uses the signals from the structure damping manifolds to infer the beam offset in the structure. With this information, each pair will be aligned to the beam after a suitable orbit has been established.

It is important to note that these beam-based alignment techniques do not need to interrupt standard operation and we have shown that the emittance growth due to slow alignment drifts (ATL-model) can be controlled by re-aligning every 30 minutes. To further stabilize the beams, there are trajectory and energy feedback loops, like those used in the SLC, implemented throughout the linacs. These feedbacks will heavily suppress the effect of low-frequency (<0.1 Hz) sources of beam jitter and drifts. At higher frequencies where the feedbacks are not effective, we do not foresee any significant sources of jitter. The primary source of high frequency jitter is due to ground motion which, as discussed in Appendix C, we have measured and found to be well below tolerance.

The estimates of the emittance growth in both NLC-I and NLC-II are based in part on detailed simulations of beam transport in the linacs that include the significant deleterious effects and have been calibrated against SLC experience. Smaller effects have been estimated analytically with guidance from SLC observations. Furthermore, we have a number of fallback methods available to control the emittance growth, such as orbit bumps, which we are confident will provide additional suppression.

In the future, we plan to continue to optimize the linac design. Additional simulations and experience from operating the NLCTA and the SLC will guide the work. In particular, the NLCTA will give us a measure of the energy gain stability and beam-loading compensation in an NLC-like multibunch environment. In addition, there are number of dedicated studies planned for the next year that will develop or verify other key design concepts. These include tests of the structure alignment technique, development of stable stripline BPMs, and monitoring of the long-term straightness of the accelerator structures. From this work, we will be able to transform the linac design to a full CDR in the near future.

References

[Adolphsen 1993] C. Adolphsen and T.O. Raubenheimer, "Method to Evaluate Steering and Alignment Algorithms for Controlling Emittance Growth", Presented at the 1993 IEEE Part. Acc. Conf., Washington D.C. (1993).

[Assmann 1995] R. Assmann, C. Adolphsen, K. Bane, T.O. Raubenheimer, R. Siemann, K. Thompson, "LIAR: A Computer Program for Linear Collider Simulations", SLAC–AP–NOTE–103 (1996); to be published as a SLAC–REPORT.

[Balakin 1983] V. Balakin, A. Novokhatsky, V. Smirnov, "VLEPP: Transverse Beam Dynamics", Proceedings of the 12th Int. Conf. on High Energy Accelerators, Fermilab, 119 (1983).

[Bane 1980a] K. Bane and P. Wilson, "Longitudinal and Transverse Wake Potentials in SLAC", Proceedings of the 11th Int. Conf. on High Energy Accelerators, CERN, Birhäuser Verlag, Basel, 592 (1980).

[Bane 1980b] K. Bane and B. Zotter, "Transverse Modes in Periodic Cylindrical Cavities", Proceedings of the 11th Int. Conf. on High Energy Accelerators, CERN, Birhäuser Verlag, Basel, 581 (1980).

[Bane 1993] K.L.F. Bane and R.L. Gluckstern, "The Transverse Wakefield of a Detuned X-Band Accelerating Structure", *Particle Acclerators* **42**, 123 (1993).

[Bane 1994] K.L.F. Bane, C. Adolphsen, K. Kubo, and K.A. Thompson, "Issues in Multibunch Emittance Preservation in the NLC", Fourth European Particle Accelerator Conference, London, England, (June 27-July 1, 1994).

[Bane 1995] K. Bane, "The Short-range NLC Wakefields", NLC–Note 9, (February 1995).

[Bowden, 1996] G. Bowden, P. Holik, S.R. Wagner, G. Heimlinger, R. Settles, "Precision magnet movers for the Final Focus Test Beam", *Nucl. Instr. and Methods* **368A**, 579–592, (1996).

[Brandt 1982] D. Brandt and B. Zotter, "Calculation of the Wakefield with the Optical Resonator Model", CERN-ISR/TH/82-13 and LEP Note 388 (1982).

[Brandt 1984] D. Brandt, "Optical Resonator Model for the Wakefields of the Deflecting mode $m = 1$", CERN-LEP Note 484 (1984).

[Gluckstern 1989] R. Gluckstern, "Longitudinal Impedance of a Periodic Structure at High Frequency", *Phys. Rev. D* **39**, 2780 (1989).

[Heifets 1989] S.A. Heifets and S. Kheifets, "High-frequency Limit of the Longitudinal Impedance of an Array of Cavities", *Phys. Rev. D* **39**, 960 (1989).

[Keil 1972] E. Keil, "Diffraction Radiation of Charged Rings Moving in a Corrugated Cylindrical Pipe", *Nucl. Instr. and Methods* **100**, 419 (1972).

[Kroll 1994] N. Kroll *et al.*, "Manifold damping of the NLC detuned accelerating structure", 6th Workshop on Advanced Accelerator Concepts, Lake Geneva, WI, June 12-18, (1994).

[Palmer 1990] R. Palmer, "A Qualitative Study of Wakefields for Very Short Bunches", *Particle Acclerators* **25**, 97 (1990).

[Panofsky 1956] W. Panofsky and W. Wenzel, "Some Considerations Concerning the Transverse Deflection of Charged Particles in Radio-frequency Fields", *Rev. Sci. Instr.* **27**, 967 (1956).

[Raubenheimer 1992] T.O. Raubenheimer and P. Chen, "Ions in the Linacs of Future Linear Colliders", presented at LINAC 92, Ottawa, (1992).

[Raubenheimer 1995] T.O. Raubenheimer and F. Zimmermann, "A Fast Beam-Ion Instability in Linear Accelerators and Storage Rings", *Phys. Rev. E* **52**, 5487 (1995).

[Schnell 1987] W. Schnell, "Microwave Quadrupoles for Linear Colliders", CERN-LEP-RF/87-24, (1987).

[Schnell 1991] W. Schnell and I. Wilson, "Microwave Quadrupole Structures for the CERN Linear Collider", Proc. of the 1991 IEEE Particle Acc. Conf., San Francisco, 3237, (1991).

[Thompson 1990] K.A. Thompson and R.D. Ruth, "Controlling transverse multibunch instabilities in linacs of high energy linear colliders", *Phys. Rev. D* **41**, 964 (1990).

[Thompson 1993] K.A. Thompson and R.D. Ruth, "Simulation and Compensation of Multibunch Energy Variation in NLC", in Proceedings of 1993 Particle Accelerator Conference, Washington, D.C. (17–20 May 1993).

[Thompson 1994] K.A. Thompson *et al.*,, "Design and Simulation of Accelerating Structures for Future Linear Colliders", *Particle Acclerators* **47**, 65 (1994).

[Weiland 1983] T. Weiland, "On the Computation of Resonant Modes in Cylindrically Symmetric Cavities", *Nucl. Instr. and Methods* **216**, 329 (1983).

Contributors

- Chris Adolphsen
- Roberto Aiello
- Ralph Assmann
- Karl Bane
- Gordon Bowden
- Franz Josef Decker
- Linda Hendrickson
- Tom Mattison
- Kiyoshi Kubo
- Yuri Nosochkov
- Tor Raubenheimer
- Kathy Thompson
- Dave Whittum
- Frank Zimmermann

8

RF System for the Main Linacs

Contents

8.1 Introduction

The basic design of the NLC main linacs rests on global experience gained from the design, construction, and 30 years of operation of the 3-km-long SLAC linac, which is powered at a frequency of 2.856 GHz [Seeman 1991, Seeman 1993]. Since its initial operation in 1966, the SLAC linac has been continuously upgraded for higher energy, higher intensity, and lower emittance.

The radio frequency (rf) system for the NLC main linacs is similar in character to the SLAC linac. The SLAC linac is currently energized by 240 high-power S-band klystrons. The klystron peak power and pulse duration are, respectively, 65 MW and 3.5 μs. The power from each klystron is compressed by a SLED pulse compressor, and then split to feed four, 3-m-long, constant-gradient S-band accelerator structures operating in the $2\pi/3$ mode.

When the SLAC linac was built, the accelerating gradient was 7 MV/ m. The original design included a future upgrade path in which the number of klystrons would be quadrupled. The upgrades that were eventually implemented involved replacing each of the initial 24-MW klystrons with a single higher-power klystron (first with 35-MW, XK-5 tubes and, later on, with 65-MW, 5045 tubes), and adding a SLED pulse compressor downstream of each klystron. For present-day SLC operations, fully upgraded with 240 SLEDed, 65-MW klystrons, the accelerating gradient has been tripled, to 21 MV/ m, and the maximum beam energy is 60 GeV (for unloaded, on-crest operation).

8.1.1 Overview

The rf power system for the NLC's two high-gradient linacs that accelerate the electron and positron beams separately from 10 GeV to 250 GeV (in the initial design), and to 500 GeV or more (after the upgrade), operates at 11.424 GHz. This system includes all the hardware through which energy flows, from the AC line to the accelerator structures. Figure 8-1 shows one module of the rf system schematically, with emphasis on the flow of energy. Electrical energy is transformed at each stage shown in the diagram: the modulator converts AC power into high-voltage pulsed DC, the klystron transforms pulsed DC into high peak power rf, the SLED-II pulse-compression system increases the peak power by about a factor of four (at the expense of a reduced pulse width), and the accelerator sections convert rf power into beam power. Because of the high average rf power required to drive the accelerator structures, it is important that the highest possible efficiency be maintained for the processing and transmission of energy at every stage of the rf system.

The primary technical choice for the rf system is the use of the 11.424-GHz frequency. This frequency, high in the X-band (8.2–12.4 GHz), is exactly four times the operating frequency of the existing SLAC 60-GeV linac. The choice of such a high frequency, relative to existing high-energy linacs, allows higher accelerating gradient, shorter linac length, and lower AC power consumption for a given beam energy. Considering the size, weight, cost, and availability of standard microwave components, we have chosen a frequency in the X-band for a design that is upgradeable from an initial 250-GeV beam energy to 500 GeV or more. This choice requires the development of klystrons capable of delivering peak power significantly greater than previously achieved by commercially available X-band sources. As described in Section 8.4, klystrons which meet the 50-MW peak-power goal necessary for the initial 250-GeV beam energy in the NLC design have been developed, and are now operating in the Klystron Test Laboratory at SLAC.

The general parameters of the high-power rf system and its major subsystems (klystrons, modulators, rf pulse compressors, and the accelerator structure itself) are specified in Table 8-1. The set of parameters has been optimized to provide high acceleration gradient (35–64 MV/ m) for trains of bunches with moderate charge per bunch (1.2–1.8 nC). This optimization keeps single-bunch wakefields under control and reduces the beamstrahlung at the collision point to tolerable levels. The upgrade to 500-GeV beam energy (1-TeV center-of-mass energy) is accomplished by doubling

the number of modulators (as shown by dashed lines in Figure 8-1), and by replacing each 50-MW klystron with a pair of 75-MW klystrons. The total active length of linac must also be increased from 16,300 m to 17,700 m. The upgrade also includes improvements in the modulator and pulse compression systems to increase the rf system efficiency. The upgrade to 1-TeV center-of-mass energy is described in more detail in Section 8.1.2.

The rf accelerator structure (discussed in Section 8.2) is designed to be very nearly a constant-gradient traveling-wave structure. The design of the structure has been optimized to reduce the wakefield seen by trailing bunches. This has been accomplished by tailoring the cell-to-cell frequency distribution of the dominant deflecting mode to yield an initial Gaussian-like decay of the wakefield amplitude. On a longer timescale, the higher-order beam-induced modes of the structure will be damped by vacuum manifolds to which each cell of the structure is coupled. This structure is designated by the acronym DDS (damped detuned structure). The damping manifolds run parallel to the beam channel and are terminated into matched loads. (The slots that couple the cells to the manifolds are cut off to the fundamental accelerating mode.) This damping scheme will reduce the typical quality (Q) factors of the deflecting modes to about 1000. The first prototype 1.8-m-long accelerator section, which was detuned but not damped, was high-power tested up to a gradient of 67 MV/ m. The effect of the detuning in that first prototype section was demonstrated experimentally by using positron and electron bunches from the SLC damping rings as probe and witness beams, respectively. Another prototype 1.8-m section that is both damped and detuned is being manufactured and will be used for a similar test before it is installed in the NLCTA.

Obtaining the X-band peak power for the NLC has required the development of klystrons (Section 8.4) capable of delivering peak power significantly greater than previously achieved by commercially available X-band sources. Both the peak power and the pulse length have already been achieved by four solenoid-focused X-band klystrons at SLAC. These klystrons will be used to power the NLC Test Accelerator (Section 8.1.3). The most recent refinement of the klystron has achieved the peak power and exceeded the pulse length required for the NLC design and upgrade, at an efficiency of 48%. The solenoid which focuses the electron beam in the prototype klystrons has a weight of 750 kg and a power consumption of 20 kW. Currently nearing completion is the first prototype of a 50-MW klystron which is focused instead by a periodic permanent magnet (PPM) array of samarium cobalt ring magnets weighing about 9 kg. It is this klystron, which operates at a higher voltage and lower beam current for compatibility with PPM focusing, which is slated as the prototype for the NLC. Based on computer projections, the tube, designated X5011, is expected to operate at about 57% efficiency with 50 MW of peak output power.

Each high-power rf station consists of a pair of PPM-focused klystrons (50 or 75 MW) and a pulsed-DC energy delivery system (modulator) that is tightly integrated in design with the electron guns of the klystrons. The modulator system (discussed in Sections 8.5 and 8.7) includes a single-thyratron switch, a Blumlein pulse-forming network (PFN), a high-efficiency power supply for charging the PFN's capacitance, and a pulse transformer. Using a Blumlein PFN allows for a relatively low transformer turns ratio (7:1), which yields a reasonably fast rise time (0.3 μs), and hence, an improved efficiency.

The rf pulse-compression system (discussed in Section 8.3) is based on the SLED-II technique which is a modification of the SLED system currently in use on the existing 60-GeV SLAC linac. For SLED-II, the energy storage cavities of SLED are replaced by resonant delay lines in order to produce flat output pulses. A prototype SLED-II system has been tested up to compressed-pulse power levels of 200 MW. Transmission of high power with low loss is accomplished by using oversized waveguide components. Three SLED-II systems are being manufactured for the NLC Test Accelerator.

To achieve a highly mono-energetic multibunch beam pulse for the final focus, the beam loading induced by the bunch train must be compensated. The initial transient can be eliminated by pre-loading the sections by shaping the input rf pulse using an approximately linear rise of the field amplitude for one filling time of the structure. In this way, the first electron bunch will see a filled rf structure that appears to be in the steady state. As will be seen, this requires phase-agile control of the rf before it is amplified by the klystron.

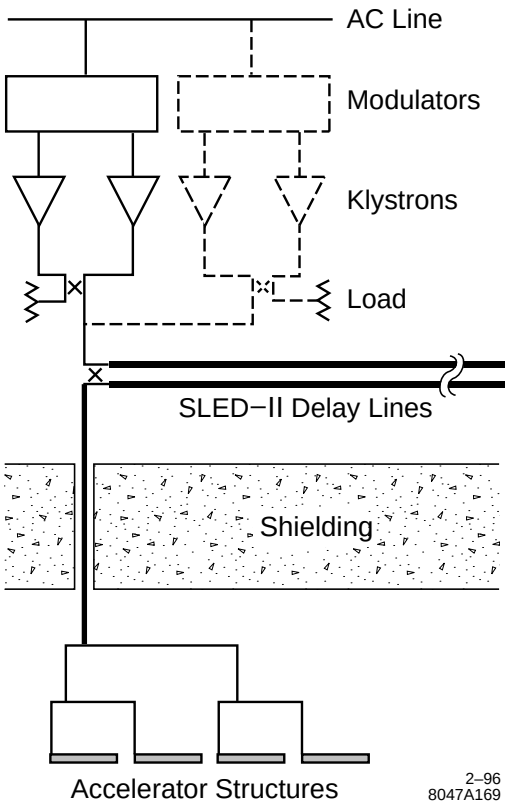


Figure 8-1. *One module of the high-power rf system for the main linacs. Dashes indicate additions for the beam-energy upgrade from 250 GeV to 500 GeV. Energy flows from the AC line to the accelerator structures.*

The design of the rf control system proposed for the NLC main linacs is based on experience derived from operating the SLAC linac. In particular, the methods of beam-loading compensation have been inspired, in part, by experience gained with beam-loading compensation while operating the SLAC linac in its long-pulse, high-average-current mode (for fixed-target experiments).

8.1.2 Upgrade to 1 TeV

As mentioned previously, the upgrade to 1-TeV center-of-mass energy is accomplished primarily by doubling the number of klystrons and modulators, and by increasing the peak power per klystron from 50 MW to 75 MW. Taking into account a small reduction in the power gain of the rf pulse compression system (which has been upgraded for higher efficiency at a lower compression ratio), this increase in rf power provides for an increase in the unloaded accelerating gradient from 50 MV/m to 85 MV/m. To obtain a center-of-mass energy of 1 TeV, the active linac length must also be increased slightly, from 16,300 m to 17,700 m. (If this additional length is not provided, the upgraded energy will be about 925 GeV).

A major part of the upgrade to 1 TeV will be to increase the efficiency of the rf system so that the AC wall-plug power is kept below 200 MW. After the 500-GeV design has been finalized, and before the 1-TeV upgrade is carried out,

	NLCTA Achieved	500-GeV Design Goal	1-TeV Upgrade
General Parameters			
Frequency (GHz)	11.4	11.4	11.4
Accel. Gradient (MV/m), Unloaded/Loaded	67/–	50/35.3	85/63.5
Overhead Factor,[a] F_{OH}		1.20	1.15
Active Linac Length[b] (m)	1.8	16,300	17,700
Total Linac Length (m) (1.08 × Active Length)		17,600	19,100
# 7.2-m RF Units and Pulse-Compression Systems		2264	2454
# Modulators		2264	4908
# Klystrons		4528	9816
Peak Power per Meter of Structure (MW/m)		50	145
RF Pulse Length at Structure (ns)	150	240	240
Repetition Rate (Hz)	60	180	120
Particles per Bunch (10^{10})		0.75	1.10
Number of Bunches per Pulse		90	90
Peak Beam Current (A)		0.86	1.26
RF Energy/Pulse at Structure Input (J/m)		12.0	34.7
Total Average RF Power at Structure[c] (MW)		34.1	71.3
Klystron			
Output Power (MW)	50, 75	50	75
Pulse Length (μs)	2.0, 1.1	1.2	0.96
Microperveance (μA/V$^{3/2}$)	1.2	0.6	0.75
Electronic Efficiency[d] (%)	48	57	60
Beam Voltage (kV)	440	465	487
Beam Energy per Pulse[e] (J)	310, 130	105	119
Focusing	Electromagnet	PPM	PPM
Cathode Loading (A/cm^2)	7.4	7.4	7.6
Overall Length (m)	1.3	1.3	1.3
Cathode Heater Power[f] (kW)		0.33	0.41
Modulator (Blumlein PFN, transformer ratio 7:1)			
PFN Voltage (kV)	48	68	71
Pulse Rise Time (ns)	500	275	175
Rise/Fall Energy Efficiency (%)		80	83
I^2R/Thyratron/Core Loss Efficiency (%)		97	97
Net Energy Transfer Efficiency (%)		77.4	80.6
$\frac{1}{2}$CV^2, two klystrons (J)		272	295
Power Supply Efficiency (%)		93	93
Net Modulator Efficiency (%)	58	72	75
Thyratron Heater + Reservoir Power[f] (kW)		1.5	1.5
Average AC Input Power (kW), Excl. Aux.	45	53	38

Table 8-1. *NLC main linac rf system parameters. (Continued on next page.)*

	NLCTA Achieved	500-GeV Design Goal	1-TeV Upgrade
RF Pulse Compression			
System Type	SLED-II	SLED-II	BPC/DLDS
Compression Ratio	5–7	5	4
Intrinsic Efficiency (%)	80.4–69.2	80.4	100
Loss Efficiency of Delay Lines, 3-db Coupler and Mode Converters (%)	92–90	95	93
Pulse Compression Efficiency (%)	73–64	76.5	93
Pulse Compression Power Gain	3.7–4.5	3.8	3.7
Power Transmission Efficiency (%)	90	94	94
Net Pulse-Compression Efficiency (%), Including Power Transmission Loss	66–58	72	87.5
Net Power Gain	3.3–4.1	3.6	3.5
Net RF System Parameters			
Total AC Power (MW), Excl. Aux.		116	181
RF System Efficiency (%), Excl. Aux.	19	29.6	39.4
Total Auxiliary Power[g] (MW)		5.4	12
Total AC Power, Including Auxiliary (MW)		121	193
RF System Efficiency (%), Including Auxiliary		28.2	37.0
Average Beam Power[h] (MW)		9.3	18.6
AC-to-Beam Efficiency (%)		7.7	9.4

[a] Includes overhead for BNS, feedback, and stations off for repair (see Table 7-1).

[b] Active length = F_{OH} ($E_0 - 20$ GeV)/(Loaded Gradient).

[c] Assumes 3% of klystrons and modulators are off (repair margin) or running off beam time (on standby).

[d] Given by simulated efficiency less 5 percentage points for 500-GeV design; equal to simulated efficiency for 1-TeV design.

[e] Useful energy in flat-top portion of pulse.

[f] Included in auxiliary power.

[g] Also includes power for modulators on standby (0.5%).

[h] Excludes injected beam power.

Table 8-1. *(continued): NLC main linac rf system parameters.*

several years of additional R&D will be possible in order to realize these potential gains in efficiency. The parameters listed in Table 8-1 for the 1-TeV upgrade are therefore somewhat less conservative than for the 500-GeV design.

In order to realize a substantial gain in the net rf system efficiency, each of the subsystems—klystrons, modulators, and pulse compression—must be examined for potential efficiency improvements. A slight gain is assumed in klystron efficiency (from 57% to 60%) by pushing closer to efficiency values given by simulations which are, in turn, expected to increase as experience is gained in klystron design. A modest improvement is assumed in modulator efficiency (from 72% to 75%), due mainly to a reduction in pulse transformer rise time. Several design approaches are being studied to reduce rise time. The greatest gain in rf system efficiency will result from an upgrade in the rf pulse compression system. The SLED-II system used in the 500-GeV design has a maximum intrinsic efficiency of 80.4%, even for lossless components, due mainly to reflected power during the period when the resonant delay lines are being charged with energy.

The SLED-II efficiency may be improved by using an active microwave switch to rapidly change the coupling (or Q) of the resonant delay lines [Tantawi 1995b]. Such a switch might be implemented as an optically-triggered silicon device operating in a low-field region of the microwave network. Experimental studies of optically-triggered silicon

devices for this purpose are underway at SLAC [Tantawi 1995c]. For a compression ratio of 4, the efficiency of the Q-switched SLED-II would be about 91%. Including the 94% power transmission efficiency, the net efficiency would be about 85% (a power gain of 3.4). Assuming the improvement in SLED-II efficiency can be realized, it remains to be determined that the power handling capability of the SLED-II configuration is adequate for the combined power of four 75-MW klystrons. (SLED-II has demonstrated its ability to handle a single 50-MW klystron. Tests with two 50-MW klystrons are expected to be performed in Summer 1996. Tests with four 75-MW klystrons will be performed when the 75-MW klystrons become available.)

An alternate upgrade path is to replace the SLED-II with a different type of rf pulse compression. By replacing this system with a Binary Pulse Compression (BPC) system [Farkas 1986, Lavine 1991], which has an intrinsic efficency of 100%, the net pulse compression efficiency (including power transmission losses) can be increased from 75% to 87.5%. At the same time, the compression ratio is reduced from 5 to 4, and the klystron pulse length is reduced from 1.2 μs to 0.96 μs. The higher compression efficiency together with the lower compression ratio results in a slight reduction in power gain from 3.6 to 3.5. The power-handling capability required of the BPC configuration is only half that of SLED-II because the BPC system has two outputs, each of which feeds only two structures, in contrast to SLED-II where a single output feeds four structures. The chief disadvantage of the BPC scheme is the longer length of delay-line pipe which is required (three times that for SLED-II). When upgrading from SLED-II, the components of the existing SLED-II delay lines can be reconfigured as the shorter of the two required BPC delay lines. However, an additional delay, twice as long, must be added to each system.

Variations of BPC can be utilized. Half of the BPC delay can be eliminated by the use of the Delay Line Distribution System (DLDS), as proposed at KEK. In the DLDS, rf energy is propagated upstream (toward the gun) by a distance equal to one half the required delay; the beam propagation time provides the other half of the delay. Loaded delay lines can also be used in principle to reduce the length of the added delay line.

Both of the above upgrade paths (adding Q switches to SLED-II, or replacing SLED-II with BPC or a variation of BPC) can be performed gradually, taking only one station offline at a time.

Since the BPC concept has been experimentally demonstrated [Lavine 1991], and the Q-switched SLED-II capable of handling four 75-MW klystrons has not, the upgrade path in Table 8-1 is based on the more conservative option of replacing the SLED-II systems with BPC systems. Additional R&D during the years after the 500-GeV design has been finalized, and before the 1-TeV upgrade, may make the Q-switched SLED-II (or other developments) possible, and more attractive.

Taken together, the above improvements lead to an increase in net rf system efficiency from about 30% for the 500-GeV design to almost 40% for the 1-TeV upgrade. Some of these potential design improvements may, in fact, be ready in time to be included in the final design of a 500-GeV machine. The 500-GeV parameters listed in Table 8-1 are, however, conservatively based on experience with, and measurements on, prototypes which exist at the present time.

8.1.3 The NLC Test Accelerator (NLCTA)

The design of the high-power X-band rf system for the NLC is based on specific experience gained from building X-band prototypes and operating them at high power, and on an rf systems-integration test—the Next Linear Collider Test Accelerator (NLCTA)—which is currently under construction at SLAC. The goals of the NLCTA project [SLAC 1993, Ruth 1993] are to integrate the technologies of X-band accelerator structures and high-power rf systems, to demonstrate multibunch beam-loading energy compensation and suppression of higher-order beam-deflecting modes, to measure any transverse components of the accelerating field, and to measure the growth of the dark current generated by rf field emission in the accelerator. The NLCTA design parameters and a possible upgrade path are summarized in Table 8-2.

The peak power and rf-system efficiency needed for the NLCTA have been demonstrated in a prototype system (discussed in Section 8.3.1). Upgrades to the NLCTA rf system will test the SLED-II design at the higher power levels and efficiencies needed for the NLC (Table 8-7). The NLCTA high-power rf system is depicted schematically in Figure 8-2. The system is comprised of four modules. Each module consists of a DC pulse modulator, up to two X-band klystrons (50 or 75 MW), a SLED-II pulse compressor, and two X-band accelerator sections. One module serves the injector. Three modules serve the linac. Power from the third and fourth modules will be combined in the pulse compressor of the third module, and then re-divided to energize the last four accelerator sections in order to test the topology that is proposed for the NLC.

The six accelerator sections in the NLCTA linac are each 1.8-m long. The two accelerator sections in the injector are similar to the linac sections, except that each is 0.9-m long to maintain beam loading comparable to the linac in the presence of approximately twice the current. All the X-band sections in the NLCTA will suppress transverse wakefields, either by cell-to-cell detuning [Thompson 1993], or by a combination of detuning and damping [Kroll 1994], as discussed in Section 7.4.2. The effect of detuning has been demonstrated experimentally with the prototype 1.8-m detuned X-band section by using positron and electron bunches from the SLC damping rings as probe and witness beams, respectively [Adolphsen 1994], as discussed in Section 8.2.10.

The high-power rf source for the NLCTA is the 50-MW, X-band klystron [Wright 1994] discussed in Section 8.4.2. Thus far, four prototype tubes (XL1, XL2, XL3, and XL4) have been manufactured and operated at 50-MW peak power for the required 1.5-μs pulse duration at 60 pulses per second.

Rf pulse compression in the NLCTA will be performed by the SLED-II technique. A prototype SLED-II system has been tested with an X-band klystron at compressed-pulse power levels up to 200 MW, to validate the design of the NLCTA pulse-compression system and its components [Nantista 1993], as discussed in Section 8.3. Power from the SLED-II prototype has been used to achieve a 67-MV/m accelerating gradient in the prototype 1.8-m-long accelerator section in the Accelerator Structure Test Area (ASTA) of the Klystron Test Laboratory at SLAC [Vlieks 1993, Wang 1994].

To achieve low rf losses, oversized circular waveguide will be used for the SLED-II delay lines and for the transmission lines that carry the rf from the klystron to SLED-II, and from SLED-II to the accelerator. The TE_{01} mode will be propagated in the circular waveguide. Matching to the TE_{10} mode in rectangular waveguide will be performed by compact, low-loss-mode transducers [Tantawi 1993], as discussed in Section 8.3.

The NLCTA rf system has been designed to accommodate the possibility of a future upgrade which would increase the accelerating gradient from 50 MV/m to 85 MV/m by replacing each 50-MW klystron with a pair of 75-MW klystrons. Each NLCTA pulse modulator is capable of accommodating a pair of either 50-MW or 75-MW tubes.

The 70-MeV, X-band injector module and two of the three 130 MeV, X-band linac modules of the NLCTA are expected to become operational in 1996, with each module powered by a single 50-MW klystron. The first accelerator physics experiments are planned for 1996–1997. The last linac rf module (including the fifth and sixth 1.8-m-long X-band sections) is planned to be installed in 1997. It is expected that the upgrade from 50-MW klystrons to 75-MW klystrons will occur gradually as the new higher-power tubes become available through the klystron development program. The initial complement of 50-MW klystrons, and the first of the 75-MW klystrons, will be solenoid-focused. Later versions of both 50-MW and 75-MW klystrons are expected to be focused by cylindrical arrays of PPMs, as discussed in Section 8.4.

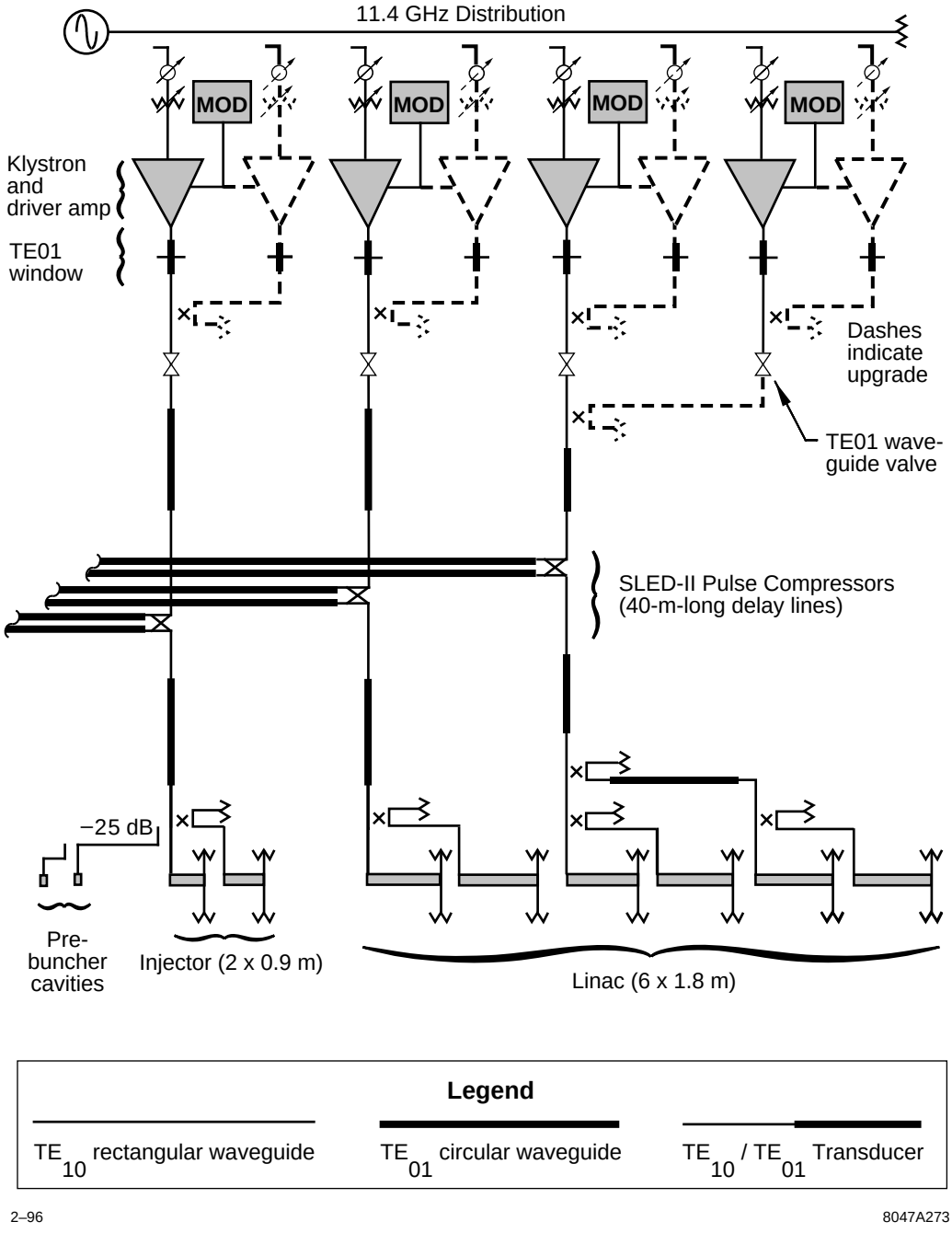


Figure 8-2. *Schematic layout of the NLC Test Accelerator's high-power rf system.*

Parameter	Design	Possible Upgrade
Beam:		
Electrons per bunch	0.4×10^9	6.5×10^9
Bunch frequency	11.424 GHz	0.714 GHz
Bunches per pulse	1440	90
Pulse length	0.125 μs	
Beam pulse repetition rate	10 Hz	
Accelerator:		
Accelerating gradient, unloaded	50 MV/m	85 MV/m
Accelerating gradient, full current	37 MV/m	64 MV/m
Filling time	0.1 μs	
Section length	1.8 m	
Sections per module	2	
Modules in linac (excluding injector)	3	
RF Pulse Compression:		
Compressed rf pulse length	0.25 μs	
Compression ratio	6	
Efficiency (intrinsic $\times$ components)	0.75 $\times$ 0.90 = 0.67	
Peak power gain	4.0	
RF Power Transmission:		
Transmission Efficiency	0.86	
Klystrons:		
Klystrons per module	1	2
Peak rf power per klystron	50 MW	75 MW
Klystron pulse length	1.5 μs	1.1 μs
Voltage	440 kV	500 kV
Perveance	1.2 μA/ V$^{3/2}$	0.75 μA/ V$^{3/2}$
Electronic efficiency	0.43	0.60
Rf pulse repetition rate	180 Hz	120 Hz

Table 8-2. *NLCTA design parameters and a possible upgrade path.*

8.1.4 Outlook

The design of the rf system for the main linacs of the NLC is supported by existing and planned developmental prototypes, and by the NLCTA. Key NLC parameters such as the klystron power, acceleration gradient, and pulse-compression power gain have been exceeded in prototype systems. The next steps in the development program are completion of the NLCTA, the first damped and detuned structure, and the first PPM klystron prototype. The design of the NLC high-power rf system is mature and is progressing toward detailed engineering considerations. Because of the magnitude of the project, special emphasis is now being placed on designing for manufacturability and for overall system reliability.

8.2 Accelerator Structure

The design of the X-band accelerator structures for the NLC is based on theoretical and experimental experience gained by numerically modeling and building accelerator structures for the NLCTA, and operating them at high gradients. One of the main challenges is to suppress the deflecting modes that will otherwise cause severe multibunch emittance growth in the NLC linacs. Suppression of the transverse wakefield will be achieved through a combination of precision alignment and by detuning and damping higher-order modes. Another challenge in the design of the accelerator structures for the NLC is suppressing field emission at the high-surface field gradients encountered in these structures. This suppression, so far, has been achieved through machining, processing, and handling techniques that minimize surface roughness and eliminate contamination of the high-gradient surfaces. Other, additional methods may be adopted later.

There is a significant amount of overlap in the discussions of the previous chapter and this present one, since both deal with the design and performance of the main linacs. The previous chapter focused on beam dynamics issues, most importantly preservation of the beam emittance and stability. This chapter is concerned with the systems needed to accelerate the beams. This section outlines an engineering design of the accelerator structures that will meet the beam dynamics requirements.

As part of the process of developing the structure design for the NLC, several 1.8-m NLC-type accelerating structures are being built for use in the NLCTA. These are of two types: some with detuning alone, and others with both detuning and damping. In both cases, the accelerator structure is a disk-loaded waveguide driven at 11.424 GHz, the phase advance per cell for the accelerating mode is chosen to be $2\pi/3$, and the detuned distribution of the synchronous lowest dipole-mode frequencies has a density in frequency space that is Gaussian with truncation at $\pm 2\sigma$ and a total detuning range of 10%. (Only damped, detuned structures will be used in the NLC main linacs.)

The frequency spread in this detuned-mode distribution results in an interference between modes that strongly attenuates the corresponding component of the wakefield that drives multibunch beam break-up. The desired Gaussian distribution of detuned modes is obtained (while also keeping the accelerating mode frequency fixed) by varying the dimensions along the structure, the main influence coming from the cell radii (*b*) and the iris radii (*a*). To reduce the smaller but non-negligible effect of the higher dipole modes, we also vary the disk thickness (t) ranging from 1 mm in the first cell to 2 mm in the last cell, in a truncated Gaussian pattern having standard deviation $\sigma_t = 0.25$ mm. (See discussion in Section 7.4.2.) Because the number of cells used to implement the Gaussian detuning pattern is finite, the wakefield resurges on a distance scale of about $c/2\Delta f \approx 30$ m, where Δf is the cell-to-cell frequency separation in the center of the distribution. To suppress this resurgence of the long-range wakefield, the damped and detuned structure (DDS) incorporates dipole-mode damping in addition to the detuning discussed above. This damping is accomplished by coupling each accelerator cell to four evacuated waveguide manifolds running parallel to the structure, symmetrically located around its circumference. The manifolds are terminated at each end by matched

loads. No manifold modes propagate at the frequency of the accelerating mode, so there can be large coupling to the dipole modes without significant damping of the accelerating mode. (We have set a limit of a few percent degradation of the shunt impedance of the accelerating mode.) Because the higher-order modes are tuned to different frequencies, one finds (Section 7.4.2) that they have a broad spectrum of phase velocities of both signs. They are therefore capable of coupling effectively to all propagating modes in the damping manifolds.

In this section, we discuss the calculation of the dimensions of the individual cells, which vary along the structure. Next, tolerances on the dimensions and alignment of the cells in the structure are discussed. Following this, the steady-state unloaded and beam-loaded gradients are calculated. In the absence of beam-loading compensation, there would be a large sag (on the order of 25%) in the beam energy during the first 100 ns (equal to the structure filling time) following beam turn-on. There are several possibilities for compensating this transient energy variation. The method chosen here is to tailor the amplitude of the rf power during the structure filling time so that the exact steady-state beam-loaded gradient (at every position along the length of the structure) is present at the moment the beam passes through. This method of multibunch energy compensation has already been introduced in Section 7.4.5, where for simplicity the approximation of a constant-gradient structure was assumed. In this chapter the scheme is extended to calculate the exact final amplitude function needed for the NLC quasi-constant-gradient structure described here.

We then dicuss in some detail the mechanical design of the structure, including vacuum and thermal calculations, high-power tests, and material handling and processing techniques.

Next, we summarize the tolerances on ripple of the phase and amplitude of the incoming rf pulse that are needed to meet given requirements on the energy and energy spread of the multibunch beam. The experimental program planned for the NLC Test Accelerator, which is designed to achieve these tolerances, is discussed.

Finally, we discuss some of the tests and diagnostics related to the detuning and damping of the accelerator structures. The Accelerator Structure SETup (ASSET) Facility at SLAC uses the SLC positron and electron bunches to probe and witness, respectively, the wakefields in accelerator structures. Excitations of the dipole modes in a structure may be used as a diagnostic to measure the alignment of the structure with respect to the beam.

8.2.1 Calculation of Structure Dimensions

The task of designing the accelerator structure includes calculating the physical dimensions of a set of cells with a common rf feed (a section) that results in a truncated Gaussian distribution for the lowest dipole-mode frequencies, while maintaining the desired frequency and phase advance per cell for the accelerating mode. The truncated Gaussian distribution has a given standard deviation σ_f and a density of frequency components near the central frequency $\bar{f}_1$ proportional to $\exp[-(f-\bar{f}_1)^2/\,2\sigma_f^2]$. This means that the spacing between adjacent modes near the ith mode is given implicitly by

$$\mathrm{erf}\left(\frac{f_{1,i}-\bar{f}_1}{\sqrt{2}\,\sigma_f}\right) = \mathrm{erf}\left(\frac{f_{1,i-1}-\bar{f}_1}{\sqrt{2}\,\sigma_f}\right) + A \quad , \tag{8.1}$$

where A is a constant, given by

$$A \equiv \frac{2\,\mathrm{erf}\,(n_\sigma/2\sqrt{2})}{N-1} \quad . \tag{8.2}$$

Here, N is the number of cells in the accelerator structure,

$$n_\sigma \equiv \Delta f_{\mathrm{tot}}/\sigma_f \tag{8.3}$$

is the full width of the truncated distribution in units of σ_f, and $\mathrm{erf}\,(x)$ is the error function:

$$\mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}\,du \quad . \tag{8.4}$$

In the central core of the distribution, the fractional spacing between adjacent frequency components is approximately

$$\frac{\delta f}{\bar{f}_1} \approx \frac{\sqrt{2\pi}}{N-1}\frac{\sigma_f}{\bar{f}_1}\,\mathrm{erf}\left(\frac{n_\sigma}{2\sqrt{2}}\right) \quad . \tag{8.5}$$

We discuss next how we arrive at structure dimensions that satisfy these requirements on the accelerating mode and higher-order modes. Increasing the cavity radius b causes both the accelerating mode frequency and the first dipole-mode frequency f_1 to decrease, while increasing the iris radius a leads to an increased accelerating mode frequency and a decreased first dipole-mode frequency. As noted earlier, we also choose to vary the disk thickness t in a specified pattern. Keeping the frequency of the accelerating mode constant (11.424 GHz) yields a unique relation between iris radius a, cell radius b, and disk thickness t. Each of these triplets (a, b, t) corresponds to a different dipole-mode frequency. If a certain detuning range of the dipole modes is given, and t is varied in the specified truncated Gaussian pattern from 1 mm to 2 mm, the triplets (a, b, t) for the two end-cells of the accelerator section can be found. It is always possible to find a unique triplet (a, b, t) to set the dipole-mode frequency to any value between the frequencies of the first and last cells and also to keep the frequency of the accelerating mode constant; we choose to vary the lowest dipole-mode frequency in a truncated Gaussian pattern, with total frequency spread of about 10% and truncation at $\pm 2\sigma_f$.

We first discuss the calculation of dimensions for the detuned structure. The overall design procedure, which uses polynomial three-parameter fits, is as follows:

1. Using the computer code YAP [Nelson 1992a], the relationship among a, b, and t, given the fixed accelerating frequency and phase advance per cell ($2\pi/3$), may be found for the structure (taking into account the effect of the rounded corners on the irises).
2. Again using YAP, the relationship among the synchronous dipole-mode frequency f_1, a, b, and t may be found (where b is fixed by step 1).
3. The desired spacings of the dipole-mode frequencies, $\{f_{1,i} - f_{1,i-1}\}$, and the distribution of disk thicknesses, $\{t_i\}$, are specified. As already noted, both of these are chosen to be truncated Gaussian distributions.
4. Given a value a_1 for the iris radius of the first cell, the central frequency $\bar{f}_1$ and all of the a_i and b_i are uniquely determined by the above constraints. We adjust a_1 to obtain the desired filling time T_f for the structure.

The resulting structure parameters are summarized in Table 8-3. When a parameter varies along the structure, the range of values from the first to the last cell is given.

This procedure must be further modified to calculate the dimensions for the DDS. The reduction of each cell diameter required to compensate for the presence of the damping manifolds is calculated using the 3-D MAFIA code. Starting from the (a, b) pairs obtained for the purely detuned structure, we calculate the increment in b that is needed to tune the fundamental mode back to the desired rf frequency of 11.424 GHz. These calculations are to be verified by microwave measurements on a series of uniform cavity stacks corresponding to different cavities in the section.

For an accelerator structure consisting of N cells, the above procedure gives the frequencies of the synchronous dipole modes for each of N periodic structures, where each such structure is constructed from cells like one of those in the actual structure. The relationship between the dipole mode frequencies in these equivalent periodic structures and the coupled-mode frequencies (and mode field distribution patterns) in the 206-cell accelerator structures is calculated using equivalent circuit models, both for the undamped detuned structure [Bane 1993a] and for the DDS [Kroll 1994]. In the case of the DDS, the damping of the structure modes due to their coupling to the manifolds is also obtained from the model. Detailed discussion of these equivalent circuit models and the resulting calculations of long-range wakefields were given in Section 7.4.2.

Accelerating mode frequency	11.424 GHz
Phase advance per cell	$2\pi/3$
Structure length	1.8 m
# of cells	205 + 1 couplers
Iris radius, a	0.572 to 0.390 cm
Cell radius, b	1.084 to 0.998 cm
Disk thickness, t	1 to 2 mm
Frequency range of dipole modes	14.312 to 15.834 GHz
Mean dipole-mode frequency, $\bar{f}_1$	15.073 GHz
$\sigma_f/\bar{f}_1$	2.5%
Total fractional spread, $\Delta f_{\rm tot}/\bar{f}_1$	10.1%
Group velocity, v_g/c	0.12 to 0.03
Filling time, T_f	100 ns
Attenuation parameter, τ	0.533
Elastance, $s \equiv \omega r/Q$	652 to 946 V/pC/m
Peak power per feed (for 50 MV/m unloaded)	89.8 MW
Q of lowest dipole mode	≈ 6500

Table 8-3. *Parameters for an NLC structure.*

8.2.2 Tolerances on Dimensions and Alignment

The tolerance on the structure dimensions, particularly on cell radius b, comes from the effect on the distribution of dipole mode frequencies, which is designed to be a truncated Gaussian. The tolerance on the frequency is roughly the core spacing of the truncated Gaussian distribution of frequencies. The tolerance on misalignments of the structures comes from the effect on the transverse emittance of the multibunch beam. This tolerance is dependent on the longitudinal correlation length of the misalignments, but it is fairly tight on all scales. (See Section 7.4.6: "Structure Misalignments".) Both these tolerances are looser for the DDS than for the undamped Gaussian detuned structure.

Tolerances on Frequency Errors

There are two extreme cases for the frequency errors: the case where the error in each frequency in the design distribution is the same in all sections (we denote this as "systematic"), and the case where the error in each frequency is totally random from section to section. Note that our definition of systematic means that the errors are the same in corresponding cells of a given structure type, but they are still random from cell to cell in each structure type. Systematic errors can lead to considerable worsening of the long-range wakefield behavior; the totally random errors are much less harmful. If care is taken to randomize the production of various cell types, it should be possible to keep the systematic components of the errors significantly smaller than the random component, perhaps by nearly an order of magnitude.

As one might expect, the transverse multibunch beam emittance growth is not much affected by the frequency errors, provided that the fractional errors in the frequency distribution are kept small compared to the core spacing.

For a single detuned accelerator section, the fractional core spacing $\delta f/\bar{f}_1$ for the fundamental dipole-mode frequencies is about 3×10^{-4}. Machining precisions for conventional machining and diamond-point machining (obtained at KEK), and alignment tolerances of stacks of cells, are given in Table 8-4. Since the cell radius is about a centimeter,

diamond point machining should produce a random $\delta f/f$ error somewhat less than the core spacing. The systematic error, as noted above, should be significantly less than this.

Misalignment Tolerances

Misalignment tolerances, on scales ranging from a few cells within a structure to several structures, were discussed in the preceding chapter. It was found that the tightest tolerances occurred for the alignment of groups of about 20–40 cells. Thus, great care must be taken in brazing together the subsections of structures after their initial assembly from individual cups.

8.2.3 Calculation of Steady-State Gradients

In this section, we discuss the calculation of the unloaded and loaded accelerating gradients. We also discuss the proposed compensation of the transient beam loading that occurs during the first filling time after the beam is injected.

Unloaded Gradient

The power flow in the accelerator structure can be expressed as

$$P(z) = P_{in} e^{-2\tau(z)} \quad , \tag{8.6}$$

where P_{in} is the input power and $\tau(z)$ is the attenuation along the structure, given by:

$$\tau(z) = \omega_{rf} \int_0^z \frac{dz'}{2Q(z')v_g(z')} \quad . \tag{8.7}$$

The shunt impedance $r(z)$ can be calculated using the code SUPERFISH. Then the accelerating electric field $E_z(z)$ can be calculated as

$$E_z(z) = \sqrt{r(z)\frac{dP(z)}{dx}} \quad , \tag{8.8}$$

where $r(z)$ is the shunt impedance per unit length. In Figure 8-3, $E_z(z)$ is shown for an input power of 100 MW. For comparison, $E_z(z)$ for a conventional constant-gradient structure and for a constant-impedance structure, with the same input power and attenuation as the detuned structure, are also plotted.

Beam-loaded Gradient

In this section, we generalize Equation 8.8 so that the discrete nature of the structure is explicitly taken into account. The analysis is then extended to include beam loading. The underlying assumption is that the structure parameters vary slowly and smoothly with length, and that the structure can be treated as locally periodic. In this case, the unloaded steady-state gradient in the nth cell is given by

$$G_0(n) = \left[2\alpha(n)r(n)P(n)\right]^{1/2} \quad , \tag{8.9}$$

where α is the attenuation parameter per unit length. It is related to the other structure parameters by

$$\alpha = \frac{\omega}{2v_g Q} = \frac{s}{2v_g r} \quad , \tag{8.10}$$

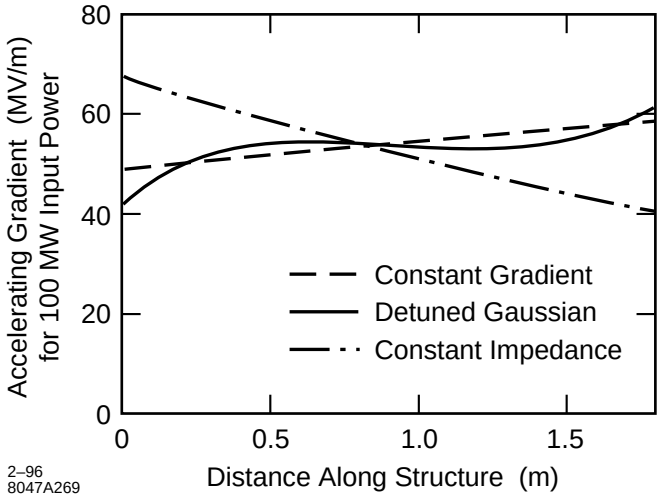


Figure 8-3. *Accelerating electric field gradient along the axis of detuned structure. Electric field gradients for conventional constant-impedance and "constant-gradient" structures with the same input power and attenuation are also shown for comparison.*

where $s = \omega(r/Q)$ is the local elastance per unit length. Using Floquet's theorem with Equation 8.9, we get

$$G_0(n+1) = \left[\frac{\alpha(n+1)r(n+1)}{\alpha(n)r(n)}\right]^{1/2} e^{-\alpha(n)d} G_0(n) \quad . \tag{8.11}$$

Here $d = L_s/N$ is the cell length, where N is the total number of cells in a structure of length L_s. The gradient in the first cell is $G_0(1) = [2\alpha(1)r(1)P_0]^{1/2}$, where P_0 is the steady-state input power. The distance, attenuation, filling time and voltage from the input to the end of the nth cell are

$$z(n) = nd \tag{8.12}$$

$$\tau(n) = d\sum_{i=1}^{n} \alpha_i \tag{8.13}$$

$$T_f(n) = d\sum_{i=1}^{n} 1/v_{gi} \tag{8.14}$$

$$V(n) = d\sum_{i=1}^{n} G_i \quad . \tag{8.15}$$

We assume that the structure is operated synchronously (no cell-to-cell phase error for a velocity-of-light electron), although the analysis can easily be extended to the nonsynchronous case.

The gradient induced by a charge Δq passing through the nth cell is $\Delta G_b(n) = s(n)\Delta q/2$. Using $\Delta q = I_0 d/v_g(n)$ and Equation 8.10, the gradient induced during the time it takes for the accelerating wave to pass through the cell is

$$\Delta G_b(n) = \alpha(n)r(n)I_0 d \quad , \tag{8.16}$$

where I_0 is the steady-state beam current. The beam-induced gradient in the first cell is $G_b(1) = \alpha(1)r(1)I_0 d$. The net beam-loading gradient in the nth cell is the sum of $\Delta G_b(n)$ and a term due to power flow from the gradient induced

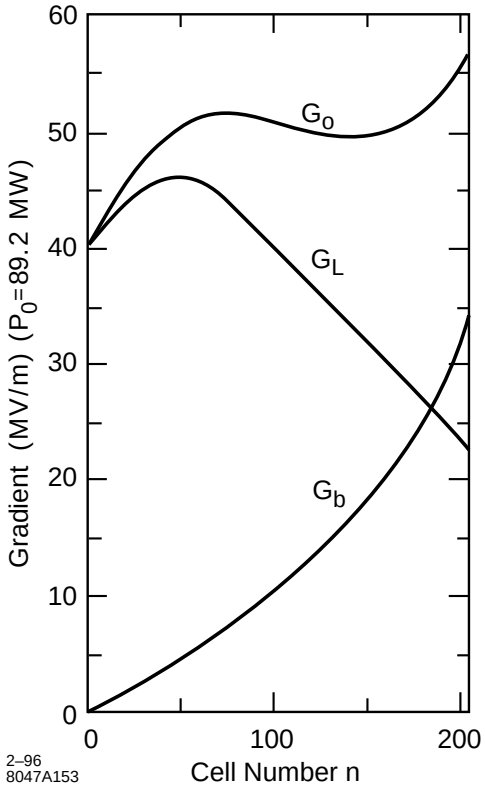


Figure 8-4. *Generator (G_0), beam (G_b), and loaded (G_L) gradients as a function of cell number n.*

in upstream cells. In analogy with Equation 8.11, the gradient in cell $n+1$ in terms of the gradient in cell n is then

$$G_b(n+1) = \left[\frac{\alpha(n+1)r(n+1)}{\alpha(n)r(n)}\right]^{1/2} e^{-\alpha(n)L_c}G(n) + \alpha(n)r(n)I_0 d \quad . \tag{8.17}$$

The net beam-loaded gradient in the nth cell is

$$G_L(n) = G_0(n) - G_b(n) \quad . \tag{8.18}$$

The functions G_0, G_b and G_L are plotted in Figure 8-4 for an average unloaded gradient $\bar{G}_0 = V_0(N)/L_s$ = 50 MV/m. It is assumed that the presence of a damping manifold results in a 3% reduction in Q. The required input power is 89.2 MW per structure (50 MW/m). The average beam-loading gradient is 14.7 MV/m for a charge per bunch of 7.5×10^9 electrons, giving a net loaded gradient of 35.3 MV/m. The beam-loading derivative is 17.13 MV/A. Further details are given in NLC Technical Notes [Farkas 1994, Farkas 1995].

Transient Beam-Loading Compensation

To compensate for transient beam loading, it is necessary to produce the steady-state beam loaded gradient profile, $G_L(n)$, by varying the generator power flow during one filling time prior to beam turn-on at time $t = 0$. This compensation is tantamount to "pre-loading" the structure. Taking into account the attenuation and propagation time

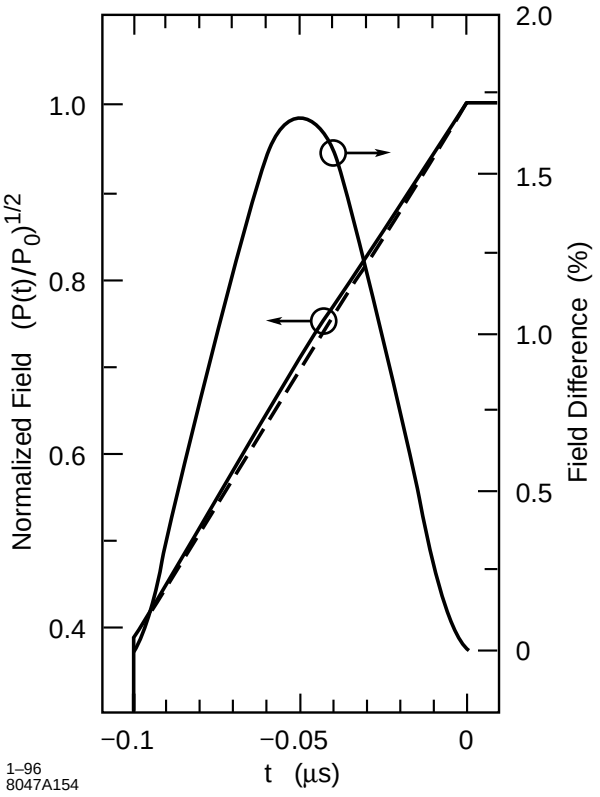


Figure 8-5. *Normalized input field $[P(t)/P_0]^{1/2}$ and the difference between it and a linearly-ramped field (dashes), vs. time. (The initial normalized input field is 0.385.)*

from the input cell to cell n, and the relation between power flow and gradient as given by Eq. 8.9, the required ramping profile is

$$P\left[-T_f(n)\right] = \frac{{G_L}^2(n)e^{2\tau(n)}}{2\alpha(n)r(n)} \quad . \tag{8.19}$$

where $T_f(n)$ and $\tau(n)$ are given by Eqs. 8.13 and 8.14.

The normalized field ramp $[P(t)/P_0]^{1/2}$ is plotted in Figure 8-5. For a true constant-gradient structure, this field ramp would be exactly linear. A linear ramp is also plotted in Figure 8-5 for comparison, along with the relative field difference on an expanded scale. The relative deviation from a linear ramp reaches about 1.5% (3% in power) at $t \approx -0.5T_f$. Use of a linear ramp instead of the exact compensation profile would result in an energy variation of about 1% along the bunch train.

8.2.4 Mechanical Design of the Accelerator Structure

This section gives a general description of the mechanical design of the damped and detuned structure (DDS). The discussion covers the design and tolerances of the cells, vacuum pumping and water cooling systems, the input coupler, and the supporting strongback. Stack assembly for diffusion bonding is described. Starting with the unavoidable geometrical complexity which is mandated by the theoretical specifications, an attempt has been made to keep the rest

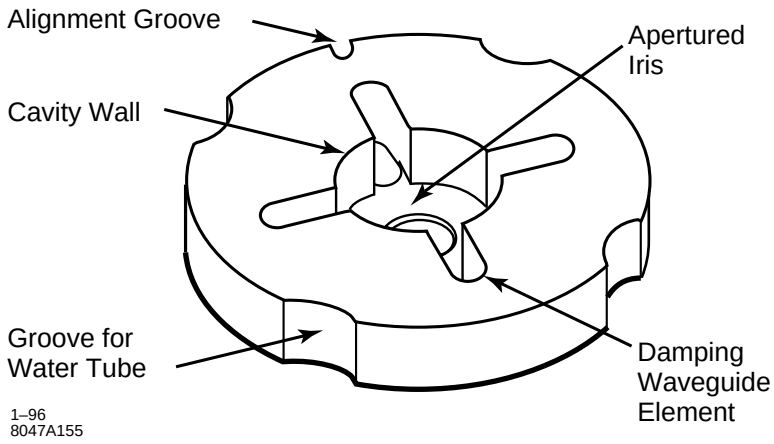


Figure 8-6. *Basic cell design.*

of the design as simple as possible to minimize costs and to facilitate automation of the production of the very large quantities of components required.

Cell Design

The basic features of the cell are shown in Figure 8-6. The central portion is the conventional "cup" consisting of a cylindrical cavity wall and an apertured iris which, when stacked in a row with other cells, forms the disk-loaded waveguide accelerator structure. The diameter of the cavity wall, the thickness of the iris, and the diameter of its aperture vary progressively from cell to cell to "detune" the beam-deflecting dipole modes and suppress short-range cumulative build-up of wakefields, while maintaining the quasi-constant-gradient characteristics of the fundamental accelerating mode. The central cavity is slot-coupled to four outer rectangular holes. When the cells are stacked, the holes form four waveguides which run parallel to the axis of the structure. When terminated in matched loads, the four waveguides become "damping manifolds." The slots (which are cut off for the fundamental accelerating mode) couple power from beam-excited dipole modes into these damping manifolds, lowering the Q and suppressing long-range wakefield build-up. Microwave signals from the damping manifolds can be used to monitor the alignment of the structure with respect to the beam.

The waveguide height and its distance from the structure axis vary progressively from cell to cell, as described in Section 8.2.1. Thus, for example, cells 10, 70, 122 and 196 appear as shown in Figure 8-7. All cells have the same thickness.

The cells have four circular-sector grooves equispaced around their periphery to accommodate four water-cooling tubes which are attached in a final brazing operation. The partially inset tubes provide adequate thermal conductance while avoiding direct water-to-vacuum diffusion bond interfaces. Since there are 205 cells in each accelerator section, forming the water channels by lining up four holes drilled through each cell would have resulted in over 800 such bonds in each section, greatly increasing the probability of failure due to leakage.

Each cell has an additional groove which is used to determine rotational alignment during assembly.

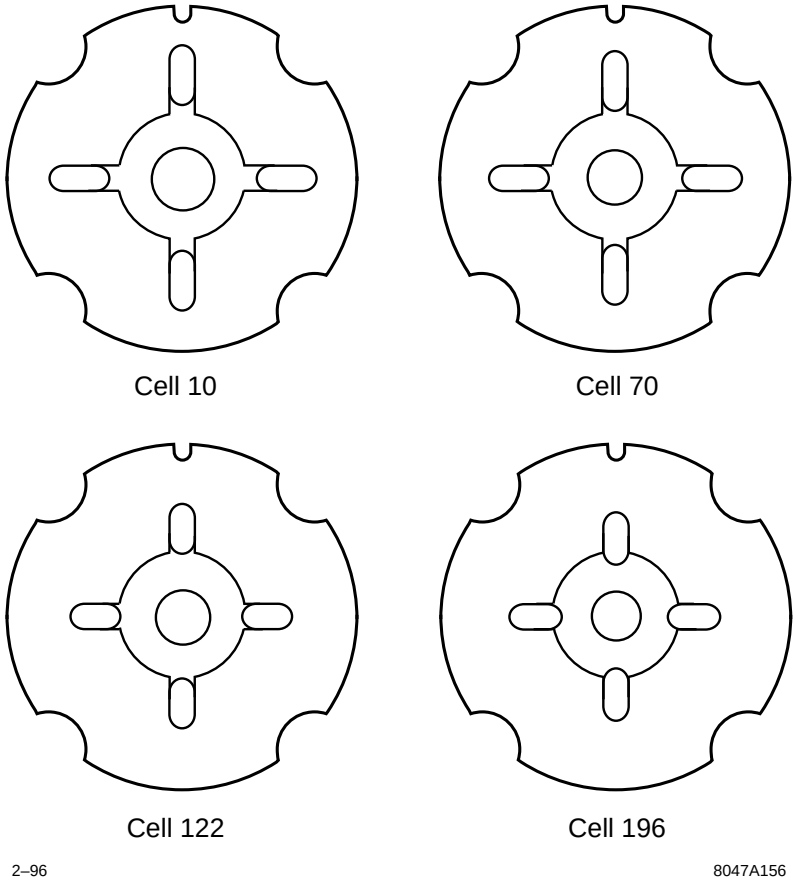


Figure 8-7. *Front view of typical cells.*

Cell Tolerances

The goal is an accelerator section 205 cells long (plus the input coupler) in which the axes of the individual cavities lie on a straight line to within 15 μm (rms).

To avoid a cumulative tolerance build-up which exceeds this goal, many other cell dimensions have to be held to tolerances in the 2 to 3 μm range. This applies particularly to the diameter and coaxiality of the outer surface of the cell, since this has to be the reference surface used in stacking the cells prior to bonding. It also applies to the perpendicularity, flatness and parallelism of the cell faces. Errors here will cause "bookshelving" [Seeman 1985] in the cell stack. Equally important is the coaxiality of the iris aperture with respect to the cavity outer wall.

These tight tolerances on cavity dimensions (including the thickness of the iris and the radius of its edge) are also necessary to achieve the design fundamental and dipole-mode frequency characteristics, because provision for tuning after assembly has been eliminated.

Flatness of cell faces is also essential to achieve diffusion bonding (discussed below) over relatively large surface areas with a very low expected failure rate.

	Detuned 1.8-m Section #1	Detuned 0.9-m Sections #1 & 2	Detuned 1.8-m Section #2 and DDS #1
Machining Technique	Conventional	Conventional	Diamond-point
Alignment Technique	Nesting	Vee-Block	Vee-Block
Diameters	± 7	± 7	± 2
Concentricity	10	10	1
Thickness	± 7	± 7	± 2
Parallelism	10	10	0.5
Flatness	10	10	0.5
Surface Finish	0.4	0.4	0.05
Cell-to-cell Alignment of Outer Cylindrical Surfaces:			
(a) Expected	22	10	3
(b) Measured after diffusion bonding or brazing	10	6	4

Table 8-4. *Cell machining and assembly tolerances achieved (in microns).*

Single-crystal diamond-point machining is to be used on all surfaces of each cavity, the cell faces and outer periphery. The hardness, high thermal conductivity, and low thermal expansion of diamond result in a superior tool which, when used in a vibration-free lathe, yield mirror-like surfaces on copper which have a roughness of 0.1 μm or less. This finish is necessary to obtain good diffusion bonding, high Q factors, low dark current and high power-handling capability. The sharpness, mechanical stability, and minimal wear of the tool allow the dimensional tolerances to be met.

Surface finish and dimensional tolerances are less stringent for the side-coupled holes which form the damping waveguides. Respectively, 0.5 μm and ± 20 μm are sufficient. The surface finish in the grooves for the water tubes can be 1 μm.

Table 8-4 illustrates what tolerances on dimensions and alignment have been achieved to date at SLAC, and what can reasonably be expected with the best technology available today. It can be seen that the most significant advances have been made by resorting to diamond-point machining and by using precision granite vee-blocks to align the stacks of cells prior to brazing or diffusion bonding. Cells designed to nest into each other have to fit loosely enough to permit assembly without galling, and this can result in unacceptable misalignment.

Our present experience is that stacks of 38 cells can have a bow of a few tens of microns after diffusion bonding. This can be reduced to a few microns by setting the stack horizontally in cradles on a granite block and applying bending forces while monitoring movement with a precision coordinate measuring machine.

Vacuum Pumping

In addition to direct pumping along the beam-interaction region, the structure is pumped in parallel by the four damping waveguides. Assuming a residual outgassing rate of 10^{-12} Torr-ℓ/cm^2/s for the copper surfaces, the combined conductance of this system is high enough to reduce the pressure to approximately 10^{-8} Torr in the middle of the structure when pumps are connected only at each end. Since lower pressure may be required, two groups of four special pumping cells will be inserted into the structure, at points one-fourth of the section length from each end. Each pumping cell has high-conductance slots cut in the middle of the broad walls of the damping waveguides. These slots, which are cut off to the fundamental accelerating mode, connect to four ducts and two pumping manifolds and pumps,

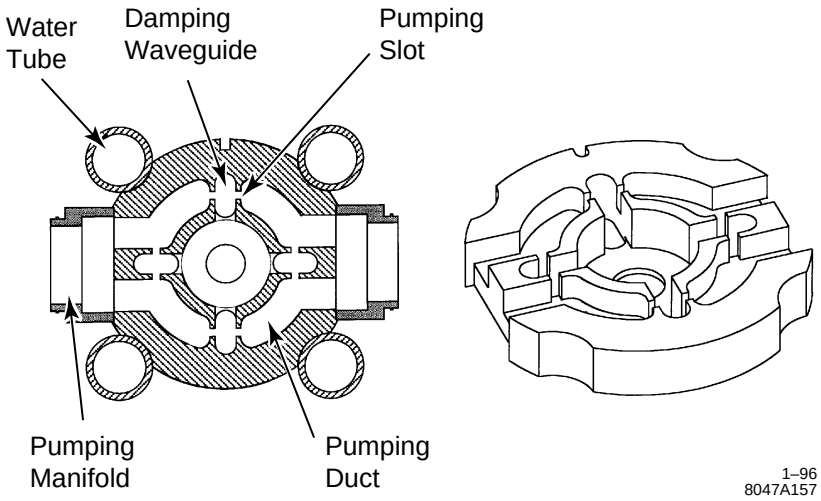


Figure 8-8. *Cross section and isometric view of pumping cell.*

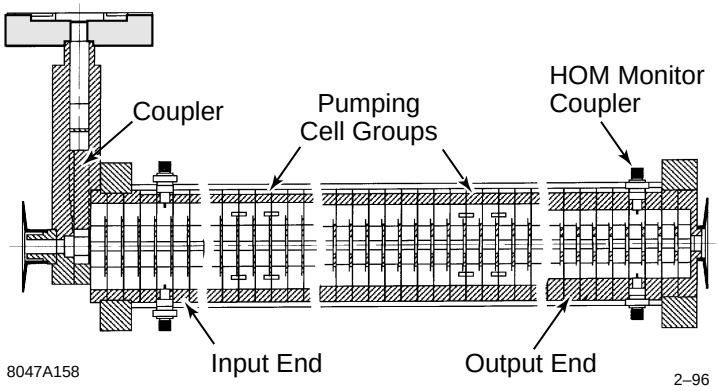


Figure 8-9. *Longitudinal cross section of accelerator section.*

as shown in Figure 8-8. By pumping only at the two quarter-points, the maximum pressure will be four times less than when pumping only at the end points (and 20 times less than when pumping from the ends of the beam-interaction region in the absence of any manifolds).

A longitudinal cross section of the accelerator section is shown in Figure 8-9. Each pumping cell group is connected to two pumps via short transverse manifolds. Each of these contains two short bellows sections to allow for a small range of accelerator movement in the vertical plane with the two pumps rigidly attached to the strongback (accelerator support beam). The horizontally opposed system ensures that there is no net transverse force on the accelerator due to atmospheric pressure.

The Input Coupler

Some details of the input coupler are shown in Figure 8-10. The coupler is a compact version of the symmetrical coupler used on the first 1.8-m sections for the NLCTA. The input cavity has double symmetrical matching irises in

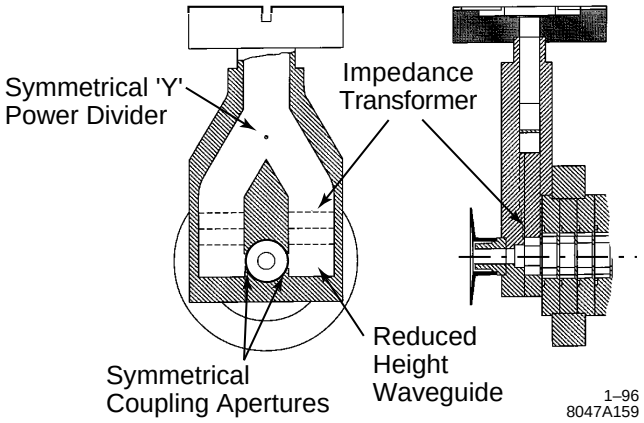


Figure 8-10. *Symmetrical input coupler.*

the horizontal plane which couple into the *sides* of two WR-90 low-impedance waveguides (instead of the ends, as used in the NLCTA design).

The waveguides are matched into standard WR-90 by step transformers, after which they combine into a common feed guide by means of a symmetrical post-matched "Y" power divider. This compact design can be machined in two blocks of copper and brazed together, ensuring electrical symmetry and minimum insertion loss.

Figure 8-9 shows no output coupler. The intention is to provide internal matched terminations for both the fundamental accelerating mode and the wakefield modes by depositing lossy coatings on the walls of the cavities and the damping manifolds in the last four cells. A coaxial loop loosely coupled to the last cavity will be used to monitor the fundamental-mode power. This signal will pass through a coaxial rf window in the cell wall, and will be used by the klystron phasing system. However, if satisfactory internal terminations cannot be developed, external loads and couplers will have to be used.

Water Cooling

After the stack of cells has been assembled by diffusion bonding, four water-cooling tubes are brazed into the grooves which are shown in Figures 8-6, 8-7, and 8-8. At each end of the section, the tubes connect to circular manifolds which provide additional conduction cooling of the input coupler and the lossy termination cavities at the output end. The direction of flow alternates from tube to tube.

Strongback Design

The standard strongback supports two 1.8-m accelerator sections. It is a 15-cm $\times$ 10-cm, aluminum box beam with a 6-mm wall. The input coupler of the first accelerator section is rigidly attached to the strongback. All other supports are flexible. Four supports are situated under the pumping cell regions of the two sections. In addition, there are supports under the middle of each section and under the input coupler and the output end of the second section. The latter is fixed in height, but all other flexible supports have adjustment mechanisms which can set the height to a few microns. These are intended only to compensate for residual elastic bowing of the accelerator-and-strongback assembly due to gravity. Any straightening of the sections which requires bending them beyond their elastic limit must be done before attaching them to the strongback.

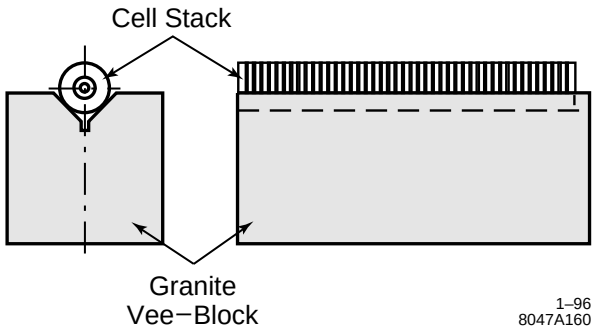


Figure 8-11. *Schematic of cell assembly on vee-block.*

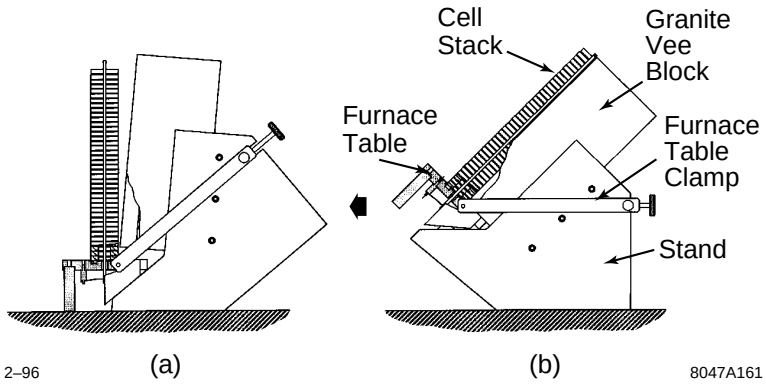


Figure 8-12. *Cells stacked on table in vee-block.*

The flexible supports are, of course, necessary to allow for differential expansion between accelerator and strongback. As presently designed, the bending radius of the flexible arm is approximately 48 mm. Thus, if the temperature of the two accelerator sections is allowed to rise 20°C above the strongback temperature, the free end will expand about 1.5 mm, which results in the accelerator axis drooping by 23 μm. An adequate solution to this problem is to maintain the strongbacks at close to the same temperature as the accelerator sections by using the same water supply in series to control their temperatures. The temperature match between accelerator and strongback does not need to be very close. Even a 5°C average difference results in a drop at the free-end of only 1 μm.

For a discussion of distortions due to transverse thermal gradients across the strongback (girder) itself, see Section 7.9.1.

Assembly: Diffusion Bonding and Brazing

The cells are presently aligned in stacks by placing them in a precision granite vee-block, as shown schematically in Figure 8-11. Actually, the cells rest on a table (Figure 8-12) which is held against the end of the block. A tension frame presses the cells against the table, so that they retain their precise alignment as the stack and table are moved away from the vee-block and transferred to the furnace.

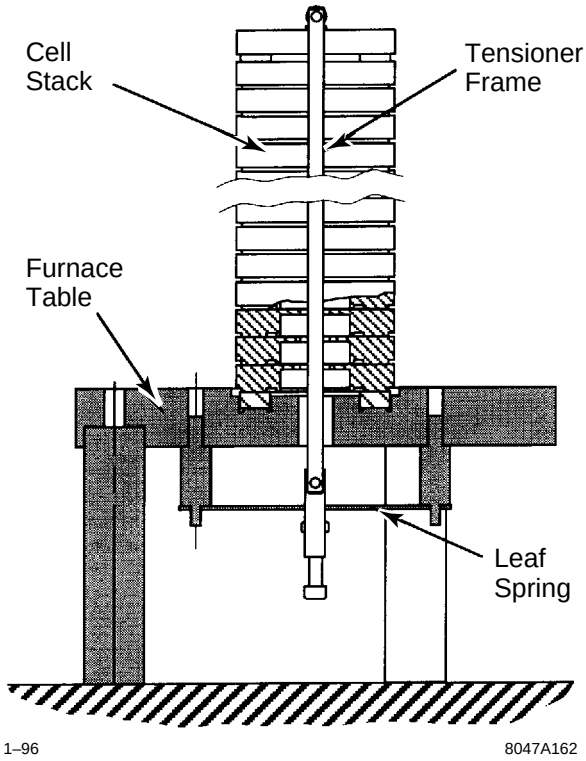


Figure 8-13. *Cells stacked for diffusion bonding.*

The cells, pressed together by the tension frame as shown in Figure 8-13, are diffusion-bonded together by heating to 1000°C in hydrogen. The number of cells bonded into a stack at one time is to be determined. Long stacks reduce the number of assembly steps but increase the difficulty of handling.

The coupler, drift tubes, and vacuum and water manifolds are added in a final furnace cycle, using a copper/gold braze. Coupling tubes are also added in two special manifold cells (one near each end of the section), into which coaxial feedthroughs (Figure 8-14) can be welded for monitoring dipole mode-induced fields in the lossy waveguides.

Leak Checking and Straightening

After the final braze, the coaxial feedthroughs are welded in. The completed section is leak-checked and then attached to a special straightening strongback. Each assembly, at present, is then transferred to a coordinate measuring machine, where the alignment of every cell is checked (with the section still hanging vertically to eliminate bending due to gravity). Corrective bending is done at intervals along the section by applying transverse forces from adjustable micrometer jacks built into the special strongback. Two sets of jacks allow bending corrections to be made in two orthogonal planes through the axis of the section.

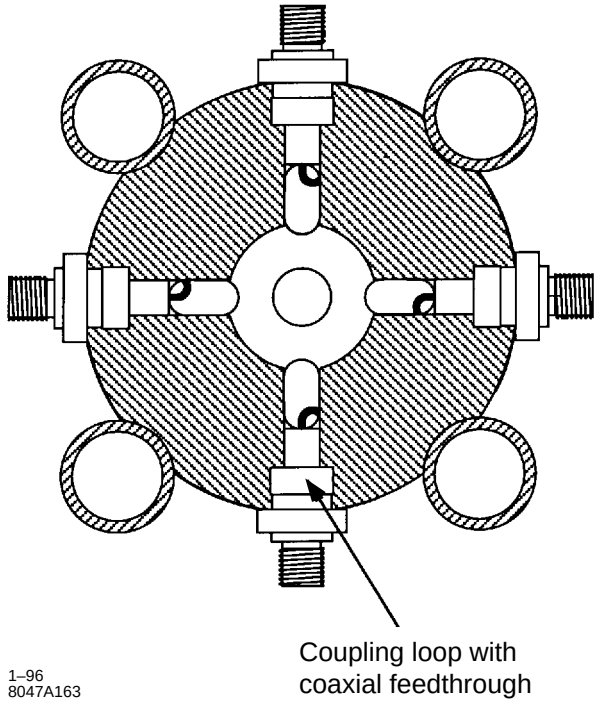


Figure 8-14. *Wakefield monitoring cell.*

Cold Testing

After straightening, the section is moved to a cold-testing station where network analyzer measurements are made to determine the input match and the phase advance per cell of the fundamental mode. The measurements are made and recorded rapidly and largely automatically (possibly by dropping a reflecting bead along the beam axis from output to input). A second set of measurements is made in the frequency band of the dipole modes, using the coaxial feeds to check the transmission characteristics from one damping waveguide across the cavities to the opposite damping waveguide. All these tests result in a simple "accept" or "reject" decision for the section, as no provision for tuning or other adjustment is made.

Bakeout

Still hanging vertically, the section is removed from the special strongback and transferred to a bake station. The section is evacuated and baked at about 500°C (in a vacuum oven) until the internal vacuum stabilizes in the low 10^{-8}-Torr region. After cooling, the section is back-filled with filtered, purified, nitrogen gas, and attached, with a second section, to a standard double-length strongback on which it will be installed in the NLC machine. The double-section assembly is shown in Figure 8-15.

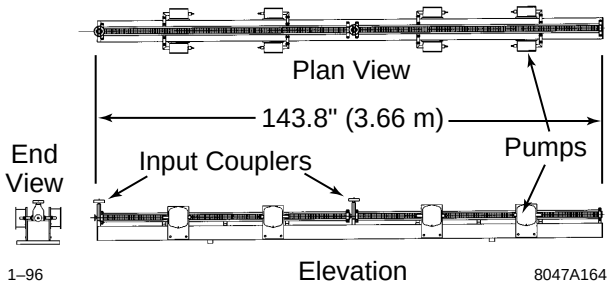


Figure 8-15. *Two NLC accelerator sections mounted on a strongback.*

Large Quantity Production of Accelerator Cells

The two main linacs will contain about 1.5 million cells, which have to be produced in about three years. Assuming that a specialized plant set up to fabricate the cells operates 50 weeks per year at two shifts per day, then, at 80% efficiency, a production rate of about 2.6 cells per minute is required. A close estimate of the time taken to machine a single cell cannot be made without reference to specific machines and machining techniques. (See, for example, the study reports prepared for the CERN CLIC cells [Read 1993] [OWGR 1993].) However, very rough estimates may be drawn from those reports, noting that the surface area of the X-band DDS cell is about three times that of the CLIC cell. All the features of the DDS cell shown in Figure 8-6 could be machined to conventional tolerances on the end of bar stock (leaving excess material on the critical surfaces to be diamond-point machined in subsequent operations), using a multipurpose CNC turning and milling machine. The cell could thus be cut from the bar using a parting-off tool, stress-relieved and delivered to a diamond-point lathe for finishing in two steps. In the first step, the cell could be gripped on the outside by a precision chuck while the back face is diamond turned. In the second step, the back face could be held in a vacuum chuck while the front face and cavity surfaces are diamond finished. Estimated operation times are: 1.5 min for conventional turning, 7 min for milling, 1 min for first-cut diamond-point turning, and 2 min for second-cut diamond-point turning. These machining times indicate that the plant will need a minimum of 22 conventional CNC turning and milling machines and 8 diamond-point lathes, with appropriate numbers of all supporting equipment (automatic handling and transfer machines, stress-relieving ovens, cleaning, QC, and packing stations) to maintain the required throughput.

8.2.5 Thermal Calculations

The thermal expansion of accelerator sections under high gradient with or without beam loading must be considered. A finite element analysis code is used to calculate the structure distortions. To maintain stabilized beam phase, a control loop is needed for adjustment of cooling water temperature.

Thermal expansion of accelerator sections under high gradient induces phase shifts with respect to the beam, and therefore detracts from beam energy stability. The heat dissipation in accelerator sections varies with the degree of beam loading and must be automatically compensated by controlling the inlet low conductivity water (LCW) temperature.

The thermal effects for the NLC accelerator sections have been calculated under the following conditions. Onto each 1.8-m-long accelerator section are brazed four cooling pipes of $0.325''$ inner radius, which are made of oxygen-free high-conductivity (OFHC) copper, aligned along the length of the structure, and spaced azimuthally 90° apart. The pipes are connected in a counterflow pattern such that one pair of 180°-separated pipes carries water from one end of

	0.5 TeV		1.0 TeV	
	Peak power/section 90 MW Total pulse length 240 ns Repetition rate 180 Hz		Peak power/section 261 MW Total pulse length 240 ns Repetition rate 120 Hz	
	No Beam	Peak Beam Curr. 0.86 A	No Beam	Peak Beam Curr. 1.26 A
Avg. dissipated power/section (kW)	2.88	2.17	5.56	4.39
Avg. temp. rise of structure ($^\circ$C)	2.03	1.55	3.93	3.10
Integrated phase shift per section ($^\circ$)	14.2	10.9	27.5	21.7

Table 8-5. *Thermal effects in 1.8-m-long X-band accelerator sections.*

the structure and the other pair carries water from the other end of the structure. The total flow rate is 16 gallons per minute. The average copper temperature is 45°.

The cooling water for the accelerator structure is provided from a manifold which supplies water to a large number of accelerator sections. The manifold will be supplied with water regulated to better than 0.1°C using a three-way mixing valve. Water flow to individual sections will be controlled using either manually adjustable control valves with optical flow meters, or flow-limiting orifices. This level of thermal control has been maintained, in the absence of varying rf power levels, for the S-band accelerator sections of the SLC at SLAC. For the NLC, the average rf power dissipated by the accelerator structure will be kept constant by combining, in phase quadrature, the power from pairs of klystrons.

In the calculation of rf power dissipation, the rf pulse in the first 100 ns is modulated to create a quasi-linearly ramped accelerating field profile at the input end of the accelerator section. The heating contribution of rf power due to the intrinsic efficiency of the SLED-II is also included. The average temperature increase of the copper structure is considered to be the average bulk water temperature change plus average film temperature drop. Table 8-5 shows the average temperature increase and corresponding integrated phase change per section for the 0.5-TeV and 1-TeV center-of-mass cases, with and without beam loading.

A finite element analysis code was used to calculate the structure distortions. The maximum asymmetry in distortion at the ends of an accelerator section is negligible (less than 0.4 μm in cavity diameter). This will not introduce excitation of higher-order dipole modes.

8.2.6 High-power Tests and Dark Current Studies

The theoretical and experimental studies on the behavior of copper accelerator structures under extremely high-rf fields have been carried out at SLAC for several years [Loew 1988a, Wang 1994a]. We have studied in considerable detail the problems of rf breakdown and dark current generated by field emission at high gradient. The dark current

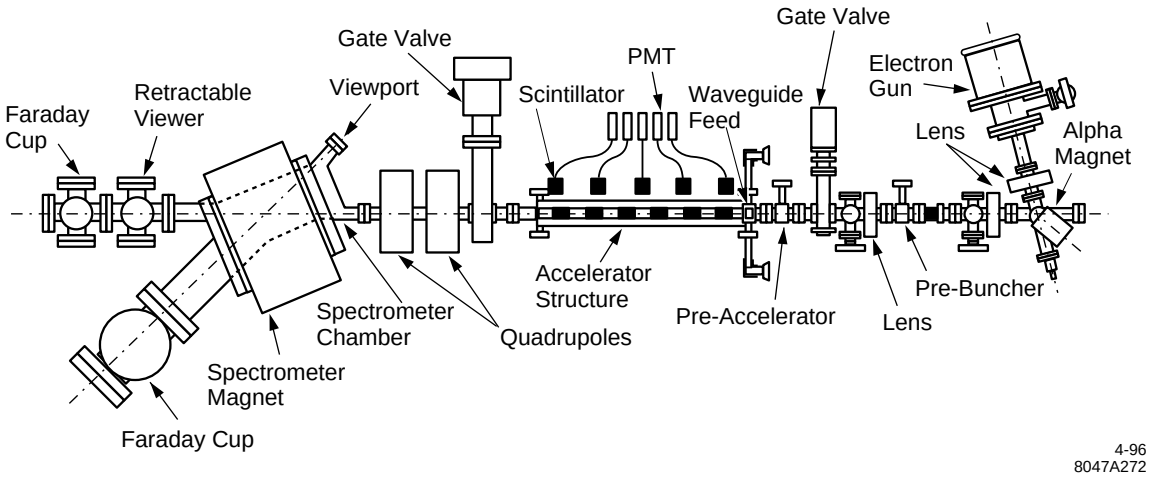


Figure 8-16. *Schematic of Accelerator Structure Test Area (ASTA).*

may absorb rf energy, get captured, and produce undesirable steering effects and detrimental x-ray radiation. Many experiments have been done to measure the amplitude and energy spectrum of the dark current and to study phenomena related to rf breakdown such as outgassing, radiation, heating, etc. We have concluded that the dark current can be minimized by improving surface finish and cleanliness, and by rf processing.

The structures which have been examined at X-band are a 6-cm standing-wave (SW) section and a 26-cm constant-impedance traveling-wave (TW) section, a 75-cm constant-impedance TW section, and a 1.8-m TW section. All of the structures were tested in the Accelerator Structure Test Area (ASTA) facility located in a shielded area. As shown in Figure 8-16, ASTA consists of a gun and beam focusing system, a prebuncher, a preaccelerator, a precision rail for mounting the accelerator sections to be tested, and a 45° spectrometer capable of analyzing electron beam energies up to 200 MeV with a resolution of $\pm 0.5\%$. Four directional couplers at input and output ports of the accelerator sections are used to pick up forward and reflected rf signals. Five collimated scintillators alongside the accelerator sections monitor localized radiation and breakdown. The accelerator structures are powered by a SLAC X-band klystron followed by a SLED-II pulse compression system.

The characteristics, main rf parameters of those structures, and high-power test results are listed in Table 8-6. The high-power tests with an electron beam have demonstrated that the 1.8-m section is properly tuned to 11.424 GHz and can accelerate a beam at a gradient of more than 67 MV/m.

Figure 8-17 shows the dark current measured in a Faraday cup as a function of average accelerating field for the 75-m constant-impedance structure and the 1.8-m detuned structure. The dark current for an accelerating gradient of 50 MV/m was found to be negligible. At 85 MV/m, it was estimated to be tolerable. The energy spectrum of the dark current was measured and found to be sufficiently low that the quadrupoles will overfocus and sweep the field-emission beam away from the primary, high-energy beam.

8.2.7 Material Handling and Processing Techniques

Experiments performed with S-band accelerators at KEK [Takeda 1991, Matsumoto 1991] have shown that the dark current produced by field emission from the accelerator disks is reduced by an order of magnitude when stringent precautions are taken to exclude dust during fabrication, assembly, processing and testing. These experiments indicate

Parameter	6 cm (SW)	26 cm (TW)	75 cm (TW)	1.8 m (TW)
Frequency, f (MHz)	11424	11424	11424	11424
Structure type	C.I.	C.I.	C.I.	Detuned
Iris diameter, $2a$ (cm)	0.75	0.75	0.856	1.134–0.786
Cavity diameter, $2b$ (cm)	2.12	2.12	2.158	2.284–2.139
Disk thickness, t (cm)	0.146	0.146	0.146	0.1–0.2
Total # of cavities	7	30	86	206
Shunt impedance (MΩ/m)	98/2	98	88	67–88
Figure of merit, Q	6960	6960	7005	7020
Coupler type	Single	Double	Double	Double
Group velocity, v_g/c	–	0.033	0.048	0.118–0.03
Filling time, T_f (ns)	80	26.5	52	100
Attenuation, τ	–	0.136	0.267	0.505
Rf pulse length (ns)	770	60	75	150
Peak input power (MW)	13.8	116	130	160
Maximum $E_{\rm acc}$ (MV/m)	–	108[b]	90 [b]	70[b]
Average $E_{\rm acc}$ (MV/ m)	110[a]	101[b]	79[b]	67[b]
Maximum surface field, E_s (MV/m)	500[a]	235[b]	206[b]	180[b]
Ratio $E_s/E_{\rm acc}$	4.55	2.275	2.292	2.37[c]

[a] Limited by rf breakdown

[b] Limited by klystron output power

[c] Ratio $E_s/\bar{E}_{\rm acc}$

Table 8-6. *Characteristics of tested structures*

that the NLC accelerator cells should be cleaned in ultra-pure, dust-free chemical solutions and rinses, assembled and bonded under clean-room conditions, and perhaps given a final rinse in ultra-pure water before vacuum bakeout and installation.

8.2.8 Multibunch Energy Spread and Compensation

Multibunch energy spread and the scheme for multibunch energy compensation were discussed in Section 7.4.5 and in Section 8.2.3: "Transient Beam-Loading Compensation". As noted in those sections, transient beam loading for both regular accelerator sections (quasi-constant gradient structures) and injector sections (over-compensated quasi-constant gradient structures) can be compensated by using a specially-modulated rf pulse, so that the rf electrical field envelope at the input end is ramped quasi-linearly during one filling time before the injection of a bunch train.

Here we summarize the tolerances (on rf phase and amplitude ripple) that come from the effect of the ripple on multi-bunch energy and energy spread. As noted in Chapter 7, the effect of dispersion of different frequency components in the rf pulse needs to be taken into account, and we have treated each bunch as a single macroparticle (*i.e.*, intrabunch energy spread is ignored in these calculations). Also, the calculations presented here do not take into account the averaging of the ripple over different rf sections, which loosens the tolerances by approximately $\sqrt{N}$ where N is the number of sections over which we average. (We expect N to be approximately the number of rf sections between linac diagnostic stations, unless there are effects that cause the ripple to be correlated over different stations.)

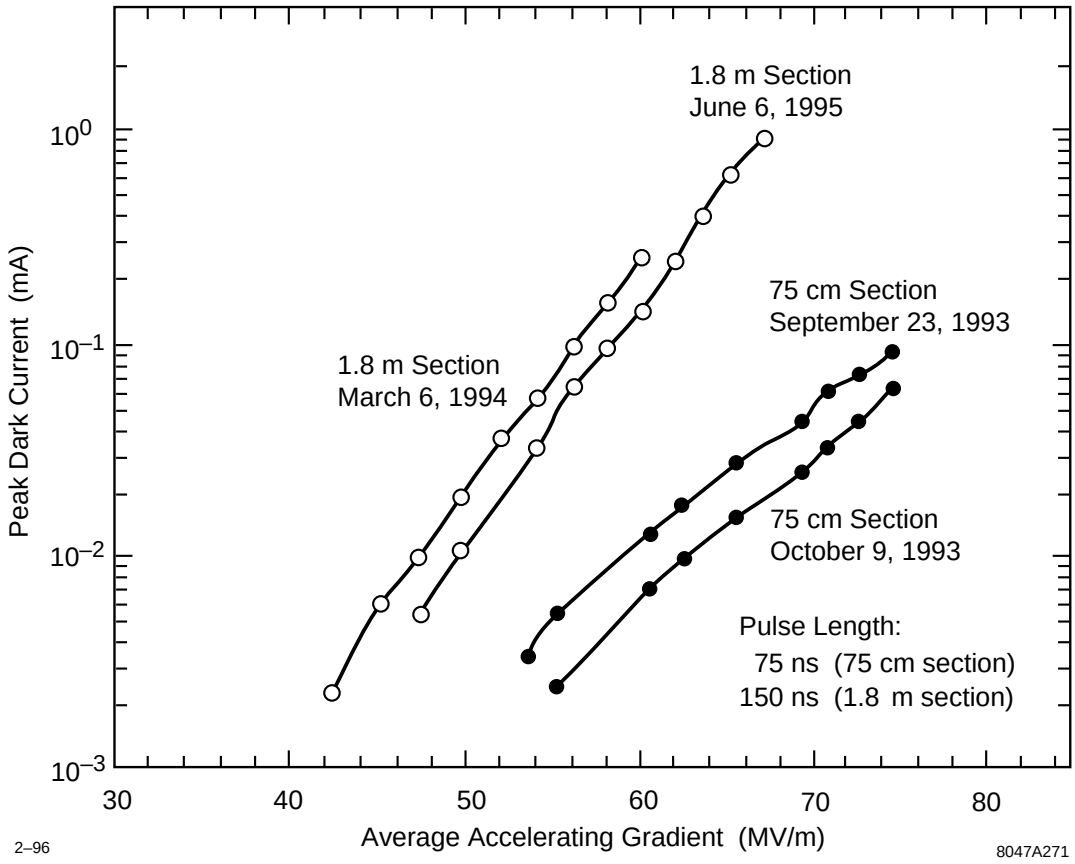


Figure 8-17. *Dark current as a function of average accelerating field for the 75-cm constant-impedance structure and 1.8-m detuned structure.*

RF Phase Variations

We considered the effect of rf phase ripple superimposed on the incoming 11.424-GHz rf pulse in the preceding chapter, where we assumed that the ripple is a cosine wave with a given period (in nanoseconds) and amplitude (in degrees at the rf frequency). For convenience, we summarize here the resulting tolerance on phase ripple as a function of ripple time scale, in order to meet given energy and energy spread tolerances on the beam pulse, consisting of a train of 90 point-like bunches. Here the tolerances on the multibunch energy spread and on the variation in the average energy of the beam have both been taken to be about $\Delta E/E \sim 10^{-3}$.

We take the tolerance on the rms fractional bunch-to-bunch energy spread to be 0.1%. As noted above, we ignore the contribution from the intrabunch spread, which would add about another 0.1% to the spread. We also take the tolerance on the energy centroid shift to be $\pm 0.1\%$. Figure 8-18 (a) shows the approximate tolerance on the rf phase ripple as a function of the period of the ripple, to meet the tolerance on the rms energy spread. Figure 8-18 (b) shows the approximate tolerance on the rf phase ripple as a function of the period of the ripple, to meet the tolerance on the centroid energy shift. The tolerances in the middle of the range of ripple timescales are approximate since they are somewhat dependent on the phase of the ripple with respect to the beam; we have recorded the most conservative values obtained in simulations.

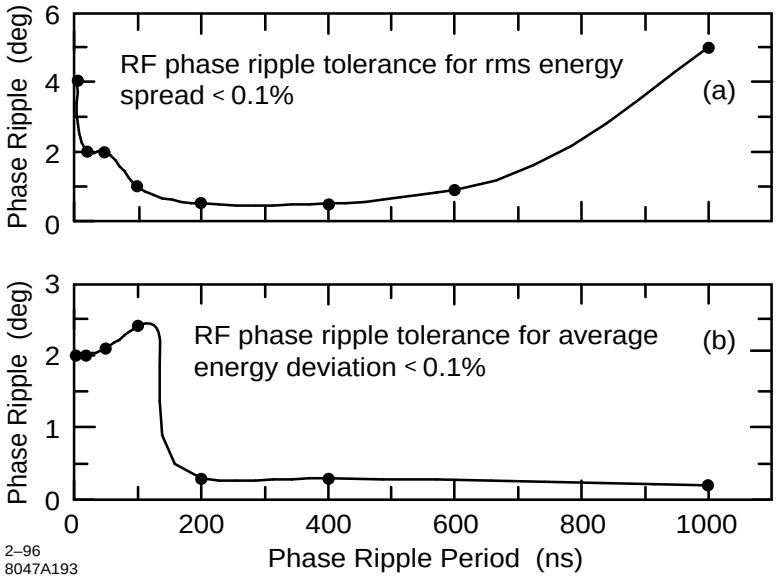


Figure 8-18. *(a) Tolerance on the rf phase ripple as a function of the period of the ripple, to meet a tolerance of 0.1% on the rms energy spread. (b) Tolerance on the rf phase ripple as a function of the period of the ripple, to meet a tolerance of 0.1% on the centroid energy shift.*

RF Field-Amplitude Variations

In Chapter 7 we also considered the effect of ripple in the amplitude of the incoming rf pulse, supposing that the ripple is a cosine wave with given period (in nanoseconds) and amplitude (in megavolts).

Figure 8-19 (a) shows the approximate tolerance on the rf field-amplitude ripple as a function of the period of the ripple, to meet the 0.1% tolerance on the rms energy spread. Figure 8-19 (b) shows the approximate tolerance on the rf field-amplitude ripple as a function of the period of the ripple, to meet the $\pm 0.1\%$ tolerance on the centroid energy shift. Again these are rough estimates, especially in the middle of the range of ripple time scales, since they are somewhat dependent on the phase of the ripple with respect to the beam.

8.2.9 NLC Test Accelerator Experiments

The goal of the NLCTA experimental program is to measure the performance characteristics of the multi-section X-band linac and high-power rf systems. The experiments will go beyond the work with single power sources and structures at the Accelerator Structure Test Area (ASTA) in the Klystron Test Laboratory; the NLCTA beam will be much more intense, will be much better defined, and will be more precisely analyzed, both before and after acceleration. In addition to system operations tests, beam-dynamics studies will be done using three types of bunch trains: a short bunch train of a few nanoseconds will be used to measure the unloaded energy-gain characteristics of the linac, a variable-length bunch train will be used to measure field-emission currents in the linac with and without injected beam, and the nominal 125-ns-long bunch train will be used to test multibunch beam-loading compensation. The specifics of these experiments are discussed in the following sections, concluded by a discussion of other types of wakefield studies that will be made possible by upgrading the NLCTA injector. An important part of the NLCTA experimental program, of great relevance for the NLC, is the commissioning of the multibunch beam instrumentation.

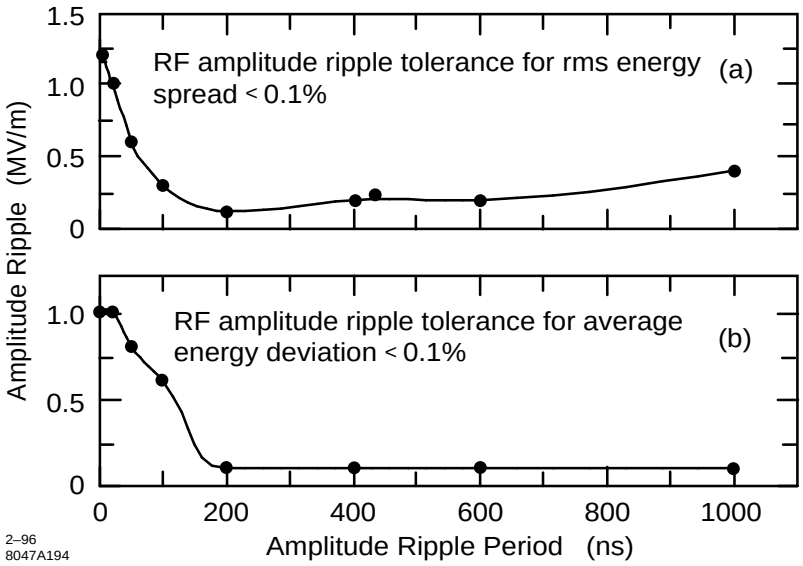


Figure 8-19. *(a) Tolerance on the rf field-amplitude ripple as a function of the period of the ripple, to meet tolerance of 0.1% on the rms energy spread. (b) Tolerance on the rf field-amplitude ripple as a function of the period of the ripple, to meet tolerance of 0.1% on the centroid energy shift.*

System Studies

The NLCTA rf system was designed to be large enough in scale, with four klystrons powering eight structures through SLED-II rf pulse compressors, that it will serve to test system control and monitoring methods applicable to NLC-scale linacs. A key component of the monitoring system will be instrumentation that can detect and digitize the phase and amplitude envelopes of the 250-ns-long rf pulses. This waveform system, which will be based on commercially available high-speed waveform digitizers, is intended to resolve 0.1% amplitude and 0.1° phase variations on time-scales longer than a few nanoseconds. The data acquisition will be controlled by local microprocessors, one of which will be assigned to each rf station, in order to permit parallel, high-level processing of the data.

We plan to monitor the transmitted and reflected rf pulses at many points along the route between each klystron and its associated accelerator structures. There will be two waveform digitizers per rf station, into which the rf phase and amplitude signals from the monitor points can be multiplexed. In this way, correlations between any two signals can be examined. An important function of the monitoring system will be to characterize the stability of the rf pulses. For this purpose, the signals will be analyzed in two basic ways. To characterize the short-term stability of some aspect of the waveform, a few hundred consecutive pulses will be sampled and analyzed. To monitor long-term stability, a sample will be taken every few minutes and stored in history buffers.

Unloaded Energy Gain

To measure the unloaded energy gain characteristics of the linac, only one klystron will be run at a time. A short bunch train, which makes the beam loading negligible, will be used to simplify the measurement. The energy gain of the bunch train depends on the rf waveform entering the two energized structures, on the rf transmission properties of the structures, and on any field emission (or breakdown) that occurs in the structures. The rf transmission properties of the structures have been computed so that, from a measurement of the input rf waveform, one can predict how the ideal energy gain should vary as a function of the transit time of the bunch train relative to the rf pulse. By comparing

this calculation to measurements done at low rf-power levels, we will test our theoretical understanding of the rf transmission properties of the accelerator sections. By comparing low- and high-power measurements, we will look for effects from possible field emission within the structures. Checks for field emission will also be made by observing the energy-gain stability at a fixed transit time for different rf power levels.

Linac Field-Emission Current

In addition to over-loading individual structures, there is a concern that field emission currents will be accelerated sufficiently to be captured by the accelerating wave and transported in an NLC-like linac. To investigate this possibility, and to provide more data on field-emission activity in the structures, the NLCTA linac will be instrumented with two types of detectors. An extended scintillation counter will be placed alongside each of the structures to provide a measure of the expelled (uncaptured) field-emission current, and toroids will be installed in the beam line at the beginning, middle, and end of the linac to measure the transported current. The time-resolution of the scintillation counters will be approximately 20 ns. The time-resolution of the toroids will be approximately 1 ns.

With the beam off, we will measure the expelled and transported current signals as functions of the rf power and of the strengths of the intervening quadrupole and dipole magnetic fields. With a strong dipole field, any wall monitor currents should be due only to field emission in the structure immediately upstream. By observing the change in the currents with no dipole field, but with different-strength quadrupole configurations, we will characterize the transport of field-emission current in the linac. Although the quadrupole field strengths available in the NLCTA will not be as great as the strengths that will be used in an NLC main linac, we believe that they will be sufficient to fully suppress the transport of the field-emission current. Studies will also be done to measure how much of the field-emission activity is associated with the presence of the beam by varying the length of the bunch train and observing the change in the time-structure of the monitor signals.

Multibunch Beam-Loading Compensation

One of the goals of the NLCTA is to demonstrate that multibunch beam-loading compensation can be readily maintained at the 0.1% level for a beam current yielding a steady-state loading of 25%. To achieve the compensation, the shape and/or timing of the rf pulses will be adjusted so that the resulting change in unloaded energy gain along the bunch train offsets the energy loss from the loading. Demonstrating compensation at the 0.1% level will require careful preparation of the bunch train so as not to generate other sources of bunch-to-bunch energy differences. The bunch train injected into the linac will be checked for tolerances on the uniformity of bunch energy, intensity, and spacing. To measure the degree of compensation, the bunch train will be analyzed in the spectrometer using a vertical kicker magnet, which will spread the bunches on a screen. From the measurements of the rf waveform entering the structures, and from our theoretical understanding of the rf transmission through the structures we should be able to predict the gross features of the bunch-to-bunch energy differences that are observed.

The method for shaping the rf waveforms will be to modulate the phase of the rf drive to the klystrons, prior to SLED-II pulse compression. Maintaining the stability of the rf pulses may require feedback which should be relatively straightforward to implement in the rf control system of the NLCTA.

Transverse Components of the Accelerating Field

The tolerances on the allowable transverse components of the accelerating field in the structures of an NLC linac are fairly tight compared to SLC standards. These tolerances differ depending on whether the transverse electric field is in-phase or 90° out of phase with the bunches, and on whether the field jitters from pulse to pulse or is static. In the

NLCTA, we will be able to measure transverse electric-field components with resolution on the order of, or smaller, than the NLC tolerances.

The basic approach in these measurements is to vary the phase of one klystron at a time and to record the change in the bunch trajectories using the BPMs downstream from the associated pair of structures. Fitting the amplitudes of the observed betatron oscillations to a sinusoidal function of the klystron phase will yield the in-phase and out-of-phase static components of the accelerating field. To measure the jitter in these components, the pulse-to-pulse rms variation of the induced betatron motion will be computed with the klystron on and off, and at the in-phase and out-of-phase settings.

To obtain accurate results will require special attention to how the data are taken. A short bunch train will be used, only one linac klystron will be powered at a time, and the quadrupoles will be turned off downstream of the associated structures so that bunch energy changes do not produce trajectory changes. The measured outgoing trajectory will be corrected pulse-to-pulse to account for any changes in the incoming trajectory. This correction will be based on the trajectory fit to the BPM data in the region between the center of the chicane and the structures being powered. With this procedure, a resolution of about 0.2 keV in the transverse accelerating field should be achieved.

The out-of-phase component of the transverse field, which will degrade the bunch emittance in the NLC due to the differential kick it produces along the length of the bunch, should be less than 1 keV to keep the emittance growth below 10%. Therefore both static and jitter components of this magnitude should be discernible in the NLCTA. Jitter in the in-phase component of the transverse field will also lead to emittance growth in the NLC, as a result of the dispersion it generates. The tolerance for a 10% growth is much smaller, about 0.2 keV, so we will not be able to put tight limits on the NLC emittance growth from this mechanism if indeed no jitter is observed.

A static in-phase component of the transverse field is fairly benign since its effect can be offset with a dipole magnet. Coupled with jitter in the energy gain, however, it also contributes to the emittance growth in the same manner as in-phase jitter. A 1% jitter in energy gain, which is reasonable for the klystrons being developed, yields a tolerance 100 times larger, or 20 keV, on the static in-phase component. Our resolution for measuring this component is also much larger since an absolute measure of the bunch angular trajectories relative to the structure axes is needed to correct the measurement for the transverse field component due to off-angle trajectories. For this correction, we will rely on the absolute BPM calibration and on the BPM alignment to the structure axes, each of which will be known to about 100 μm. These errors translate into a 10-keV uncertainty in the static field strength so, again, the measurement resolution is near the tolerance.

Transverse Multibunch Beam Breakup

The NLCTA (like the NLC) would not operate stably if conventional constant-impedance accelerator structures were used. In both cases the offset of the bunches would be amplified by many orders of magnitude by the end of the linac. The detuned structures for the NLCTA will eliminate this blowup. This will be tested experimentally using the straight-ahead line of the spectrometer and the vertical kicker magnet.

Using corrector magnets in the chicane, the bunch train will be launched into the linac with a horizontal position or angle offset. The resulting positions of the bunches exiting the linac will be measured using a profile monitor in the straight-ahead line of the spectrometer in conjunction with the vertical kicker. This will allow us to distinguish the bunches along the train by their vertical offset while observing the amplification of the initial horizontal betatron amplitude.

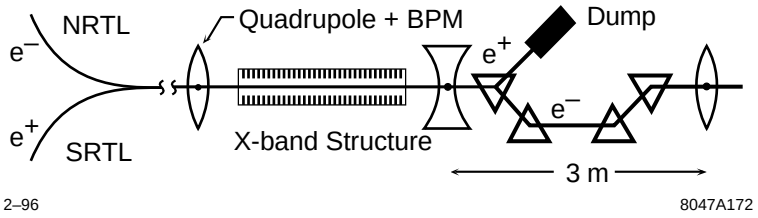


Figure 8-20. *Layout of the ASSET facility in the SLC.*

Studies with an Upgraded NLCTA Injector

There are a number of NLC linac beam-dynamics issues which cannot be readily addressed with the NLCTA as designed. The small bunch intensity produced by the X-band injector makes single-bunch longitudinal- and transverse-wakefield effects too small to be measured easily. Also, the small bunch spacing makes it extremely difficult to accurately measure the strength of the long-range transverse wakefield from its effect on betatron motion. We plan to upgrade the NLCTA injector in the future to produce NLC-like bunch trains with 0.75×10^{10} electrons per bunch and 1.4-ns bunch spacing. If a photocathode rf gun is used, we may be able to produce an adjustable-length bunch train (from 1 to 100 bunches) with a fairly small normalized (invariant) emittance ($\epsilon_n < 10^{-5}$ m-rad) and NLC-like bunch lengths (100 μm after compression).

With such an injector operating in a single-bunch mode, the mean energy of the bunch in the spectrometer would be measured as a function of bunch intensity to infer the short-range longitudinal wakefield strength. The size of this quantity is of particular interest since it has some theoretical uncertainty for such short bunches. Similarly, the strength of the short-range transverse wakefield would be inferred by measuring the change in the transverse bunch profile, using the wire scanner before the spectrometer bend, as a function of bunch intensity and betatron amplitude in the linac. To study the long-range transverse wakefields, a long bunch train would be used and the effects of the wakefields on the betatron motion of the bunches in the linac would be measured.

Until a new injector is built, some of these issues will be addressed with the Accelerator Structure SET-up (ASSET) facility at the SLC. (See Section 8.2.10.)

8.2.10 ASSET Measurements

To test the suppression of the transverse wakefields in the proposed NLC accelerator structures, a new facility has been incorporated into the SLC. This facility, the Accelerator Structure SETup (ASSET) [Adolphsen 1992] [Adolphsen 1995a], takes advantage of the low-emittance e^+ and e^- bunches that can be produced in the SLC, with individual control of the two bunch intensities and relative timing. Using the e^+ bunch as a probe, and the e^- bunch as a witness, the wakefields can be measured over an interbunch separation from zero to several hundred nanoseconds, with the sensitivity that is needed to verify the wakefield suppression required in the NLC.

The layout of ASSET is shown in Figure 8-20. One of the 3-m-long S-band accelerating sections has been removed from the SLC, and the X-band structure to be tested is inserted in its place, followed by a combination of a beam dump (for the e^+ probe bunch) and chicane (for the e^- witness bunch). The wakefield kick from the probe bunch is determined from the betatron oscillations it induces on the witness bunch.

Measurements of the wakefield amplitude in a 1.8-m-long detuned X-band structure are shown (points) in Figure 8-21, along with theoretical predictions of the amplitude (solid curves), with and without cell frequency errors included.

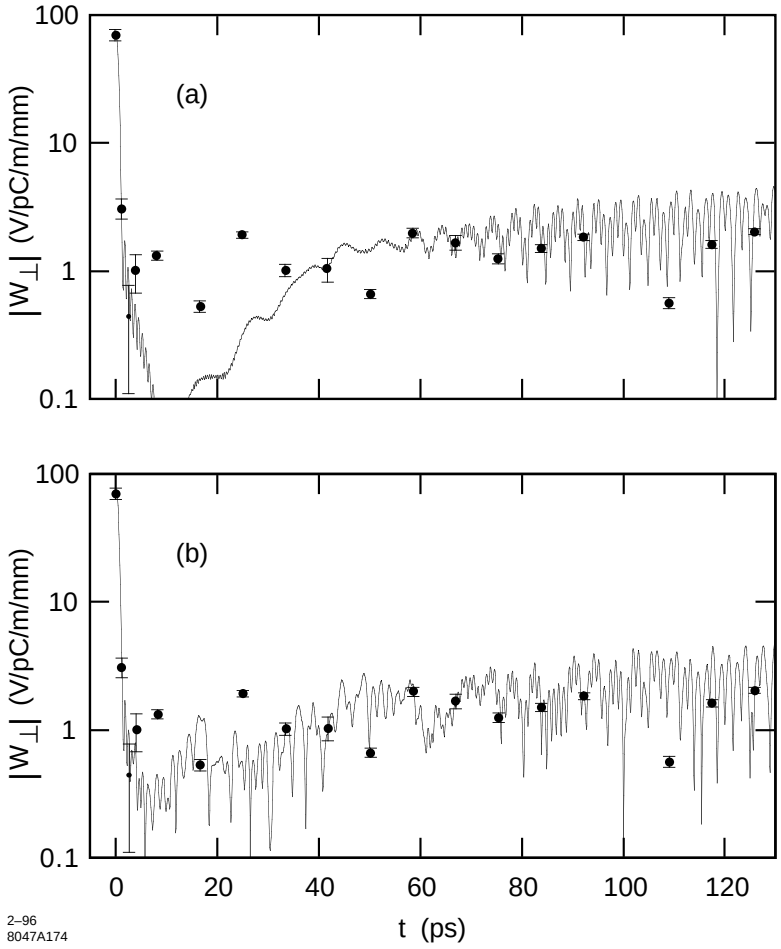


Figure 8-21. *Dipole wakefield amplitude measurements and prediction from equivalent circuit model (a) without cell frequency errors and (b) with* 1.5×10^{-4} *rms fractional frequency errors.*

The fractional frequency errors have an rms of 1.5×10^{-4}. The agreement, when frequency errors are included, is reasonably good, although the wake is somewhat more suppressed at long distances than was predicted by the equivalent circuit model, assuming Q values of 6500 for the lowest-dipole deflecting mode.

Measurements of the wakefield in the DDS will be made using ASSET when the first prototype DDS becomes available.

8.2.11 Use of Beam-Excited Modes to Monitor Alignment of Structures

Each dipole mode in the DDS (as well as the structure with detuning alone) is localized in a particular region of the structure, depending on the mode's frequency and where that frequency is cut off in the structure. Thus it is proposed to pick up the frequency of two modes, one localized near each end of the structure, and use them to measure the alignment of the structure with respect to the beam.

The desired dynamic range for the detection of position offsets in such a system is large—from about a millimeter down to a micron. Details on the design of such structure beam position monitors are given in Section 7.10.3.

8.3 RF Pulse Compression and Power Transmission

This section describes the use of rf pulse compression to multiply peak power in the NLC linac high-power rf system. The 1.2-μs-long X-band klystron pulses will be compressed by a factor of five—to 0.24 μs—using the SLED-II rf pulse compression technique [Wilson 1990]. SLED-II is a modification of SLED [Farkas 1974]. SLED is used to increase the rf power into the 60-GeV SLAC linac by a factor of about 3. In this section, we discuss the performance required, present a physical layout, and analyze its efficiency.

8.3.1 Performance

A SLED-II pulse compressor works by storing microwave energy in a pair of high-Q resonant delay lines for most of the duration of the klystron pulse. The round-trip transit time of the rf in the delay lines determines the duration of the compressed pulse. The SLED-II microwave network is shown schematically in Figure 8-22. The delay lines are coupled to input and output waveguides by a four-port 3-dB coupler (such as a "Magic T"). In normal operation with a pulse-compression ratio of five, the delay lines are filled with microwave energy for four round-trip-delay times. The microwave energy stored in the delay lines is discharged by reversing the phase of the drive power to the klystron. The stored energy then is discharged through the fourth port of the 3-dB coupler, where it combines constructively with power still coming from the klystron, forming the compressed, high-power pulse.

To achieve low losses, oversize circular waveguide, propagating the TE_{01} mode, has been used for delay lines in SLED-II prototypes. Delay-line quality (Q) factors between 4.3×10^5 and 1.0×10^6 have been achieved [Tantawi 1995a, Tantawi 1996]. Converting between the TE_{01} circular-waveguide mode and the TE_{10} rectangular-waveguide mode has been performed by compact, flower-petal type, low-loss mode transducers (see below) [Lanciani 1953, Hoag 1993, Tantawi 1993]. Figure 8-23 shows the demonstrated performance of a prototype high-power SLED-II pulse compressor at SLAC. The small variations on top of the compressed pulse are not a property of the SLED-II pulse compressor, as demonstrated by low-power testing that produced a flat pulse [Tantawi 1995a], but are due to mode impurities in the transmission line from the klystron to the pulse compressor. The design of the mode transducer has been modified, to better suppress mode impurities in the oversized TE_{01} waveguide, by adding a choke for TE_{41} and a pair of interacting chokes for TE_{11} and TM_{11} [Tantawi 1995b]. The effectiveness of this modification has already been demonstrated in a low-power "cold test" of the new mode-transducer design.

In general, there is a trade-off between the peak-power gain achieved by pulse compression, and the efficiency of the pulse compressor. For SLED and SLED-II pulse-compression systems, the trade-off is shown in Figure 8-24 [Wilson 1992]. The intrinsic inefficiency of SLED-type pulse compressors results from the inability of a SLED energy storage cavity, or a SLED-II resonant delay line, to charge or discharge completely, in a finite amount of time, through the coupling aperture that is shown schematically in Figure 8-22. A SLED-II pulse compressor with a compression ratio of five has a maximum (intrinsic) efficiency of 80.4%, which is achieved when the voltage-reflection coefficient of the coupling aperture is 0.651. (Hence, 0.651^2, or 42% of the power incident on SLED-II is reflected at the beginning of the klystron pulse.)

The net peak-power gain is the product of the pulse-compression ratio (5), the intrinsic efficiency of the pulse compressor (0.804), and the cumulative efficiency of all of the waveguides and other microwave components in the pulse compressor and transmission lines. The minimum peak power gain required to achieve the NLC-design

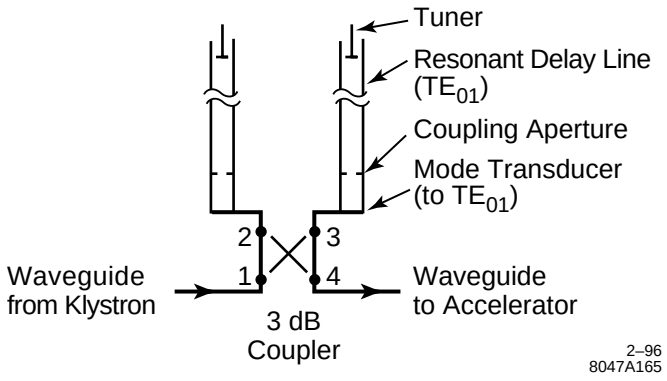


Figure 8-22. *Microwave network of a SLED-II pulse compressor.*

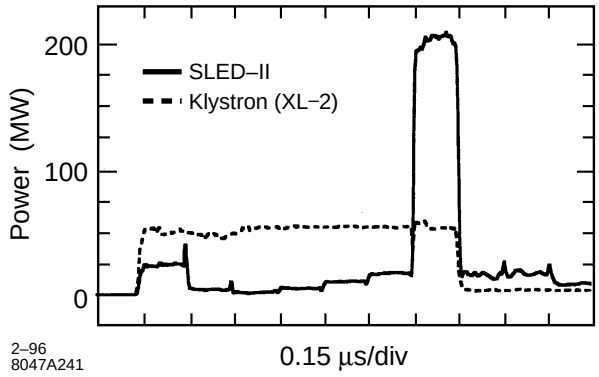


Figure 8-23. *Demonstrated performance of a prototype high-power SLED-II pulse compressor.*

accelerating gradient is 3.6. Therefore, the cumulative losses that occur due to finite conductivity and mode conversion in the entire system must be limited to 10% to achieve the desired power gain product:

$$5(0.804)(1-0.10) = 3.6 \quad . \tag{8.20}$$

Developmental SLED-II systems have been tested at SLAC with high-power X-band klystrons to validate the design and its components [Nantista 1993, Vlieks 1993, Wang 1994]. The peak power and efficiency needed for the NLCTA have been demonstrated in a prototype system. Further validation of the SLED-II design at the higher power levels and efficiencies needed for the NLC will be performed in upgrades to the NLCTA. Table 8-7 summarizes the performance of the SLED-II prototype and the plans for using SLED-II in the NLCTA and in the NLC. The rf pulse compression systems in the NLC and NLCTA are designed to accommodate future upgrades in which the peak rf power is nearly tripled, by replacing each 50-MW klystron with a pair of 75-MW klystrons, as indicated in Table 8-7.

8.3.2 Physical Layout

Figure 8-25 illustrates the physical layout of the waveguide network for rf pulse compression and power transmission to four 1.8-m-long linac sections for the two-tunnel (“SLAC-type”) configuration shown in Figure 7-11.

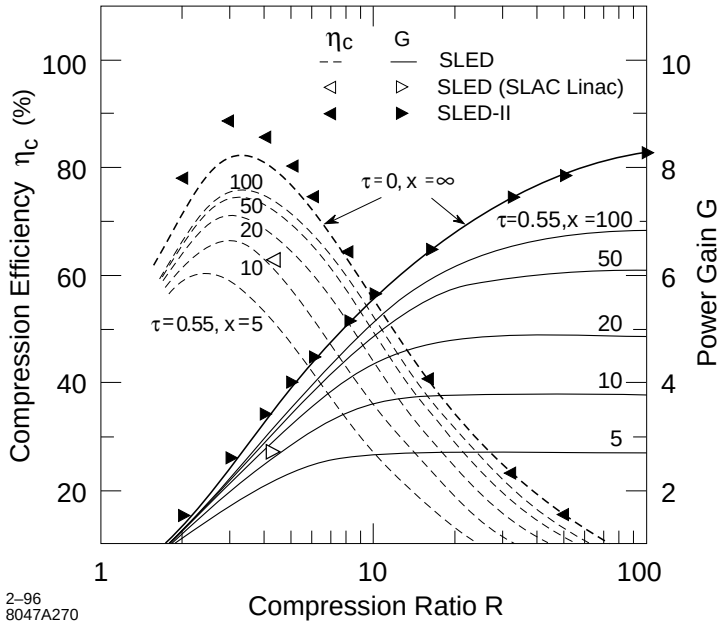


Figure 8-24. *Power gain (solid curves) and compression efficiency (dashed curves) for a SLED-II rf pulse compressor driving an accelerator structure with τ and x as indicated. τ is the voltage attenuation parameter (nepers) for the structure. x is the unloaded time constant of the pulse-compression energy-storage cavities ($2Q/\omega$) normalized to the structure filling time. The top curves show gain and efficiency for SLED as implemented for the SLAC linac. The solid triangular points are for a SLED-II system with lossless components.*

To achieve low rf power losses, oversized circular waveguides will be used for the SLED-II delay lines, and for the high-power transmission lines. The TE_{01} mode will be propagated in the circular waveguides that comprise the delay lines, in the transmission lines from the klystrons to the power combiners, and in the transmission lines from the klystron gallery to the linac tunnel. The waveguide diameters that have been chosen to avoid spurious-mode cut-offs near the operating frequency are 4.75 in (WC475) for the delay lines and 2.93 in (WC293) for the transmission lines from the klystrons. The HE_{11} "hybrid" mode will be propagated in flexible corrugated circular waveguides of approximately two-inch inner diameter that distribute power to the four accelerator sections powered by each SLED-II pulse compressor.

Rectangular waveguides, propagating the TE_{10} mode, also will be used in parts of the transmission system. The rectangular waveguide sizes that have been chosen are WR100 (1.0 in $\times$ 0.5 in) and oversize WR100 (1.0 in $\times$ 0.9 in).

Conversion between the TE_{01} circular-waveguide mode and the TE_{10} rectangular-waveguide mode will be performed by compact, flower-petal type, low-loss mode transducers [Lanciani 1953, Hoag 1993, Tantawi 1993]. Transducers of this type, shown in Figure 8-26 have been analyzed, measured, and used extensively at SLAC.

The four-way power splitter that will divide the compressed rf pulses into four equal parts has been designed by computer simulation [Kroll 1995], but the design has not yet been manufactured and tested. The four-way splitter is shown in Figure 8-25; an enlarged view is shown in Figure 8-27. The device accepts the TE_{01} circular-waveguide mode, and divides the power equally into four TE_{10}-mode rectangular waveguides.

Matching the HE_{11} "hybrid" mode in corrugated circular waveguide to the TE_{10} mode in rectangular waveguide may be performed by relatively simple tapers, since the two modes have essentially the same character.

	Prototype 3/95[a]	NLCTA Design	NLCTA Upgrade	NLC Design	NLC Upgrade
Klystron pulse width	1.05 μs	1.5 μs	1.5 μs	1.2 μs	1.1 μs
Compressed pulse width	0.15 μs	0.25 μs	0.25 μs	0.24 μs	0.22 μs
Compression ratio	7	6	6	5	5
Klystron peak power	55 MW	50 MW	$4\times$(50–75) MW	2×50 MW	4×75 MW
Intrinsic efficiency of SLED-II without losses[b]	0.692	0.746	0.746	0.804	0.804
Efficiency of SLED-II components	0.92	0.90	0.90	0.95	0.95
SLED-II gain (Max. possible)	4.42 (4.82) [c]	4.0 (4.48)	4.0 (4.48)	3.8 (4.02)	3.8 (4.02)
Efficiency of transmission lines	0.84	0.90	0.90	0.95	0.95
Transmitted power	205 MW	180 MW	720–1080 MW	360 MW	1080 MW
Net power gain	3.7	3.6	3.6	3.6	3.6

[a] Achieved.

[b] Assumes the delay-line reflection coefficient is optimized for the stated compression ratio.

[c] The delay-line reflection coefficient was optimized for a compression ratio of 8 in the prototype test.

Table 8-7. *Performance of SLED-II rf pulse compressors.*

The waveguide layout is sufficiently flexible to absorb the absolute motion of the accelerator sections due to their mechanical movers (± 1 mm) and the relative motion between the klystron gallery and the accelerator tunnel due to settling (± 10 mm).

The SLED-II waveguide networks are located in the low-radiation environment of the klystron gallery so that the rf pulse-compression system may be modified, station by station, if necessary for upgrade to higher beam energies, simultaneously with colliding beam operations.

Since the round-trip delay time in the 4.75-in-diameter waveguide (at the group velocity, 0.964c) must equal the rf pulse duration (0.24 μs), the physical length of the SLED-II delay lines is 34.7 m. The physical layout can accommodate this delay-line length by spatially overlapping the delay lines from adjacent rf stations. Since the total length of linac fed by each rf station is approximately eight meters, each pair of delay lines will overlap parts of five other pairs of delay lines.

The alternate tunnel configuration, the semi-elliptical highway-style tunnel shown in Figure 7-12 is topologically similar to the physical layout presented above. The waveguide configurations differ only in the lengths of very low-loss oversized circular waveguides. The circular waveguide from each klystron to the power combiner is longer (3 m *vs.* 1 m) since it must pass through the shielding wall. The waveguide from the pulse compressor to the accelerator is much shorter (1 m *vs.* 10 m) since, in this configuration, the pulse compressors are in close proximity to the linac.

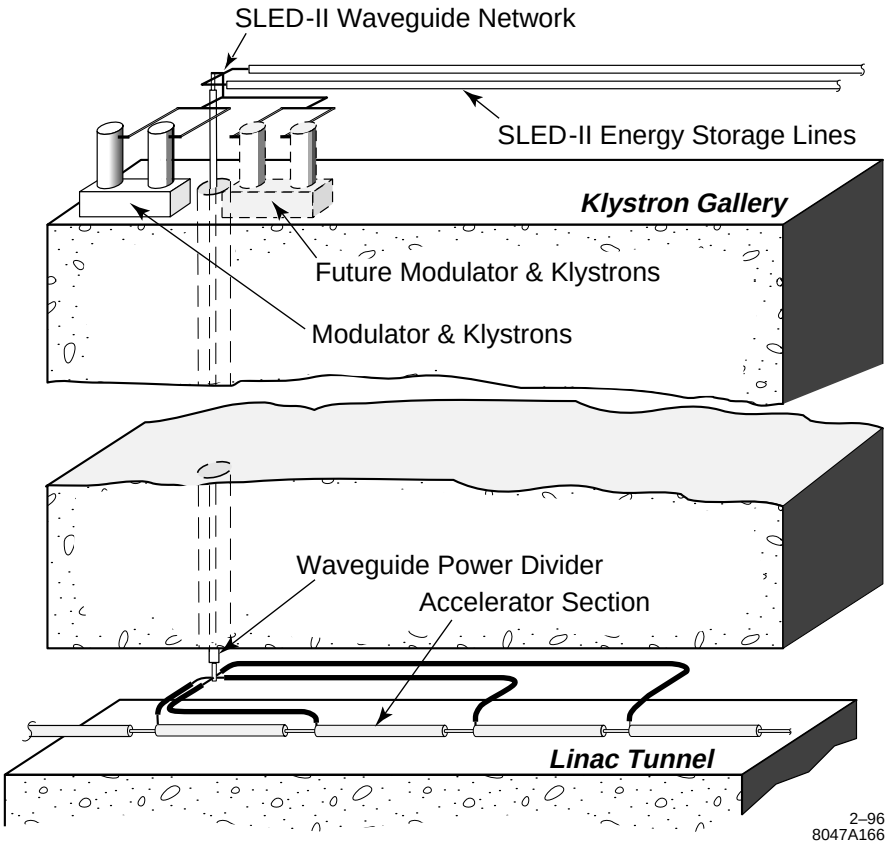


Figure 8-25. *Physical layout of the waveguide network for rf pulse compression and power transmission.*

8.3.3 Power Losses

To realistically estimate the peak power gain of the physical layout illustrated in Figure 8-25 we have constructed a “loss budget” for the rf system based on the separate components from which the above physical layout might be assembled. The losses assumed for the individual components are based on a combination of actual experience with identical or similar devices, and theoretical estimates derated for imperfect conductivity and manufacturing tolerances.

Ohmic losses in all waveguides are assumed to exceed standard theory [Ramo 1984] by 20% to allow for imperfect conductivity and surface roughness. This has been confirmed empirically for the NLCTA's WC293 transmission lines.

The 3-dB power combiner/divider is assumed to be a WR100 “Magic T” with one-way power loss of 1%. Our experience with WR90 Magic Ts consistently has yielded losses of about 1.2%, due to a combination of ohmic loss, energy trapped in the junction, and finite reflection. Using a WR100 is expected to reduce the ohmic contribution by the ratio of ohmic losses in WR100 compared to WR90, which is 3/4.

The power loss assumed for the flower-petal-type mode transducer, 0.7%, is consistent with the measured loss of the NLCTA mode transducers.

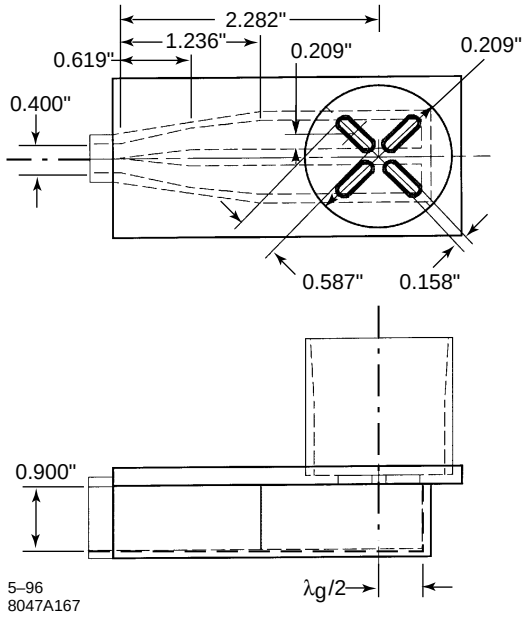


Figure 8-26. *Mechanical drawing of the "flower-petal" mode transducer.*

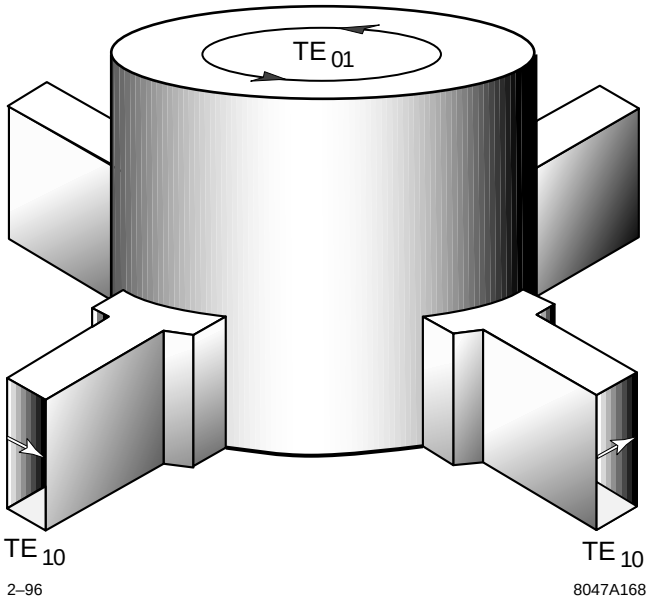


Figure 8-27. *Sketch of the four-way power splitter designed by computer simulation.*

Power losses in the four-way power splitter are assumed (in the absence of a measured prototype) to be the same as for the flower-petal-type transducer, 0.7%. Less loss is expected due to the symmetry of the device, which guarantees mode purity (in the absence of mechanical imperfections).

The loss budget for the physical layout shown in Figure 8-25 is summarized in Table 8-8. In the table, power losses and transmission fractions are tabulated for individual components and for groups of components, including the SLED-II pulse compressor, the transmission lines from the klystrons, and the transmission lines to each accelerator section. The net power gain is tabulated separately for each of the four accelerator sections. The table presents a detailed accounting of the power losses for each of the four accelerator sections. The average power transmission is 89.6%, resulting in a net power gain (averaged over the four accelerator sections) of 3.60.

For the semi-elliptical highway-style tunnel configuration shown in Figure 7-12, the waveguide configuration differs only in the lengths of very low-loss oversized circular waveguides. The loss budget and the resulting net power gain are nearly identical to those in the two-tunnel (SLAC-type) configuration.

8.4 High-Power Klystrons

The design of the 50-MW (X5011) klystron is the product of an ongoing development program which began at SLAC in 1988. At that time, the peak power, frequency, and pulse duration required in combination were well beyond the state of the art. As a result, rf circuit breakdown and output window failures plagued early development efforts at SLAC. Through a steady stream of improvements, the failure modes are largely under control. SLAC presently has four 50-MW solenoid-focused klystrons in operation. Three of these are suitable for use in the NLC Test Accelerator. The PPM-focused X5011 klystron, in which a periodic permanent magnet (PPM) array replaces the solenoid, is in the final stages of computer simulation, mechanical design and fabrication. Those parameters which have not yet been demonstrated have been shown to be achievable through computer simulation.

The upgrade of the NLC to 1 TeV center-of-mass energy requires a 75-MW peak power klystron. The preliminary electrical and mechanical design of this klystron, designated X7511, has been carried out. The criteria applied in the design of the 75-MW klystron were not to exceed the design limits imposed on the 50-MW klystron and, to the greatest extent possible, to stay within the gradients and stress levels demonstrated in the X5011. The upgrade is discussed in Section 8.4.3, “Upgrading Performance”.

Design and performance parameters for the X5011, 50-MW X-band klystron are given in Table 8-9. Key features of the tube are identified in Figure 8-28.

8.4.1 Design Features

Table 8-9 lists the design and performance parameters which make up the preliminary specification for the X5011 klystron. The design features and innovations which make the specification achievable are described below.

Electron Gun

Figure 8-29 shows the electrode configuration of the electron gun and the electron beam which is produced, based on the computer code EGUN. The cathode diameter (5.72 cm) is a compromise. A larger cathode would reduce the maximum voltage gradient (230 kV per cm, which is high, but acceptable), but would increase the beam area

Component	Unit	Loss	Transmission Subtotals				Net Gain[a]	
		(Linear)	(Linear)	(dB)	(dB)	(Linear)	(dB)	(Linear)
Definitions of Components								
TE_{01}–TE_{10} Transducer	(each)	0.0070	0.993	−0.031				
WR100 "Magic T"	(each)	0.0100	0.990	−0.044				
WR100 ($1.0'' \times 0.5''$)	(m)	0.0205	0.979	−0.090				
Oversize WR100 ($1.0'' \times 0.9''$)	(m)	0.0133	0.987	−0.058				
Corrugated WC220 (HE_{11})	(m)	0.0011	0.9989	−0.0048				
WC293 (TE_{01})	(m)	0.0010	0.9990	−0.0042				
WC475 (TE_{01})	(m)	0.00022	0.99978	−0.00093				
Trans. Line from Klystron					**−0.137**	**0.969**		
WC293 (TE_{01})	1.0	0.0010	0.999	−0.004				
Mode Transducer	1	0.0070	0.993	−0.031				
Oversize WR100	0.5	0.0067	0.993	−0.029				
WR100 "Magic T" Combiner	1	0.0100	0.990	−0.044				
Oversize WR100	0.5	0.0067	0.993	−0.029				
SLED-II Pulse Compressor					**−0.233**	**0.948**		
WR100 "Magic T"	2	0.0199	0.980	−0.087				
Mode Transducer	2	0.0140	0.986	−0.061				
WC475 Delay Lines (TE_{01})	34.7	0.0193	0.981	−0.084				
Trans. Line to Linac Tunnel					**−0.070**	**0.984**		
Mode Transducer	1	0.0070	0.993	−0.031				
WC475 (TE_{01})	10	0.0021	0.998	−0.009				
Mode Transducer (4-way split)	1	0.0070	0.993	−0.031				
Trans. Line to Section #1					**−0.023**	**0.995**	**5.580**	**3.614**
WR100 Taper	0.1	0.0021	0.998	−0.009				
Corrugated WC220 (HE_{11})	1.0	0.0011	0.999	−0.005				
WR100 Taper	0.1	0.0021	0.998	−0.009				
Trans. Line to Section #2					**−0.030**	**0.993**	**5.572**	**3.608**
WR100 Taper	0.1	0.0021	0.998	−0.009				
Corrugated WC220 (HE_{11})	2.5	0.0028	0.997	−0.012				
WR100 Taper	0.1	0.0021	0.998	−0.009				
Trans. Line to Section #3					**−0.040**	**0.991**	**5.563**	**3.600**
WR100 Taper	0.1	0.0021	0.998	−0.009				
Corrugated WC220 (HE_{11})	4.5	0.0050	0.995	−0.022				
WR100 Taper	0.1	0.0021	0.998	−0.009				
Trans. Line to Section #4					**−0.049**	**0.989**	**5.553**	**3.592**
WR100 Taper	0.1	0.0021	0.998	−0.009				
Corrugated WC220 (HE_{11})	6.5	0.0072	0.993	−0.031				
WR100 Taper	0.1	0.0021	0.998	−0.009				
Average over all four sections					**−0.475**	**0.896**	**5.567**	**3.603**

[a] Net gain includes SLED-II compression ratio (5) and intrinsic efficiency (0.804).

Table 8-8. *Summary of the microwave power losses, power-transmission fractions, and power gains. The losses assumed for the separate components are defined in the topmost section.*

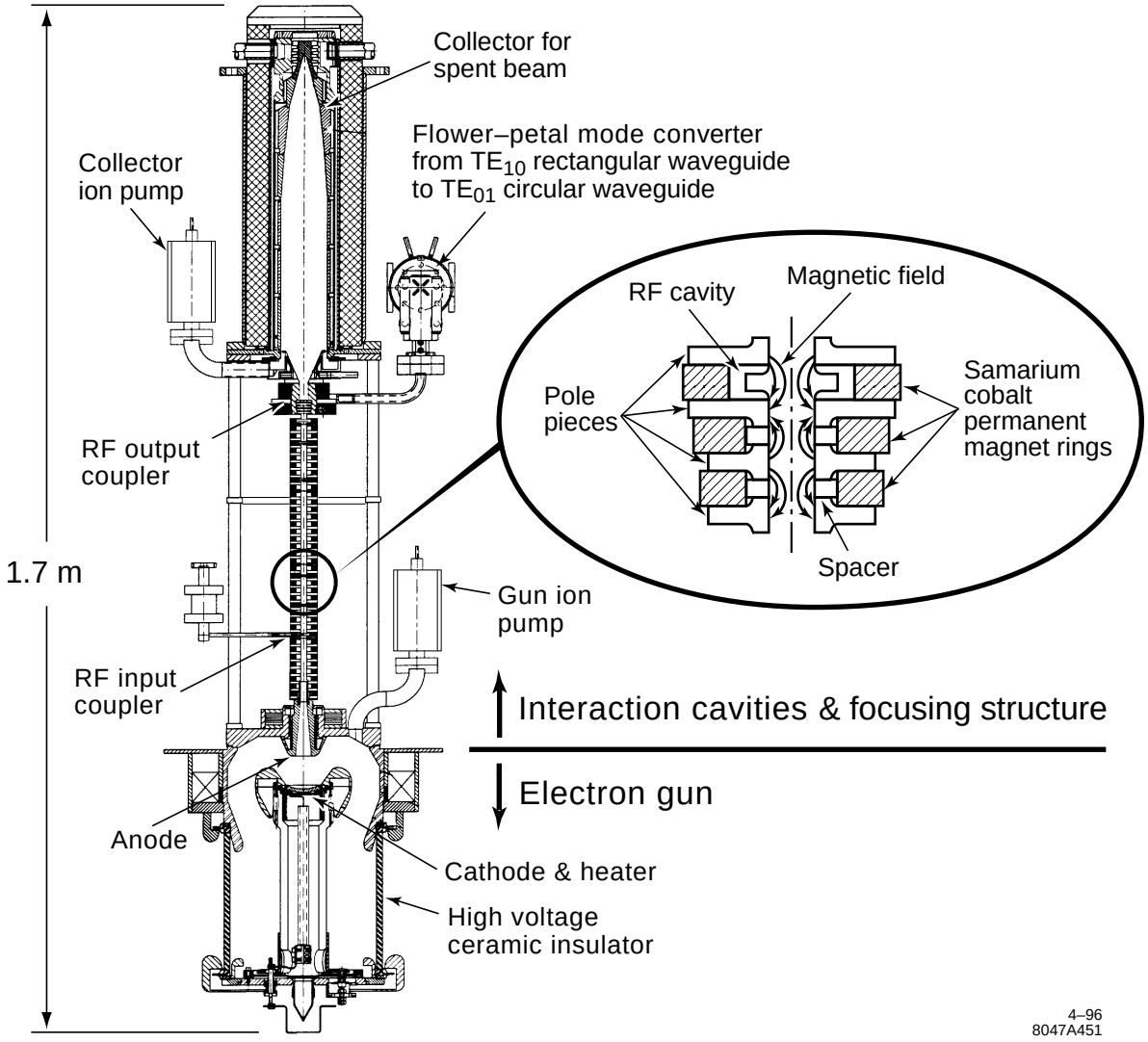


Figure 8-28. *Preliminary layout of the X5011 klystron, with water jacket removed. The insert shows one representative buncher cavity, to illustrate the use of oversize permanent magnets.*

Parameter	Units	Value
Electrical:		
Frequency	(GHz)	11.424
Peak Power	(MW)	50
Beam Voltage	(kV)	465
Beam Current	(A)	190
Beam Microperveance	(μA/ V$^{3/2}$)	0.6
Gain	(dB)	57
Bandwidth	(MHz)	100
Beam Modulation		Cathode Pulsed
Pulse Duration	(μs)	1.20
Rep Rate	(Hz)	180
RF Efficiency	(%)	57
Cathode Current Density	(A/cm^2)	7.4
Heater Voltage	(V)	15
Heater Current	(A)	21.5
Mechanical and Magnetic:		
Beam Focusing		Periodic Permanent Magnet (PPM)
Tube Weight	(kG)	32 (without lead)
Tube Length	(m)	1.3
Cathode Diameter	(cm)	5.72
Cathode Half Angle	(°)	25
Beam Area Convergence		144

Table 8-9. *Design and performance parameters for the X5011 klystron.*

convergence (currently 144/1, also high, but achievable). The resulting cathode current density, which averages 7.4 A cm^2, with a peak value of 10 A/ cm^2 at the cathode edge, is consistent with long cathode life. The 7.4 /Amp/ cm^2 compares with 6 A/ cm^2 in the 5045 SLC klystron, which has a mean time between failures (MTBF) in excess of 40,000 hours.

Elimination of the solenoid made possible by the PPM beam focusing removes any constraint on the size of the high-voltage ceramic, which allows for conservative high voltage-seal gradients. The absence of the solenoid also makes room for an ion pump which will be used to pump the gun directly through the anode, further reducing the risk of high-voltage gun breakdown.

Beam Focusing

Figure 8-28 shows details of the periodic magnetic structure and magnets which supply the beam focusing magnetic field. The magnetic circuit is designed to minimize the volume of the relatively high-cost samarium cobalt ring magnets. Oversized magnets are required at the location of the gain and buncher cavities as shown. A segmented magnet is used to supply the unidirectional field for the extended interaction output circuit (not shown).

Two key features of the magnetic circuit design make it possible for the periodic magnetic field to provide focusing, with and without rf drive, which is comparable with that provided by a high field solenoid.

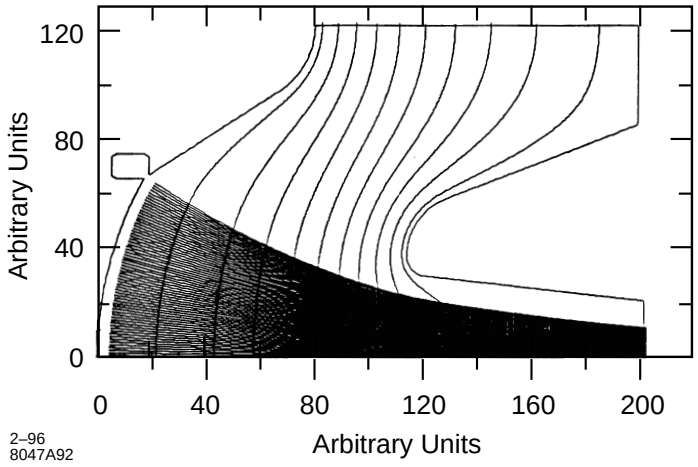


Figure 8-29. *Computer printout of gun electrodes and electron beam for the X5011.*

- **Short Magnet Period:** The periodic length of the PPM focusing structure used here is about one-half that customarily used in traveling-wave tubes, most of which are PPM focused. As a result, the stop-band voltage, the voltage below which all focusing ceases, a feature of all axially symmetric PPM focusing, occurs at about 6% of operating voltage where it does not materially worsen intercepted beam current during the pulse. By contrast, traveling-wave tubes with their longer magnet periods are forced to operate with stop-band voltage ratios of 25–50% of operating voltage.

  An even more important benefit of the short magnet period is that it produces a contoured field, having an rms value at the beam tunnel wall as much as twice that on the axis, which results in a high degree of beam stiffness. The higher field at the wall is capable of focusing a beam having a current density almost four times that which can be focused on the axis.

- **Tailored Field Amplitude:** Another advantage of PPM focusing over solenoid focusing is that the field amplitude can be tailored to a much greater degree to match the known beam conditions. As the beam is bunched and charge density increases in passing through the buncher cavities, the on-axis magnetic field required to maintain optimum focusing increases. The required tailoring of the magnetic field amplitude is easily carried out by simply using either larger diameter or higher-energy product magnets in the region where a higher field is required.

Figure 8-30 is a plot of the on-axis magnetic field which has been found in computer simulation to produce the best rf performance while avoiding beam interception. Figure 8-31 is a computer simulation of the focused beam with no rf drive.

RF Circuit

Figure 8-32 shows the layout of the rf circuit cavities approximately to scale and in their correct relative positions. The circuit consists of an input plus two gain cavities, three buncher cavities, and a five-cell, disk-loaded, extended-interaction, traveling-wave output circuit. These are described in more detail below:

Gain Cavities: The input and gain cavities are tuned near synchronism for maximum gain but with enough stagger to provide the bandwidth needed to ensure proper operation of the pulse compression system.

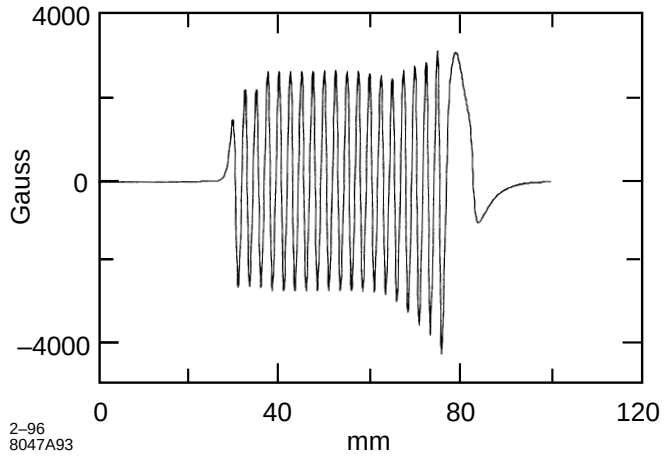


Figure 8-30. *Magnetic field profile for best efficiency without beam interception.*

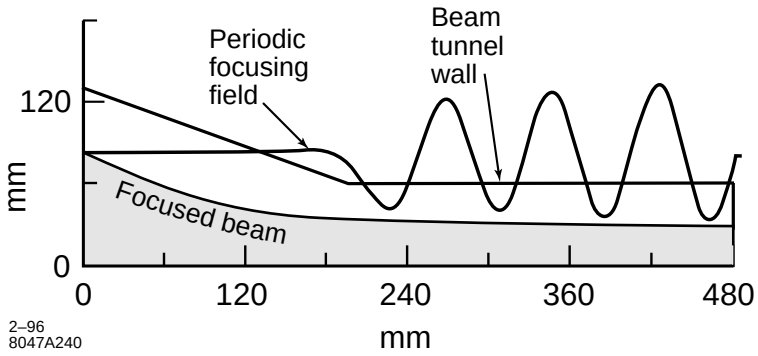


Figure 8-31. *Computer simulations of X5011 PPM-focused beam in the launch region. The periodic focusing field is also shown.*

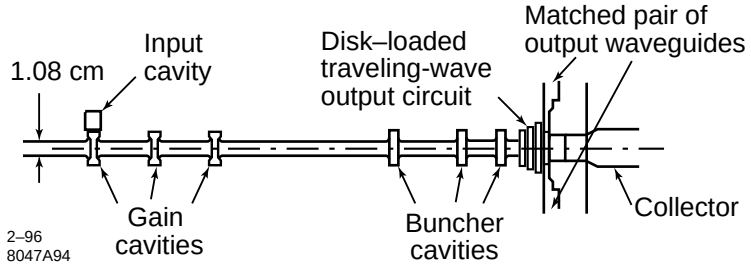


Figure 8-32. *Rf circuit cavity layout, drawn approximately to scale.*

Figure 8-33. *Computer simulation of bunched beam passing through the X5011 disk-loaded output circuit. The vertical scale is magnified by 3 to emphasize the growth of bunch amplitude in the output circuit.*

Buncher Cavities: The three buncher cavities replace the single penultimate cavity found in most klystrons. The buncher function was split into three parts to maintain an acceptable rf surface gradient. It was discovered that this approach has the further advantage that it provides a more adiabatic bunching of the beam with a higher rf current (I_1/I_0) which enhances efficiency.

Output Circuit: For the same reason that three buncher cavities are used, it was found necessary to employ an extended interaction type of output circuit to control rf breakdown. For high efficiency, rf voltage across the total length of the output circuit must be comparable with the 465-kV beam voltage. The maximum surface gradient is necessarily much higher than the on-axis gradient.

The output circuit is tapered in both impedance and velocity. Tapering the impedance from cavity to cavity by a factor of about five, from highest in the first cavity to lowest in the fifth cavity, produces a nearly constant rf voltage per gap. The velocity taper ensures that the wave remains in synchronism with the slowing beam for maximum energy extraction. It is fortuitous that the impedance taper is achieved with a physical taper of the circuit, with each successive cavity disk having an increased inner diameter, as efficiency is enhanced by allowing the highly bunched beam to expand as it traverses the output circuit. The circuit taper accommodates the beam expansion, free of beam interception. The computer simulation "snapshot" of the bunched beam traversing the output circuit under saturation conditions of Figure 8-33 clearly illustrates this phenomenon. The computed efficiency in the simulation was 63%. We have derated the specification efficiency to 57%, pending hot tests. A 10% reduction in measured versus simulated performance is not uncommon.

Output Coupler and RF Window

Rf power exits the klystron and passes through a pair of symmetrically-disposed, rectangular waveguides. It is then combined in a single rectangular waveguide before being converted to the TE_{01} circular waveguide mode, using a flower-petal mode converter. The power exits the klystron through a TE_{01} disk window. Figure 8-34(a) is a sketch of the circular disk window and matching irises which complete the rf output coupling from the tube. The matching irises and quarter wavelength ceramic disk window in combination are designed to produce a traveling-wave through the ceramic disk, as opposed to the standing wave which is used in the more conventional half-wave resonant window. The advantage of the traveling wave is that the rf gradient on the window surface is half as great for a given power flow as that for the standing-wave window. The lower gradient reduces the chance of window breakdown. The advantage

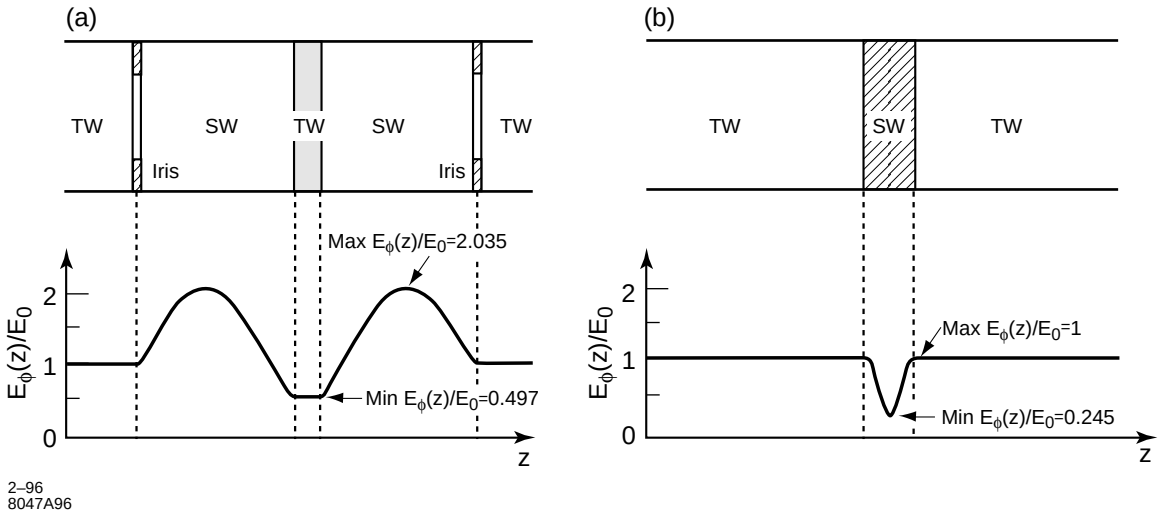


Figure 8-34. *Comparison of the reduced-field TW window (a) with a half-wave resonant SW window (b), both operating in the TE_{01} circular mode. In addition to the TW version having substantially lower rf electric field at the window surface, the integrated dielectric power losses are only 23% those of the half-wave resonant window. The horizontal scale is expanded for clarity.*

of coupling the power through the window in the TE_{01} mode is the total absence of rf electric field terminating on the metal wall of the waveguide, again enhancing power handling capability.

8.4.2 Results To Date

Status

SLAC currently has four operating 50-MW, X-band klystrons. Three of these will serve as klystrons for the NLCTA. The first of the series, XL-1, is not suitable for the NLCTA because of insufficient bandwidth. XL-2 and XL-3 operate at 50-MW peak power, at approximately 440-kV beam voltage, and at 36%–40% rf efficiency, depending upon operating settings. Preliminary testing of XL-4 has demonstrated up to 48% rf efficiency. These solenoid-focused klystrons use a microperveance 1.2 electron gun, with an input cavity, two gain cavities, and three buncher cavities. Both XL-1 and XL-2 have three-cell, disk-loaded, standing-wave, extended-interaction output circuits while the disk-loaded circuits in XL-3 and XL-4 are traveling wave. The 2.5D CONDOR computer simulation for XL-1 and XL-2 predicts a maximum efficiency of approximately 43% which yields about 10% more power than obtained experimentally. The maximum efficiency measured with XL-3 was 43% with a simulated value of 50%. XL-4 incorporates rf cavity design refinements and spurious mode suppression techniques which result in a simulated efficiency of 52% compared with the preliminary measured value of 48%.

Figure 8-35 is a pictorial overview of XL-1, including a drawing of the tube and solenoid, a table of design and performance parameters, and an oscillographic recording of beam current and rf input and output pulses.

XL-4, the most recent refinement of the solenoid-focused 50-MW klystron designs, has produced clean, stable rf pulses of 50-MW peak power for pulse durations up to 2 μs, and 75-MW peak power for pulse durations up to 1.1 μs. This performance exceeds the pulse durations required for the NLC design and upgrade. The 75-MW, 1.1 μs pulse is shown in Figure 8-36.

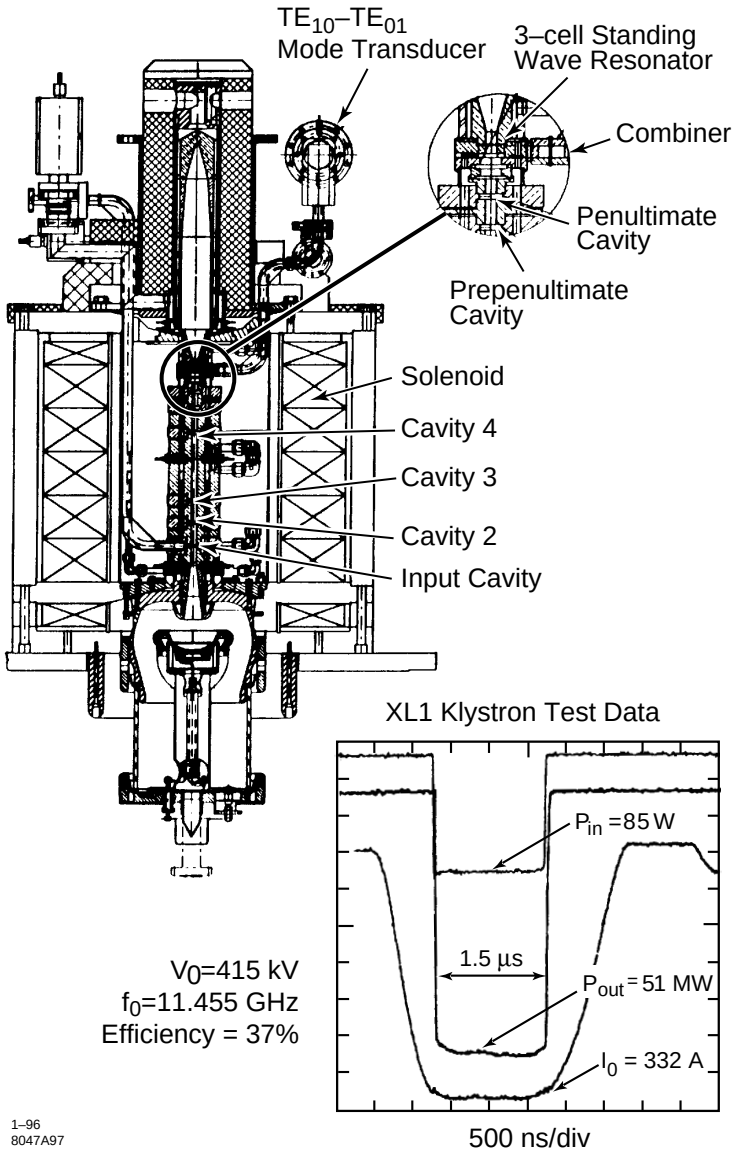


Figure 8-35. *Overview of XL-1, 50-MW klystron, and high-power test results.*

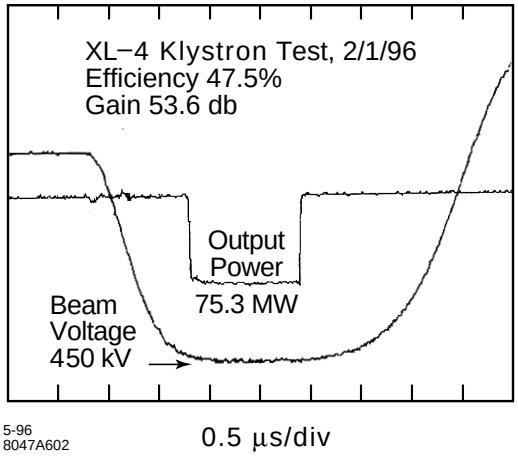


Figure 8-36. *High-powered test of XL-4 klystron. Output power, 75.3-MW. Pulse duration, 1.1 μs.*

Development History

All experimental results to date are based on solenoid-focused tubes. The XL series of klystrons is the most recent. These were preceded by the XC series of eight higher current tubes (microperveance 1.8). It was the experience gained from this earlier series of tubes that narrowed the choice of rf circuits suitable for the required peak power and pulse length, and that led to gun and rf window improvements which resulted in functional, long-pulse, 50-MW klystrons. Highlights of the development history follow.

Choice of Voltage and Current. The initial objective was to achieve 100 MW of peak power at 1 μs pulse length. It was necessary that these initial tubes match the capability of the modulators on hand at SLAC, which were limited to 440 kV. This, in turn, required high beam currents (1.8 microperveance) to obtain sufficient beam power to have a chance of producing 100 MW of peak power. The plan was to convert to higher voltage, lower current designs (for higher rf efficiency) when an appropriate modulator came online.

The choice, four years into the program, to develop 50-MW tubes for the NLCTA, made it possible to reduce the beam microperveance to 1.2 at the same 440-kV voltage with a consequent improvement in efficiency. This led to the XL tubes described earlier. In order to obtain equivalent power from PPM-focused tubes, the beam perveance needed to be reduced still further. This adjustment was imposed by the limit on the rms magnetic field obtainable from permanent magnets. The available range varies from microperveance 0.47 on the low side, dictated by the 500-kV maximum desirable beam voltage, to 0.9 on the high side, dictated by magnetic field considerations. The first PPM klystron, the X5011, described in Table 8-9, is currently in the design stage and will be tested in mid-year 1996. Computer simulations, showing up to 63% rf efficiency, have been carried out. The tube is designed to produce 50 MW of peak power at 465 kV at a microperveance of 0.6.

Rf Circuits. Figure 8-37 shows all of the rf circuit types that were tested in the XC phase of the development program. They include a conventional, single-reentrant cavity; two reentrant cavities; both inductively coupled and uncoupled; and three versions of disk-loaded (capacitively coupled) extended-interaction circuits; four- and five-cell traveling-wave and four-cell standing-wave. The single and double inductively coupled reentrant cavities were limited by rf breakdown to a pulse length of less than 200 ns. The uncoupled reentrant cavities performed well, producing

86-MW peak power at 36% efficiency at 600 ns of pulse length. This design approach was not pursued further because of the necessity of combining the power from four output waveguides, with an uncertain impact on rf stability.

The disk-loaded, traveling-wave circuit showed the best pulse length capability at 1 μs with 51-MW peak output power. The low measured efficiency, about 24%, also demonstrated the need for a more sophisticated design approach. This was achieved using 2.5D computer simulations which incorporated circuit tapers in both impedance and velocity. The first disk-loaded, standing-wave circuit, the four-cell XC8, exhibited a destructive zero mode oscillation which was eliminated in the subsequent XL-1 and XL-2 klystrons. The XL-1 and XL-2 use three-cell, standing-wave circuits.

One important design criterion to emerge from the XC series of the development was the need for good axial symmetry in the output circuit, including output and cavity-to-cavity coupling. All of the circuits with single-coupling slots or single (one-sided) output waveguides broke down at pulse lengths below 200 ns, in general with asymmetrical cavity damage. Symmetrical, disk-loaded cavities with two balanced output ports, and uncoupled reentrant cavity pairs with four balanced output ports performed best. It appears that field asymmetries cause beam steering which becomes destructive in much less than the required 1.5 μs.

Rf Windows. The rf window development and refinement is the most nearly complete of the major development efforts associated with the X-band klystron. The use of the traveling-wave, TE_{01} mode window as contrasted with the TE_{10} standing-wave, modified pill box window of our early X-band klystrons has had a remarkable effect on power handling capability. In the resonant ring, the traveling-wave TE_{01} windows have demonstrated the capability of transmitting 100 MW of rf power at 1.5 μs pulse length. By contrast, breakdown in the TE_{10} mode windows has been observed at as little as 25-MW peak power at a pulse length less than half of that achieved with the TE_{01} window.

The development progression was from TE_{10} modified pill box to the TE_{01} standing-wave window, followed by the TE_{01} traveling-wave window and finally the latter window using a high-purity, isostatically-pressed window disk. Isostatically-pressed windows have yet to fail in the resonant ring, even at a peak power above 100 MW. One traveling-wave window of standard grade alumina was punctured at 100 MW. At the present time, the high-purity, isostatically-pressed window is limited by vacuum leaks through the brazed joint. This has been of sufficient concern that no isostatically-pressed window has been used in a tube. The development effort to obtain better metalizing is continuing. Meanwhile, the 100-MW maximum capability demonstrated by the standard grade alumina disk provides an adequate margin.

8.4.3 Ongoing R&D

The remaining R&D is of three types:

- Validating basic design concepts.
- Strengthening those areas where more performance margin could lead to longer life and higher reliability.
- Upgrading performance.

Validating Basic Design Concepts

Disk-Loaded, Traveling-Wave Output Circuit (XL-3). At this writing (February 1996), the solenoid-focused XL-4 with the four-cell, disk-loaded output circuit, has undergone preliminary testing. In computer simulation, this circuit

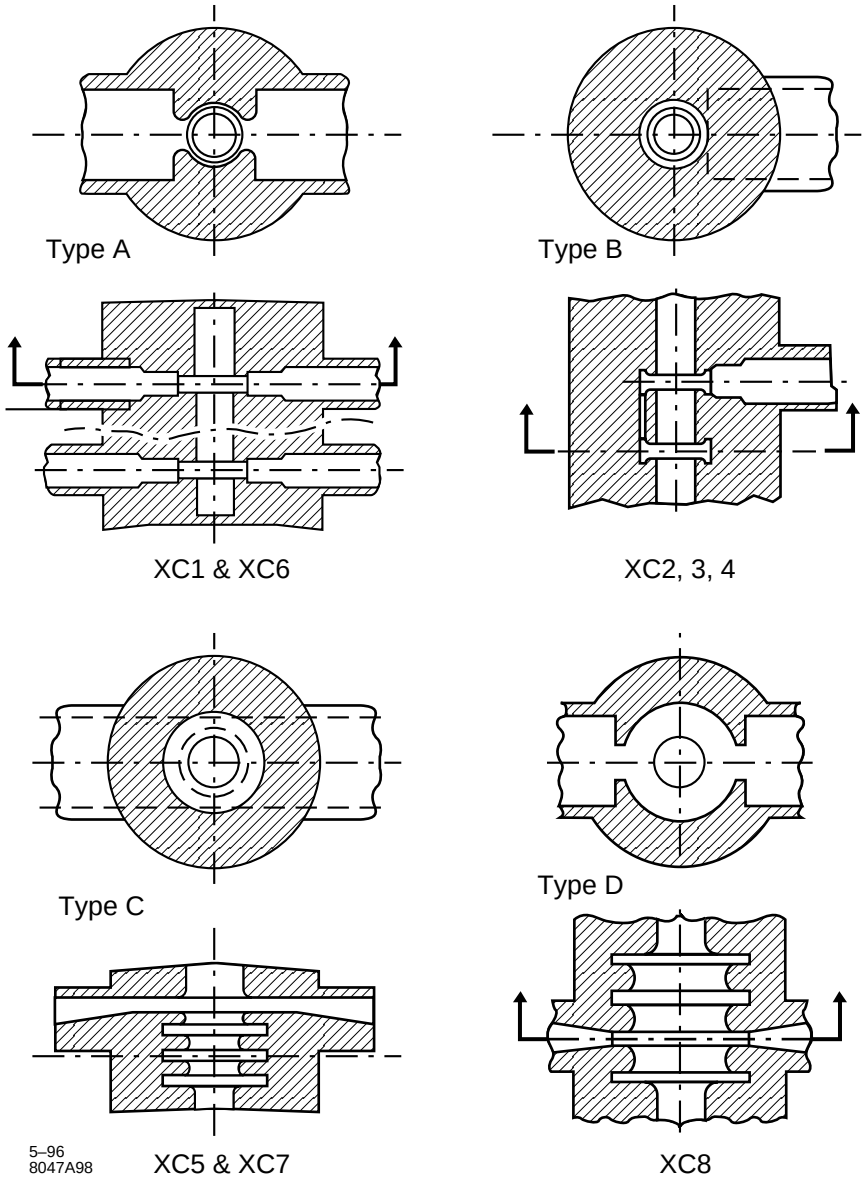


Figure 8-37. *Comparison of the rf output circuit types tested on the high-current XC series of NLC klystrons.*

showed the highest rf efficiency (52%) of the XL solenoid-focused series. This compares with a measured value of 48%.

PPM Focusing. Two PPM-focused test vehicles are currently under development. The first is the beam stick pictured in Figure 8-38. It is essentially a klystron without cavities. The beam stick is being used to test DC transmission (no rf) in two modes of operation:

- Brillouin Focusing (No flux threading the cathode)
- Partially Immersed Flow (50% flux threading the cathode)

All solenoid-focused, high-power klystrons use extensive cathode immersion (up to 95%) to achieve a high degree of beam stiffness to prevent rf driven beam interception. PPM simulations to date indicate that the combination of short magnet period which produces a contoured field, and the use of a profiled field (increase in amplitude with increased bunching) make it unnecessary to use partially immersed flow to further enhance beam stiffness. Because the simulations do not take into account a possible current halo around a real beam, we have also carried out PPM-focusing simulations, and the required design modifications to achieve beam flow with 50% cathode immersion.

The second test vehicle will be the X5011 klystron, which is designed to produce 50-MW output power at 465 kV with a beam microperveance of 0.6. Computer simulations of this tube predict an efficiency of 63% and no beam interception.

Performance Margin

Output Circuit RF Voltage Hold-off. An experimental program has been launched to identify the source of rf breakdown at relatively low-surface gradient which has been observed in all X-band klystron output circuits at SLAC. The output circuit breakdown occurs at about 20% of the rf surface gradient to which SLAC accelerator structures have been successfully subjected. At the present time, it is believed that the degradation in hold-off gradient is the result of the presence of the high-current electron beam, perhaps in conjunction with the solenoidal field which is used to confine the beam. The presence of the confining magnetic field alone has been ruled out as the cause, by initial tests of a cavity in the X-band resonant ring with and without magnetic field. A series of additional cavities is being tested in the resonant ring, to determine the effect on breakdown voltage of diamond-turned surfaces and more exotic surface cleaning procedures as well as the use of both sputter coatings and alternative cavity materials such as Elkonite (tungsten-copper) and perhaps Glidcop.

Use of Internal Loss. Most high-voltage klystrons display high-frequency anomalies in the form of observable oscillations at frequencies that are capable of propagating in the beam tunnel. The degree to which these occur appears to increase with increasing beam voltage. At SLAC, they became most pronounced in the first of two recently completed 535-kV, 150-MW, S-band klystrons developed for DESY in Germany. The problem was largely eliminated by substituting lossy stainless steel for copper in the beam drift tunnel in the second klystron. Similar oscillations, which limit the range of usable operating parameters, such as beam voltage and focusing field were encountered in the testing of XL-3. The objective of this R&D effort is to incorporate loss in the beam tunnel for the purpose of preventing, or at least damping out, such oscillations. The simple expedient of using stainless steel has been added to XL-4 in a manner similar to that of the second DESY klystron. It is too early in the test cycle to be sure that the modification has been completely successful.

The testing of a more invasive, but potentially more effective approach is planned. The initial test-bed for these experiments is the 65-MW, 5045, S-band klystron used in the SLC at SLAC. Lossy sleeves are to be inserted in a

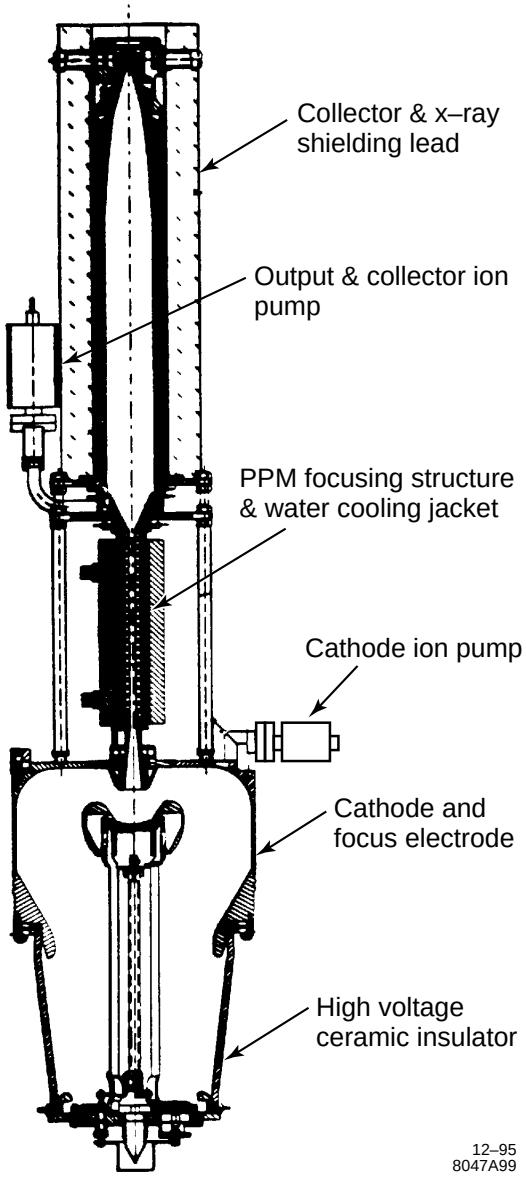


Figure 8-38. *PPM-focused "beam stick" which simulates the X5011 without rf drive. Transmission will be tested with and without flux through the cathode by means of separately tunable electromagnets (not shown).*

modified 5045 klystron to determine whether there are any deleterious effects, and whether any improvement can be detected. A pair of sleeves ground from single-crystal doped silicon having a resistivity between 1 and 4 Ω-cm have been procured. Should there be no deleterious effects, the experimental tube will be placed in a nonsensitive area in the accelerator gallery where many thousands of hours of testing can be observed. In the meantime, we will proceed with testing this increased oscillation damping loss in one or more future experimental X-band tubes.

Diffusion Bonding. Diffusion bonding of the individual cavities which make up the klystrons may result in a significant reduction of manufacturing cost. Diffusion bonding is somewhat complicated in the case of the klystron because of the need to bond spacers and pole pieces which are of different materials (monel or copper for the spacers and iron for the pole pieces). These materials also have different thermal coefficients of expansion. An experimental program, intended to develop the process technology (varying temperature, pressure and time) to obtain strong, reliable, vacuum-tight bonds has been launched. Diffusion bonding of a short stack of copper spacers and copper-plated, iron pole pieces of the type to be used on the PPM test diode has been carried out. This work has already demonstrated that vacuum-tight bonds can be made using this bonding technique.

Upgrading Performance

The upgrade to 1-TeV center-of-mass energy requires a 75-MW peak power klystron. The preliminary electrical and mechanical design of the 75-MW klystron, designated X7511, has been carried out. This includes rf simulations which show 60% operating efficiency. The criteria applied in the design were not to exceed the design limits imposed on the 50-MW klystron and, to the greatest extent possible, stay within the gradients and stress levels demonstrated in the X5011.

These restrictions ruled out simply increasing the beam voltage, which would exceed the self-imposed 500-kV limit while also increasing cathode current density and electron gun gradients to unacceptable levels. To stay within the 500-kV beam voltage limit, it was necessary to increase the beam microperveance from 0.6 to 0.75. To maintain cathode current density, the cathode diameter was increased from 5.72 cm to 6.75 cm. Finally, in order to maintain a realizable cathode area convergence and stay within the capability of existing permanent magnets, it was necessary to increase the diameter of the beam tunnel from 0.95 cm to 1.08 cm.

With these changes, cathode current density and gun gradients are approximately the same for the two tubes. Also, a common set of magnets will focus both beams. Based on 2.5D computer simulations, the rf efficiency drops by only 2 percentage points from 63% to 61%. Because of the reduction in repetition rate from 180 Hz to 120 Hz and a small decrease in pulse length, the average power required of the two klystrons is the same. The margin provided by the rf output window, which has demonstrated 100-MW peak power capability, remains adequate. Table 8-10 summarizes the design and performance parameters for the X7511 PPM-focused klystron.

8.4.4 Manufacturing

Because of the large numbers of klystrons required during the NLC construction phase, the cost of manufacturing and testing each tube is significant to the feasibility of the NLC project. Diffusion bonding is one example of possible cost reduction. The substitution of a low-cost oxide cathode for the present dispenser cathode is another possible cost reduction. After completion of the NLC, monthly tube replacement during operation is estimated to be at least 15 times that of the SLC. On-site production and test capability must be considered in detail. Automation of klystron construction and testing, as well as the accelerator structure construction and testing, is key to the economics of the overall project.

Parameter	Units	Value
Electrical:		
Frequency	(GHz)	11.424
Peak Power	(MW)	75
Beam Voltage	(kV)	490
Beam Current	(A)	255
Beam Microperveance	(μA/V$^{3/2}$)	0.75
Gain	(dB)	57
Bandwidth	(MHz)	100
Beam Modulation		Cathode Pulsed
Pulse Duration	(μs)	1.0
Rep Rate	(Hz)	120 (180 for NLCTA)
RF Efficiency	(%)	60
Cathode Current Density	(A/cm^2)	7.4
Heater Voltage	(V)	15
Heater Current	(A)	27
Mechanical and Magnetic:		
Beam Focusing		Periodic Permanent Magnet (PPM)
Tube Weight	(kg)	32 (without lead)
Tube Length	(m)	1.3
Cathode Diameter	(cm)	6.75
Cathode Half Angle	($^\circ$)	29
Beam Area Convergence		124

Table 8-10. *Design and performance parameters for the X7511 klystron.*

An "Alliance for the Advancement of Robotics Technology" (AART) has been set up among several DOE National Laboratories which seeks to foster collaborative research ties between these organizations. The AART has agreed that its first project is to study the concept of building an NLC klystron factory relying on robotic machining, cleaning, inspection, assembly, brazing, processing and testing of klystrons. With funding of the study provided by both the DOE and DOD, the work will include a complete design and cost analysis.

8.5 Klystron Pulse Modulator

Each high-power rf station includes a pulsed energy delivery system that is tightly integrated in design with the beam generation section of the klystron, and a high-efficiency power supply for charging the PFN (pulse forming network) capacitance. A preliminary design is shown in the diagram in Figure 8-39.

The subsystems of each rf station, discussed in this section, include a single-thyratron Blumlein modulator (Section 8.5.2) driving two PPM-focused klystrons, the modulator power supply (Section 8.5.3), and the station-cooling and oil-circulation systems (Section 8.5.4). The protection systems will be discussed in Section 8.7. All of these subsystems are closely interlinked.

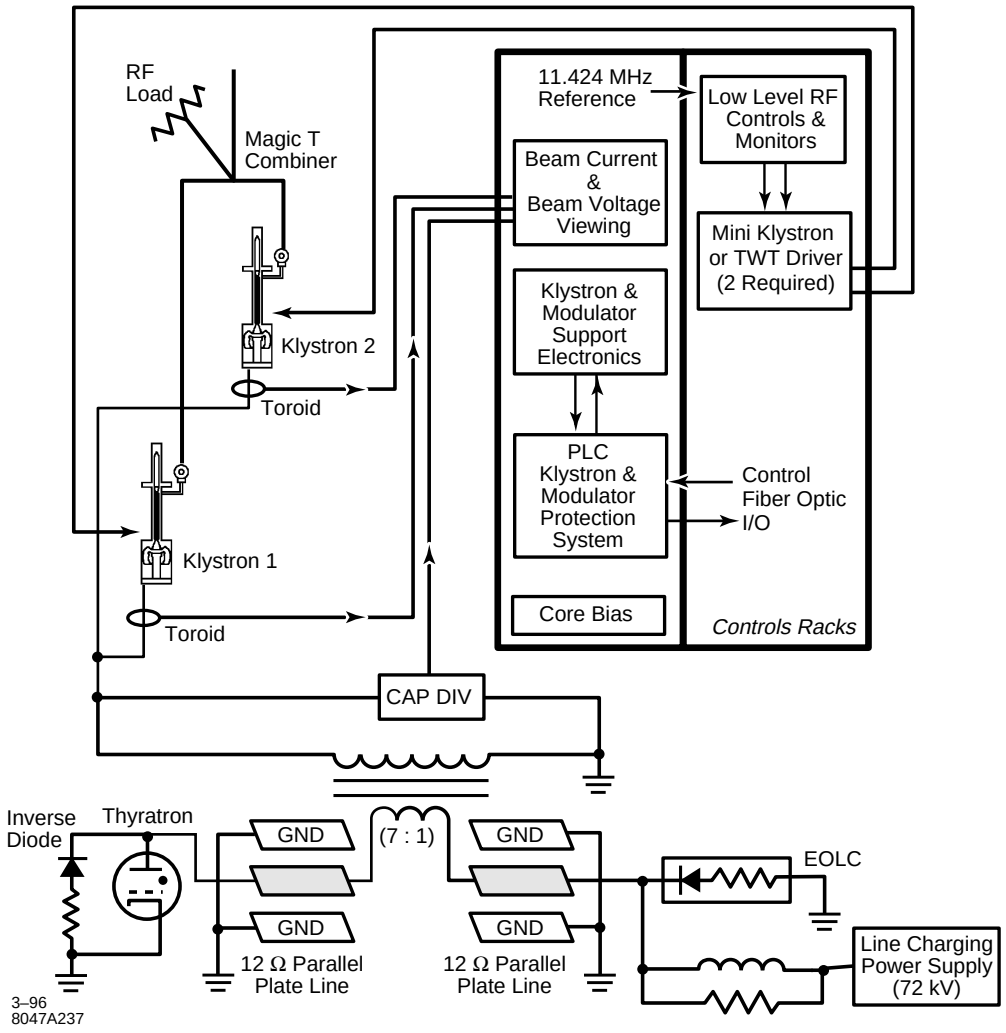


Figure 8-39. *NLC RF Station Block Diagram*

There is much basic research and development that must be accomplished to characterize a final system that will maximize efficiency at minimal cost, and will be reliable and serviceable. Some elements of a program for developing prototype modulator components are outlined in Section 8.5.6. In the meantime, modulator performance can be estimated through simulations. These simulations are based on the best knowledge available for modulator components, in particular the pulse transformer. Simulation examples are given in Section 8.5.5; these are used to obtain the efficiency estimate for the NLC modulator design goals as shown in Table 8-11.

8.5.1 Modulator Requirements

Specifications for the NLC modulator are given in Table 8-11. These specifications are driven by the requirements of the NLC klystron, as listed in the first part of the table. Although the initial design goal for this klystron is an

output power of 50 MW, as required by the 500-GeV NLC design, an improvement program is in progress which will eventually increase the power output to the 75 MW required for the 1-TeV upgrade. The modulator must therefore be capable of driving two klystrons at this higher power level. The repetition rate for the initial 500-GeV NLC design is 180 Hz. To save AC power, this drops to 120 Hz for the later 1-TeV machine. However, during the transition to the higher energy, there may be a mix of 50-MW and 75-MW klystrons on the linac, operating at 180 Hz. In fact, the higher power tube may be available even at the time construction begins on the 500-GeV collider. In either case, the modulator must be capable of driving the higher-power tube at the 180-Hz repetition rate.

The first two columns in Table 8-11 list the design goals for the 50-MW and 75-MW klystrons, and a corresponding set of parameters for a modulator capable of driving two such klystrons. However, as mentioned above, the 75-MW klystron and the modulator must also be capable of operating at a 180-Hz repetition rate, as reflected in the third column of Table 8-11. The efficiency of the 75-MW klystron may also fall short of the 60% efficiency goal, at least initially. To be conservative, the modulator parameters have therefore been chosen so that the modulator can drive a klystron with an efficiency as low as 55% to the required 75-MW output power.

A key consideration for a modulator in a linear-collider application is the efficiency with which power is transferred from the AC line to the klystron beam power in the flat-top portion of the high-voltage output pulse. A 1% decrease in efficiency at any point in this efficiency chain results in an increase by between one and two megawatts in the AC line power required for the NLC rf system. An important component in this efficiency is the ratio of the useful energy in the flat-top portion of the pulse to the total pulse energy, including the energy in the rise and falltime portions of the pulse. This rise- and fall-time energy efficiency is given (very roughly) by the ratio $\mathrm{T}_K/(\mathrm{T}_K + 1.1\mathrm{T}_R)$, where T_K is the flat-top pulse width and T_R is the rise time (the constant 1.1 depends on the precise definition of rise time). In turn, the rise time is determined in large part by the physical design of the pulse transformer, and in particular by the leakage inductance (due to imperfect coupling between the primary and secondary windings). In practical pulse transformer designs, it is observed that the leakage inductance tends to decrease as the transformer turns ratio is decreased. However, a lower transformer turns ratio implies a higher value for the charging voltage for the pulse-forming network (PFN), and a correspondingly high hold-off voltage for the thyratron and other high-voltage components. For a modulator using a standard PFN design which can deliver a 500-kV output pulse, this charging voltage would be about 145 kV for the 7:1 pulse-transformer turns ratio listed in Table 8-11. However, by using a Blumlein-type of PFN, the charging voltage can be reduced by a factor of two to about 72 kV. Although a Blumlein PFN design is somewhat more complex, it makes possible both a low pulse-transformer turns ratio and a reasonable value for the charging voltage.

In a standard modulator design with output pulse lengths longer than a microsecond or so, lumped elements (capacitors and inductors) are used for the PFN. Such a lumped network has the advantage that it can be readily tuned to adjust the pulse shape, although it is still difficult to eliminate all of the ripples to attain a truly flat pulse unless a very large number of elements are used. Also, it is difficult and expensive to manufacture pulse capacitors with a very low series inductance and long life, especially when the polarity must reverse during the pulse, as required by the Blumlein configuration. For these reasons, the use of lengths of smooth transmission line (termed here a pulse-forming line, or PFL) is being proposed for the NLC modulator. The major disadvantage of such lines (long length) can be ameliorated by choosing an appropriate packing geometry and good mechanical design (Section 8.5.2).

The rise- and fall-time efficiency is the major component which determines the overall modulator efficiency. In addition, a voltage drop across the thyratron, eddy currents and hysteresis losses in the transformer core, and resistive losses in the transformer windings and connecting leads contribute an additional loss of several percent. The charging voltage on the PFL must be slightly higher (about 1.5%) to compensate for these losses. The corresponding loss in efficiency (97%), multiplied by the rise- and fall-time efficiency, gives the net efficiency with which energy stored in the capacitance of the PFL is transferred through the pulse transformer into the useful flat-top portion of the output pulse. A discussion of the projected efficiency for the NLC modulator design, based on simulations of the output pulse shape is given in Section 8.5.5.

Description: Blumlein PFL configuration with individual line-charging power supply and 7:1 pulse transformer turns ratio.			
	50-MW Klystrons	75-MW Klystrons	Design Min/Max
Klystron Parameters:			
Klystron Peak Output Power (MW)	50	72	
Klystron Microperveance (μperv)	0.60	0.75	
Klystron Beam Voltage (kV)	465	490	
Klystron Efficiency (%)	57	60	55 Min
RF Pulse Width (μs)	1.2	1.0	
Repetition Rate (Hz)	180	120	180 Max
Modulator Parameters:			
Output Pulse Voltage (kV)	465	490	510 Max
Output Pulse Current, for two klystrons (A)	380	510	550 Max
Flat Top Pulse Width (μs)	1.2	1.0	
Repetition Rate (Hz)	180	120	180 Max
Pulse Flatness (%)	1	1	
PFL Voltage (kV)	68	71	75 Max
Est. Energy Transfer Efficiency[a] (%)	77.5	80.5	76 Min
PFL Stored Energy (J)	274	310	370 Max
Pulse Energy Width (μs)	1.5	1.2	
Pulse Rise Time (ns)	275	175	
Thyratron Current (kA)	5.3	7.15	7.7 Max
PFL Impedance[b] (Ω)	12.4	9.9	9.7 Min
Auxiliary AC Power[c] (kW)	2.2	2.3	2.5 Max
Est. Power Supply Efficiency (%)	93	93	92 Min
Net Modulator Efficiency (%) (excluding auxiliary power)	72	75	70 Min
AC Input Power (kW) (excluding auxiliary power)	53	40	72 Max

[a] See Section 8.5.5. The energy-transfer efficiency is given by the rise- and fall-time energy efficiency times a series loss efficiency (97% is assumed here).

[b] Adjusted to match impedance of production klystrons.

[c] Includes thyratron heater, thyratron reservoir, and klystron cathode heater supply.

Table 8-11. *Klystron-modulator specifications (two klystrons per modulator).*

8.5.2 Pulse Modulator Design Outline

The pulse modulator described in the section below is presented as an example of a design that is physically realizable with present design technology or modest extensions to that technology. In Section 8.5.6, an R&D program is outlined whose results will undoubtedly change this design.

The station klystron modulator drives two PPM-focused klystrons. The modulator is of the Blumlein type, and uses as its energy storage elements distributed-type transmission lines rather than lumped element artificial lines. The design is driven by the need to synthesize the klystron cathode pulse with the highest efficiency consistent with reasonable cost. In a modulator, the major areas where energy is lost are in the rise and fall times of the cathode pulse where energy is dissipated, but no useful rf energy is produced, and in the IR drops and capacitive and inductive stored energy that is dumped after each pulse ends. There is also the power lost in the thyratron voltage drop, and the energy necessary to operate the thyratron heater and reservoir.

In a thyratron-PFN modulator, the element determining the rise time is usually the pulse transformer. Because of the voltage hold-off necessary between the primary and secondary of the transformer, there is always leakage flux generated by the primary that does not couple to the secondary. This shows up as leakage inductance in series with the primary which limits the voltage rise time of the secondary. There is a minimum stray capacity associated with the klystron cathode, and the high-voltage secondary of the pulse transformer. This capacity must be charged on each pulse, and the charging energy is lost during each pulse fall time. The transformer core does not have infinite magnetic permeability, and so a real inductance which is not infinite appears in parallel with the primary of the transformer. This has the effect of lowering the load impedance as a function of pulse length, leading to droop at the tail end of the pulse. The transformer leakage inductance can be minimized by keeping the pulse transformer turns ratio as low as possible. A low ratio dictates a high-primary pulse voltage, and the primary voltage is limited by the switching capability of the thyratron and the voltage holdoff of the pulse forming line.

In the standard-design modulator, a single PFN switched by a single thyratron drives a high-ratio (typically 23:1) pulse transformer. A modulator using this conventional design was optimized for efficiency and minimum rise time in the Test Stand 13 position in the Klystron Test Lab at SLAC. The rise time was less than 600 ns. This is typical of conventional design technology. A modulator design in which two PFNs are used, charged in parallel, and discharged in series (called the Blumlein design) allows the use of a low ratio (7:1) pulse transformer and a thyratron with reasonable voltage hold-off capability. In a lumped element PFN, internal inductance in the capacitors limits the shortness of the rise time that can be obtained. If a distributed transmission line is used instead of the LC lumped line, this limit can be circumvented. Figure 8-40 shows the simple Blumlein modulator concept using two 10-Ω distributed lines, a 75-kV thyratron, and pulse transformer with a 7:1 turns ratio and a primary impedance of 20 Ω. This circuit can provide a pulse rise time of less than 300 ns, delivering a 490-kV pulse of about 500 A to drive two 75-MW klystrons.

The Blumlein pulser circuit as shown in Figure 8-40 operates as follows:

1. A capacitor-charging power supply charges both inner plates of the two lines to a DC potential of +72 kV. The two inner plates and the primary of the pulse transformer are now at a +72-kV potential.

2. The thyratron is triggered and presents a short circuit to the left end of the left transmission line. The 10-Ω transmission line presents to the thyratron a voltage source of 72 kV in series with an internal impedance (line characteristic impedance) of 10 Ω. The current in the thyratron is thus 7.2 kA. In order to match the current and voltage conditions at the thyratron after switching, a -72-kV, 7.2-kA wave must be launched traveling to the right on the transmission line. After a time T (where T is the electrical length of the line), this wavefront reaches the end of the line at the connection to the pulse transformer. There it sees a load impedance of 30 Ω (the pulse transformer primary in series with the 10-Ω impedance of the second line). From transmission line theory,

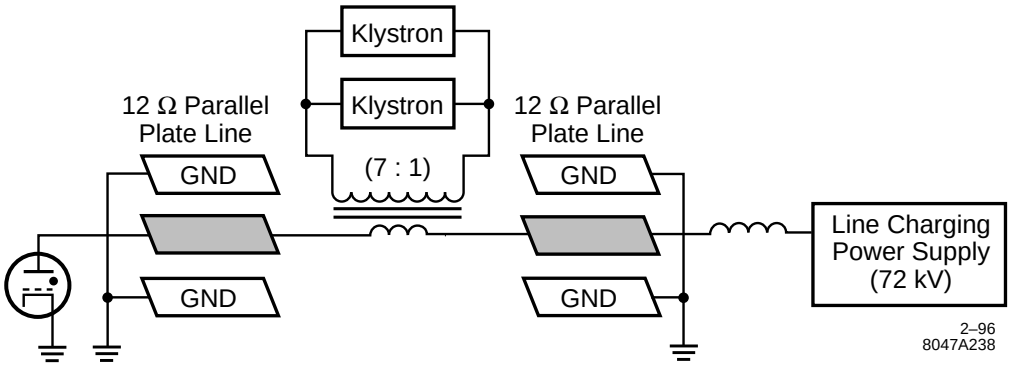


Figure 8-40. *NLC modulator Blumlein schematic.*

this mismatch launches a reflected wave of -36 kV, which travels back toward thyratron carrying a current of 3.6 kA.

3. The remaining current in the initial wave ($7.2\,\text{kA} - 3.6\,\text{kA} = 3.6\,\text{kA}$) will start to flow through the primary of the pulse transformer at time $t = T$. It will also launch a wave with an amplitude of -36 kV traveling to the right in the second transmission line. The net voltage at the right side of the transformer primary is $72\,\text{kV} - 36\,\text{kV} = +36\,\text{kV}$. The net voltage drop across the primary is therefore 72 kV (also given by $3.6\,\text{kA} \times 20\,\Omega$). The transformer output voltage is this voltage drop times the 7:1 turns ratio, or about 500 kV.

4. The two -36-kV waves, traveling to the left on the first line and to the right on the second line, leave behind them voltages of -36 kV and $+36$ kV, respectively. Note that the polarity has reversed on the first line; this reversal imposes an additional burden on the design of pulse capacitors for Blumlein networks.

5. After a second period T, the two waves reach the left and right ends of their respective transmission lines. At that time a reflected wave of $+36$ kV (short circuit termination) is launched in the first line, and a -36-kV wave (open circuit termination) is launched in the second line. These waves travel back toward the pulse transformer during a third period T, and wipe out all the remaining energy stored in the two lines. At time $t = 3T$, the waves reach the pulse transformer and cancel the current flowing in the primary.

Distributed energy storage lines are most familiar as coaxial cables. There are many low-power applications where coaxial cables are used for smooth-pulse energy discharge. At high power, there is a good technology development in water-filled pulse lines for very short pulses at low impedance. These water-filled energy storage lines make use of the very high dielectric constant of pure water, $\epsilon_r > 80$, but they must be charged very rapidly because water as a dielectric cannot withstand high electric fields for more than tens of microseconds before it becomes lossy. An oil-filled line with high dielectric oil ($\epsilon_r = 7$) can in principle hold off a high voltage, but the support structure of the inner coaxial element is subject to creepage breakdown. For the NLC application, physically realizing a 10-Ω distributed line as a coaxial cable requires a very large diameter cable, since high-voltage creepage across the inner support insulator would otherwise limit the charging voltage. Any breakdown across the inner support insulator would be difficult to repair or replace.

An oil-filled five-layer parallel plate transmission line as shown in Figure 8-41, is physically realizable and relatively easy to manufacture and repair. The structure as shown contains two lines of 10-Ω impedance each. The dielectric constant of the oil is just that of normal transformer oil, 2.6. The oil containment for the transmission line is shown as separated from the klystron and pulse transformer oil so that high-dielectric ($\epsilon_r = 7$) oil can be considered in the design development or upgrade. Note that with normal transformer oil, $\epsilon_r = 2.6$, the transmission lines shown match klystrons of 0.6 μperv. By increasing just the dielectric constant of the oil, lower impedance lines result that can match

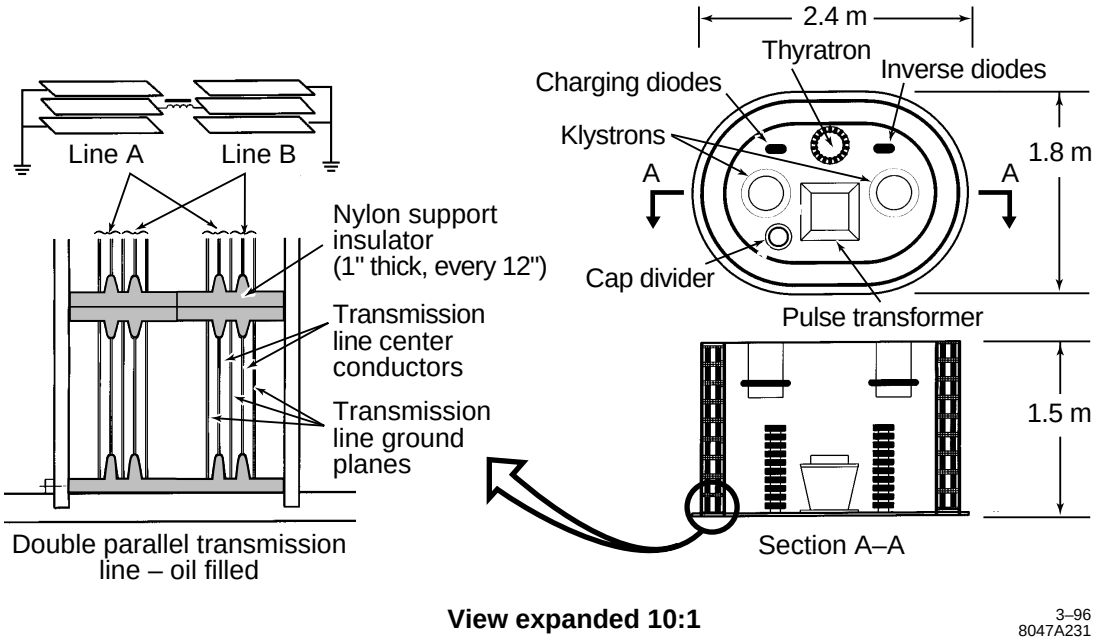


Figure 8-41. *NLC modulator tank assembly showing two klystrons, thyratron, pulse transformer, and two multilayer plate transmission lines.*

higher perveance klystrons. These lines would also be electrically longer by the square root of the dielectric constant ratio.

A proposed modulator tank supporting two klystrons is shown in Figure 8-41. The parallel plate transmission lines are shown mounted in the space between the inner oil tank and the outer containment tank. The total length of this double transmission line assembly is about 460 ft. The oil is contained in an aluminum racetrack-shaped inner tank that also serves as the support for the transmission line assembly. The transmission line assembly consists of continuous aluminum strips supported by molded nylon support insulators approximately one-inch-thick spaced at 12-in intervals. This construction allows easy servicing and repair of the pulse line if an arc destroys one of the support insulators. By supporting the parallel plates from the end of the aluminum strips, the creepage path between the plates can be made long enough to prevent discharge while still keeping the plate-to-plate gap small enough to take advantage of the greater breakdown properties of the oil.

There are a number of secondary elements also housed in the modulator tank: the end-of-line clipper diode stack, the charging diode (if one is used), the current viewing transformers and voltage viewing capacity divider, the core reset inductor, and the thyratron inverse clipper diode. All of these elements are contained in the inner tank adjacent to the thyratron and pulse transformer.

A double-rack enclosure adjacent to the modulator tank contains all the support electronics for the modulator, and the rf-drive electronics. The support electronics includes the klystron cathode heater supply, the thyratron cathode supply, the thyratron reservoir supply, the control power distribution, the rf low-level driver and interlock system, and a custom designed programmable logic controller (PLC) that provides the control and monitor functions for both the pulse modulator and the klystron rf drive and protection systems. Depending on the construction design, this set of racks may also contain the primary capacitor charging power supply.

Charging Voltage	72 kV
Pulse Forming Line Capacitance	0.12μF
Joules/pulse, 2 klystrons	310
Repetition Rate	180 pps at 500 GeV; 120 pps at 1 TeV
Charging Voltage Regulation	0.1%
AC Line	480 V, 3-phase, 60 Hz
AC Line Stability	3%
Power Supply Efficiency	93%

Table 8-12. *Charging power supply specifications.*

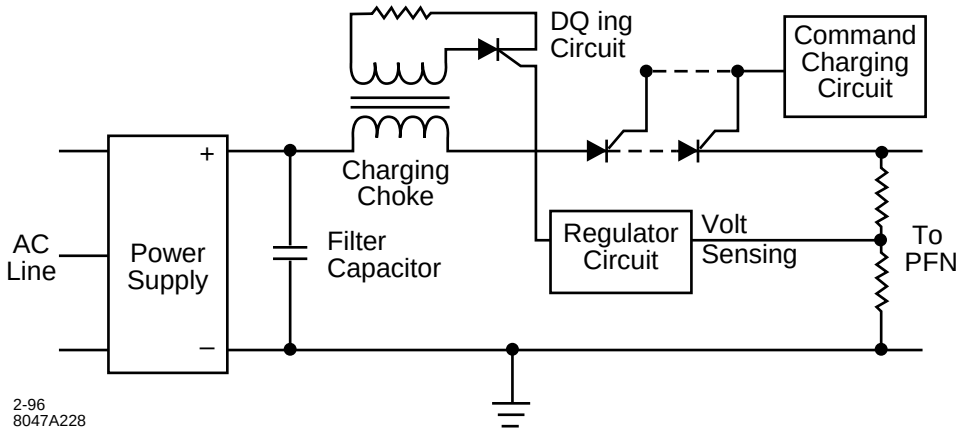


Figure 8-42. *Conventional PFN charging power supply.*

8.5.3 Charging Power Supply Design Outline

Specifications for 500-GeV Operation

The power supply that charges the PFL must fulfill the requirements shown in Table 8-12.

Conventional Charging System

Traditionally, the power supply that charges the PFL capacitance consists of an AC to DC power supply, a filter capacitor bank, a charging choke, a charging diode or a command-charging SCR circuit, and de$Q$ing circuit that regulates the PFN charging voltage (Figure 8-42).

If we assume an AC line stability of 3%, then we must de$Q$ at least 6% to meet the regulation requirements. If the de$Q$ing power is dissipated, the efficiency of the power supply charging system is diminished. An energy recovery scheme could be developed to recover the energy and feed it back to the power lines, but it would be expensive and would only recover a portion of the energy. With energy recovery, we could only achieve a total power supply efficiency of about 90%. The output voltage could also be regulated by using a primary SCR line regulator instead of the de$Q$ing circuit.

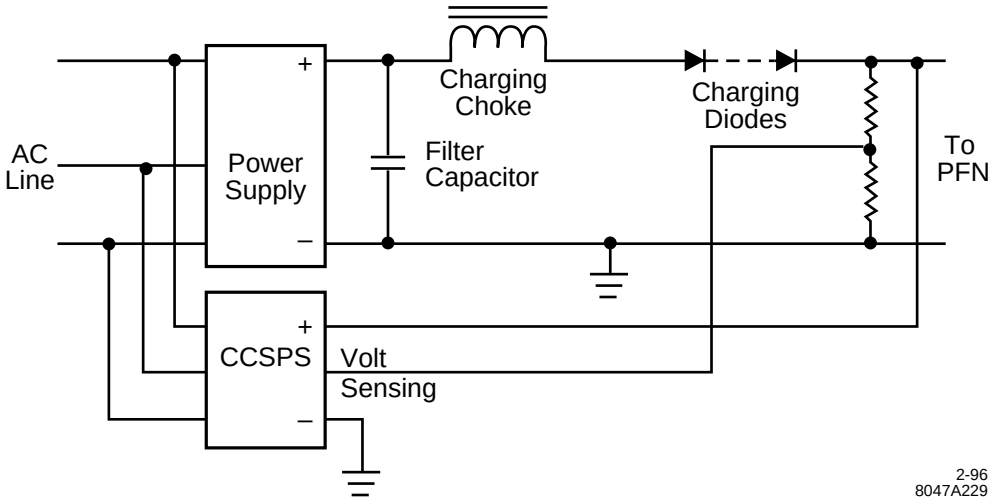


Figure 8-43. *Hybrid circuit consisting of a conventional power supply and a small capacitor charging switching power supply (CCSPS) in parallel.*

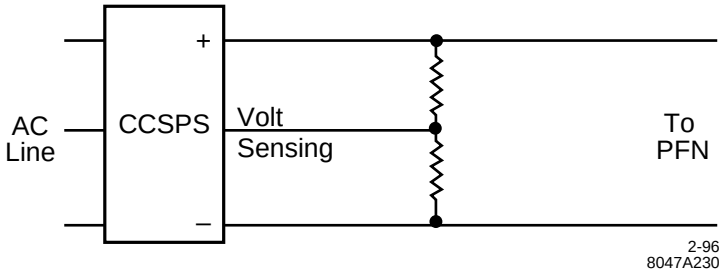


Figure 8-44. *Capacitor charging switching power supply (CCSPS) scheme.*

Hybrid Charging System

A better scheme would be to use a small capacitor charging switching power supply (CCSPS) in parallel with the charging choke output (Figure 8-43). At the lowest line voltage, the PFN would then be charged to 94% of its peak voltage by the charging choke, and the switching power supply would be used to charge the final 6% of the voltage. What we have done is substitute an additive process for a subtractive one with a resulting increase in efficiency. With this additive scheme, power supply efficiency could be as high as 93%.

Capacitor Charging Switching Power Supply System

An even more attractive idea is to use a CCSPS to charge the PFN directly (Figure 8-44). The charging cycle starts with the PFL capacitor at zero volts. The CCSPS starts charging the PFL at constant current until the desired voltage is reached. At that time, the supply becomes a constant-voltage power supply and keeps the PFL charged at the desired value. The charging waveform looks like a linear ramp which flattens out at the desired voltage.

The CCSPS supplies are becoming commercially available now and will be common in the near future. A single 50-kV power supply is presently available which will charge the PFL to 50 kV at 120 pps, will tolerate a voltage reversal of

20%, and has an efficiency of approximately 93%. Research is required to increase the charging voltage to a minimum value of 72 kV. To be able to charge at 180 pps, two CCSPSs, operated in parallel using master/slave connections, are required.

This solution will greatly simplify the charging circuit topology since there is no need for a filter capacitor bank, charging choke, command charging circuit, and high-voltage blocking SCR strings. The elimination of the filter capacitor bank reduces the amount of energy storage and the possibility of thyratron damage in a latch-up condition (when the thyratron fails to turn off, or breaks down during PFL charging). With a filter capacitor bank, a false trigger or a thyratron breakdown will cause the filter bank to discharge completely through the thyratron, and the power supply will then short circuit itself through the thyratron. With the CCSPS, a thyratron breakdown would only result in a maximum current equal to the nominal charging current since the CCSPS reverts to constant current operation when it is short circuited. As an additional advantage, the CCSPS is much more compact than a conventional power supply. For example, the 50-kV supply cited is rack mountable, $19''$ wide, $22''$ deep, and only $12.25''$ high.

Other manufacturers have built power supplies of this power level, housed in a single chassis and having an efficiency of approximately 85%. Additional development is needed to increase the efficiency and to develop a reliable 75-kV supply.

8.5.4 Station Cooling System and Oil Circulation

There are several cooling circuits that remove heat from the klystron and modulator. The klystron has just one water cooling circuit that is a combination of all the passages in the klystron. Temperature monitor points on parts of the klystron (body, anode, collector, etc.) provide interlocking for over-temperature conditions. The modulator contains two oil systems, one for the transmission lines oil which may have a higher dielectric than the main tank, and the other for the low-dielectric oil for the klystron guns, the thyratron, the end-of-line clipper (EOLC) and charging diodes, and the pulse transformer. Each of these systems will have a small oil-circulation pump and an oil-to-water heat exchanger appropriately interlocked.

There is no cooling required in the rf system support racks unless the charging power supply is mounted in these racks and it needs water cooling. There may be a need for water temperature stabilization if the rf components cannot be made to be phase-stable with temperature changes.

8.5.5 Simulations and Efficiency Projections

At the time of this writing, the efficiencies of the various components and subsystems are being analyzed, and the numbers will become clearer as R&D progresses. Pulse response simulations based on projected parameters of various lines, discrete capacitors, and pulse transformers are shown in this section.

Distributed Blumlein

There have been several preliminary network simulations run, using the program ISPICE, which show the general waveshapes on a Blumlein circuit using one or several segments of transmission line, as shown in Figure 8-45. By using several segments of varying impedance, the voltage at the tail of the pulse can be boosted to compensate for transformer droop. A simulated pulse shape is shown in Figure 8-46 for the case of a modulator driving a 50-MW

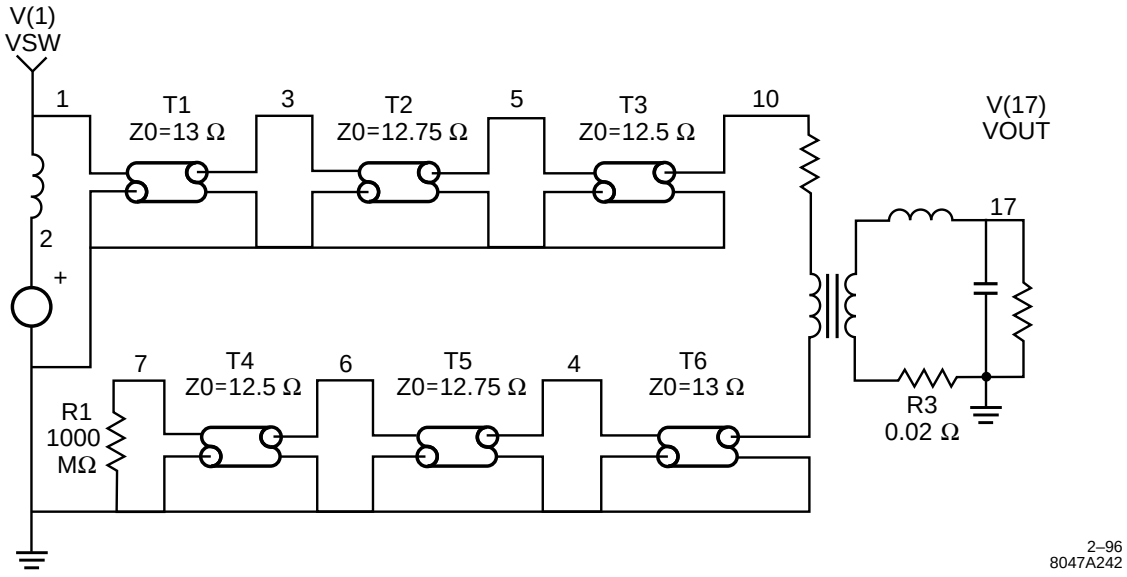


Figure 8-45. *Example of a three-section Blumlein circuit.*

klystron at 465 kV. The flat-top pulse length is about 1.4 μs, the rise time is about 250 ns, the energy efficiency is about 83% and the peak-to-peak ripple about 2%.

More detailed simulations are planned to look at waveshapes using more complex equivalent circuits that approximate real lines and pulse transformers. All the secondary parameters such as klystron stray capacity, the real characteristics of the various diode stacks and monitor elements, and the actual characteristic of a switched thyratron will be added to the model to get a better picture of what the real pulse response will be.

The dielectric constant of the various normal and high-dielectric oils must be characterized as a function of frequency, including dielectric losses, since the propagation of the wavefront on the lines will be limited by these parameters. High frequency losses limit the wavefront rise time, and dielectric constant variations with frequency can smear out the rise time, and cause pulse overshoot if all frequency components of the wavefront do not arrive at the end of the line at the proper times. It will take more development work with a simulation program to allow the dielectric constant to have a real and imaginary part corresponding to loss and propagation velocity.

Discrete Element Blumlein

While smooth distributed transmission lines are attractive substitutes for lumped element lines with discrete capacitors, there are some advantages to using lumped elements, especially if the capacitors can have high-energy storage density, low inductance, and easy manufacturability leading to low cost. Strontium Titanate (ceramic type) capacitors have high-energy storage density, and because they can be automatically manufactured, can have very low per-unit cost. A lumped element line using as many as two hundred elements can approximate the characteristics of a smooth oil-filled line, but can be packaged into a much smaller volume because of the high energy storage density in the Strontium Titanate dielectric. Much R&D will be necessary to physically realize such a line and to produce low-inductance capacitors.

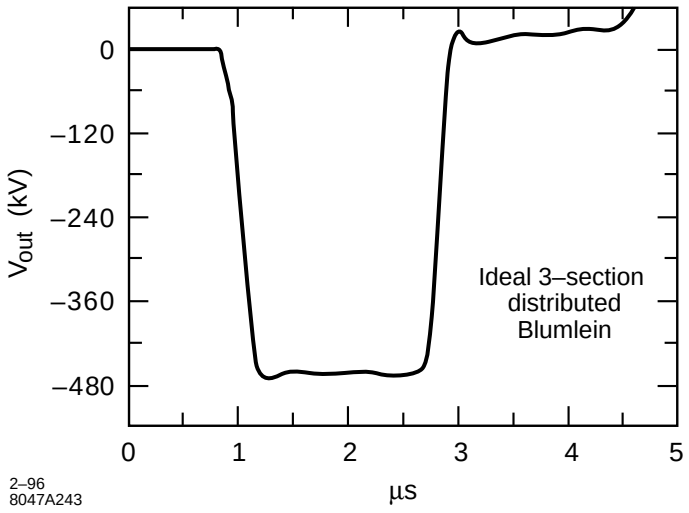


Figure 8-46. *Voltage pulse obtained from simulation of an ideal three-section distributed Blumlein.*

8.5.6 Prototype Modulator Development and Performance

In the Klystron Test Lab at SLAC, there are two modulators which can give some experimental check points to compare with simulations. Test Stand 13 is a high-power conventional modulator which contains close-coupled PFNs and a direct connection to a 23:1 pulse transformer. It produces 550-kV pulses driving a 1.8 microperveance klystron for a 3-μs flat-top pulse.

Test Stand 3 contains a three-section Darlington line modulator driving a 6:1 pulse transformer with all primary pulse-forming elements in one large oil tank. It also is designed to produce 600-kV pulses, but it has not run at full voltage in oil as yet. The capacitors that make up the three PFNs have too much series internal inductance to allow a very fast rise time, flat-top pulse. Additional testing will be done on this modulator to get better data. This test stand will be used as the location for new R&D work on distributed and lumped element pulse line modulators.

As the paper design of a possible NLC modulator evolves, the R&D areas where the new ideas must be tested start to surface. At the present time, we can divide the R&D effort into two categories: System concept development, and component and subsystem design and testing.

System Concept Development

Conventional klystron beam-delivery pulse-modulator systems proceed in design with microwave-oriented klystron engineers developing the klystrons, and electronics-oriented modulator engineers designing modulators to meet klystron specifications. While this cooperative effort can sometimes produce well-coordinated results, more often performance, cost, and efficiency savings are lost in the interface process. By considering the klystron beam formation design as an integrated system with the pulse transformer and primary pulsed energy storage system, trade-offs can be made at an early stage in the design development resulting in a better optimized system.

New concepts can be considered when the system is an integrated whole. As an example, the klystron interface with the pulsed high voltage is traditionally a large, cylindrical ceramic interfacing the high vacuum of the klystron to the oil insulation of the pulse transformer tank. Consider the advantages of a pulse tank under high vacuum instead of

being filled with oil. With vacuum on both sides of the klystron cathode seal, the seal can be redesigned to be smaller, radial rather than cylindrical, and much easier to manufacture. The capacity of the gun structure decreases because of the smaller cathode system, and the absence of high dielectric oil. Inside the tank, the pulse transformer size might be reduced because of the better insulating properties of high vacuum. This smaller size, together with the lower dielectric constant, would reduce the distributed capacitance between windings and other stray capacitances, and possibly also the leak inductance. All these reductions in stray energy storage work to decrease the rise and fall times of the cathode pulse, the whole system becomes more efficient.

To physically realize such a system, new R&D on a variety of components and subsystems must be conducted. These efforts are included in the proposed R&D program outlined below.

Component and Subsystem Design and Testing

In the component and subsystem area, R&D programs are needed to develop both oil-immersed and high-vacuum systems. Efficient charging supplies are a separate area of R&D where cooperation with industry should produce reliable, high-efficiency units that can be adapted for large-scale production. Some of the pulsed energy storage and delivery development efforts are as follows:

Thyratrons: The present design of two- and three-gap thyratrons is well advanced and with minor modifications can meet the needs of the NLC pulsed power switch. Long cathode life, on the order of 50,000 hours, has been achieved with a large dispenser cathode and increased-size hydrogen reservoirs. The R&D effort for thyratrons includes mechanical packaging for low-cost fabrication and quick changing in the field, and the reduction of power needed to heat the cathode and hydrogen reservoirs. This can be done as a joint effort with industry. We are fortunate in having at least three interested thyratron manufacturers with the skills and facilities to enter this development and optimization process.

Distributed Energy Storage Lines: Several different energy storage lines have been suggested, including oil-filled Andrews cable, either coaxial or triaxial, and the parallel plate open network described earlier. The oil dielectric-medium used in these lines requires further study to understand the breakdown properties and the dispersive characteristics of the various oils and other possible dielectric liquids.

The lines themselves need further detailed mechanical and electrical design analysis, including building some full-voltage test models to see how these lines perform. The manufacturability, fabrication cost, and maintainability of such lines must also be studied.

Lumped Element Quasi-distributed Storage Lines: Strontium Titanate (ceramic) high-voltage capacitors have very large energy storage density and in quantity are simple to manufacture. The temperature coefficient is rather large, but in a temperature stabilized oil environment, this drawback is not severe. A quasi-distributed energy storage line can be physically realized with as many as 100 elements which looks much like a distributed line, but is much smaller. R&D would be conducted on examples of this type of line to establish the feasibility of using it for the NLC pulsed power storage.

Pulse Transformers: Since all the pulsed energy storage systems envisioned for the NLC operate below 100 kV, a pulse transformer is required to transform the low-impedance line pulse to the higher impedance 500-kV level needed by the beam-forming elements of the klystron. Traditionally, this pulse transformer has operated in oil for insulation. The core material has been 2-mil tape-wound iron. Improvements can be made in pulse transformers both in the winding of the primary and secondary of the transformer, the use of oil versus vacuum insulation in the windings, and in the properties of the core material. This area is central to the R&D effort to produce a more efficient pulsed energy delivery system.

An oil-immersed transformer is fairly well understood. A transformer using high vacuum as the insulation medium is a challenge in that the materials used in the transformer and core are not usually good vacuum elements. However, with some innovative design both for the core and the windings, and for a bare ion pump contained within the pulse tank vacuum enclosure, a significantly improved pulse transformer and klystron cathode high-voltage seal can be physically realized. The resulting klystron and pulse transformer tank package would be much smaller and lighter than the present 5045 klystron assembly, and would look more like the early-generation, SLAC XK-5 klystron, without the barrel permanent magnet.

Primary Materials Developments: Improvements in the understanding of performance and limitations of basic materials such as insulators, oil and general dielectric fluids, and core materials such as iron, nickel alloys, Metglas, etc. are necessary to develop the pulsed-energy storage elements envisioned above. With better understanding of basic material properties, more compact and reliable devices will be developed. The appropriate materials science research will be part of the modulator and klystron development program.

8.6 RF Drive and Phasing Systems

The design of the rf drive and phasing systems for the NLC will be based heavily on experience gained from existing systems developed for the SLAC linac and its 60-GeV upgrade for the SLC, with further extensions based on the design of the NLCTA. The challenges to be met for the NLC arise from its greater length, number of components that must be controlled, required tolerances, maintainability, and reliability. The requirements imposed on the rf drive and phasing systems are summarized below:

- The drive system must provide stable, adjustable, and reliable input power to 4528 klystrons (approximately 1-kW pulsed peak power per tube). The layout of the system must be such that individual subsystem failures do not cause the beam energy to decrease to less than 85% of its operating level. This requirement is dictated by potential collimator and other damage caused by off-energy beams.
- The amplitude of the drive power must be adjustable so as to individually saturate the high-power klystrons.
- The phase of the drive power must be adjustable in several ways:
  - Slowly, to take care of phase drifts and drifts caused by length changes in equipment and terrain. Included here are the couplers, coax lines, intermediate amplifiers, high-power klystrons, SLED-II, waveguide components, and the changes in time at which the bunch trains are injected into the X-band linacs by the S-band bunch compressors.
  - Quickly (within each 1.2-μs klystron rf pulse), to create the proper resultant amplitude and phase profiles necessary to produce the desired fields in the accelerator structures and to place the bunches at the desired phase positions with respect to the X-band waves. These positions are dependent on bunch number and charge, which can change from pulse to pulse, and are dictated by BNS phasing, single bunch and multibunch beam loading, and other considerations.
  - The phasing system design must also include initial adjustments (upon installation) of the waveguide runs which feed the four accelerator sections driven by one pair of klystrons.

The phases of high-power rf fields which act on bunch trains should be adjusted to about 0.5° at 11.424 GHz. A schematic layout of the rf drive and phasing system for the NLC is shown in Figure 8-47.

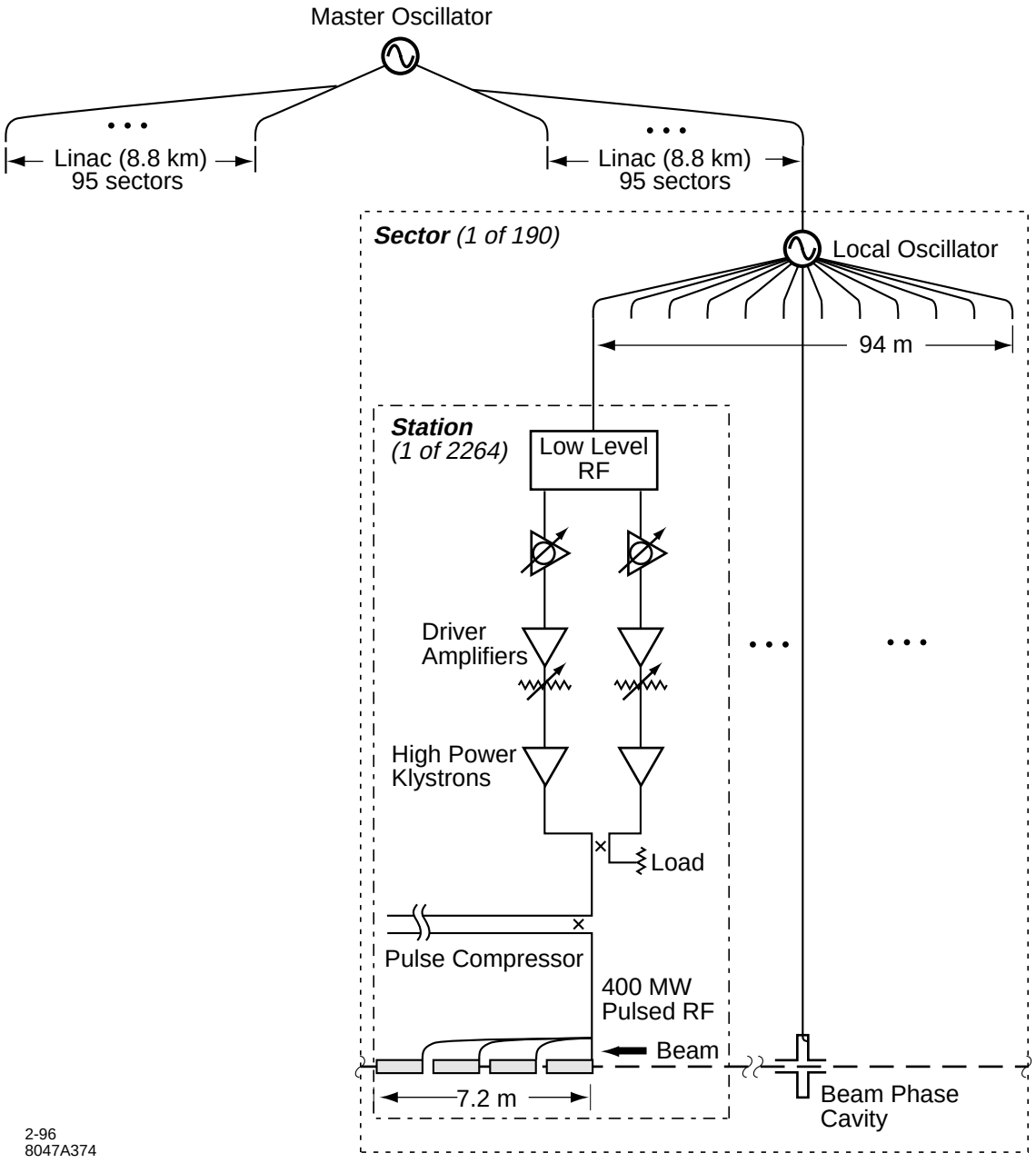


Figure 8-47. *Schematic of rf drive system. The master oscillator provides phase-stable rf to local oscillators in each sector over a fiber distribution system. Local sources in each sector provide 11.424 GHz to each rf station in the sector using coaxial lines. A beam-phase cavity uses the beam to determine the rf phase for optimum operation. Power levels are low (several milliwatts) until the final drive for the high-power klystrons.*

Master Oscillator Quantity: 1	Logical Sectors Quantity: 190	RF Stations Quantity: 2264
• 11.424-GHz reference signal • Fiber optic distribution	• Local oscillator • Beam phase correction • Coaxial rf distribution within sector • 12 rf stations per sector • 94 meters sector length	• One modulator • Two klystrons • SLED-II rf pulse compression (5:1) • Four accelerator sections • Phase coupler for klystron and beam-induced rf • Rf phase detectors

Table 8-13. *Contents and granularity of rf sub-systems.*

8.6.1 Functional Overview

For NLC, the rf system must meet the challenges listed above, while giving the collider control system the tools required to model the machine performance accurately, and to change the operating parameters as conditions mandate. The machine rf requirements and the proposed solutions which follow constitute a proposal to meet these goals.

The sub-systems of the NLC rf drive and phasing systems are shown in Table 8-13 and outlined below. The individual subsystems are fully described in later sections.

The master oscillator for the accelerator generates the accelerator's primary 11.424-GHz reference and uses a fiber-optic based transport system to provide the frequency reference to each sector's local rf source. This signal is used as a diagnostic reference and serves to maintain system stability in the absence of a transported beam.

Each accelerator is divided into 95 sectors of up to 12 rf stations each. The sector size is determined primarily by two operational factors: the practical length of coaxial distribution systems for rf systems, and the requirement of the machine protection system that no single failure of a subsystem result in a destructive, non-transportable beam. For comparison, SLAC has 30 sectors consisting of eight high-power klystron/modulator stations.

Each sector has a local oscillator to generate a harmonically-clean rf signal which is phase locked to the beam for additional stability. This signal is distributed to each station for rf drive and monitoring purposes.

The smallest unit in the accelerator rf systems is the station. Each station consists of a pair of high power klystrons powered by a single modulator. The klystrons are differentially phase modulated to achieve the phase and amplitude control required for SLED-II pulse shape control and to address thermal considerations associated with beam loading (see below). The high power klystron outputs are combined using a four-port power combiner with a high power load on the “unused” quadrature-phase output. The “used” output of the combiner drives a single pair of SLED-II energy compressor delay lines and four accelerator sections.

The amplitude output of the SLED-II compressor is shaped by phase modulation of the two klystrons in order that the power extracted by each electron bunch is exactly matched by the input rf power. In this way all bunches in the bunch train experience equal accelerating gradients. Indeed, the operational goal is to provide an accelerator gradient that will accelerate all bunches equally, allowing for changes in pulse length, shape and total charge.

• The beam energy gain must be stable and understood.
• The rf entering the accelerator must be modulated in amplitude and phase as a function of time to produce the design acceleration of the entire pulse-train. (Transient beam loading compensation was discussed in Section 8.2.3).
• The phase of the rf at the accelerator must be adjustable. Machine operation requires that sectors operate slightly off phase.
• The overall rf amplitude and phase of a sector must be controlled, even in the presence of large changes in beam loading.

Table 8-14. *Basic rf control functions for NLC operation.*

8.6.2 System Functional Requirements and Specifications

The acceleration of long trains of electron bunches places heavy demands on the control and monitoring of each sector's rf sources. The general monitoring and control functions for basic machine operation are listed in Table 8-14. Additional diagnostics and monitoring are required, with some specific items detailed in the sections that follow.

Performance specifications for rf devices, monitoring, and control functions depend on the nature of potential disturbances of the accelerating phase and amplitude, and general beam dynamics considerations. Disturbances with uncorrelated noise spectra ("gaussian processes") act weakly on the beam while disturbances which are strictly correlated across larger sections of the machine have a stronger effect on the beam.

Disturbances which have high frequency components (greater than $1/2\pi T_{fill}$ where T_{fill} is the rf filling time) lead to further relaxation of the tolerances which depends on the power spectrum of the disturbance (see Section 8.2.8) and the characteristics of the SLED-II and accelerator structures. Disturbances with very low frequency components (less than the feedback bandwidth) which are corrected by feedback systems also have relaxed tolerances.

Examples of several rf noise sources for consideration are:

- Errors in SLED-II shaping due to klystron bandwidth limits. These errors have both high and low frequency components, and are correlated and stable.
- Modulator thyratron induced phase and amplitude noise. Thyratrons are the primary source of modulator-induced errors, and are uncorrelated with an intermediate frequency spectrum.
- Local oscillator phase noise. These are broadband and are correlated within a sector.
- Thermally induced phase errors due to changes in average beam current. These have an intermediate frequency spectrum and are strongly correlated among all accelerating devices in the accelerator.
- Thermally induced phase errors due to tunnel temperature changes. These are typically very slow changes and are strongly correlated.

As seen by the above examples, no single set of tolerance values can apply. An interesting metric for gauging a disturbance's effect is obtained by comparing the coherence length of the effect (correlation length of a station, a sector, a linac) to the "interaction length" of the process. The device tolerances may generally be reduced statistically by the square root of the ratio of the interaction length to the coherence length of the disturbance. (For the SLC at SLAC, the effective interaction length is on the order of several betatron wavelengths or, near the injector, the length to achieve an energy gain of several 10's of percent.) Because of this scaling, specifications for station and

Parameter	Accuracy	Stability & Resolution	Units
Energy profile	± 0.5	± 0.1	% voltage
Energy gain knowledge	± 5	± 0.1	% voltage
Phase readback	± 1	± 0.1	degree
Phase stability	N/A	± 0.1	degree

Table 8-15. *Summary of rf control specifications for each station.*

sector performance are looser than the accelerator performance goals. Since different errors have different interaction lengths, different relaxations of tolerances apply.

An example is found in the tolerance on the energy gain of each station, where the modeling code must know the beam's energy to scale the magnetic lattice. Errors in the energy result in the phase of the betatron oscillation "slipping" from the modeled values, generating errors in correction by steering and cascaded feedback systems.

Phase and Amplitude Specifications

As noted above, an additional complication (and relaxation of tolerance) comes from optimizing tuning procedures and feedback systems. Sensitivity to accuracy errors in phase and amplitude are quite low, while slow changes and fast jitters may range from annoying to debilitating for machine operation, depending on the power spectrum and the beam energy. Table 8-15 summarizes these goals.

Energy Profile Control

Knowing the expected beam current and pulse profile, a shaped rf voltage output is computed and programmed into the SLED-II energy compression system. Ramping allows a pre-loading of the accelerator with rf energy prior to the passage of the first bunch, with a reduced energy gain for the first bunch. The accelerator rf power immediately following the start of the passing beam would be near the full available power. This is done to allow the wakefields (which diminish the energy gains) of the leading bunches to be compensated before the subsequent bunches arrive.

The ramping of the power from the saturated klystrons can be done by either of two methods: identical phase modulation or complementary (opposite) phase modulation of the two klystrons in a pair. Identical phase modulation is achieved by applying identical time-dependent phase modulation to each of the two klystrons and thus to the SLED-II compressor input, such that four delay cycles of stored power leaving the delay line ports, plus the incoming klystron power, results in the desired rf voltage and phase modulation at the input to the accelerator during the beam's passage. Alternatively, the complementary phase modulation method takes advantage of the symmetry of the power combiner joining a pair of klystrons. Pulse shaping is achieved by phase modulation of the paired klystrons by equal amounts in opposite directions, resulting in a pure amplitude modulation at the SLED-II compressor input, with unused power being dissipated as heat in the combiner load. Complementary phase modulation need not apply any time-dependent phase modulation at the SLED-II input. In the absence of system bandwidth considerations, both methods can generate the same waveforms for the duration of the rf filling time and beam passage.

For the NLC, we plan to use the method of complementary phasing (in opposite directions) of the accelerator klystrons to program the desired time-dependent SLED-II output. This method has several advantages over the identical

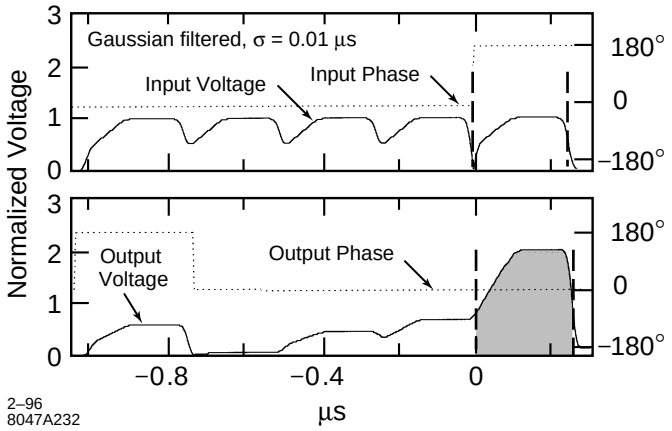


Figure 8-48. *Illustration of SLED-II compression with pulse shaping achieved by complementary modulation of the phases of paired klystrons (equally, but in opposite directions) to achieve the input power modulation. Note that the phase of the SLED-II input is zero (modulo 180°), allowing use of simple nulling phase detectors. Klystron impulse response modeled as a gaussian with* $\sigma = 0.01\ \mu s$.

phase-modulation method from the perspectives of improving performance and simplifying rf control design. These advantages are:

- Reduction of power dissipated in the accelerator. The unused power is dissipated in the power combiner load and not in the accelerator.
- A simpler correction for average power effects when the bunch count or the beam current is temporarily reduced.
- The bandwidth and finite rise time of the klystrons and drive systems do not result in complicated waveforms which may decrease performance and increase diagnosis and maintenance costs.
- A simpler monitoring system. Since the phase input and output of SLED-II are approximately flat (modulo a 180° phase jump), lower cost phase detectors may be used: either simple mixer-type phase detectors or quadrature-type phase detectors operating at a fraction of their potential capability. Phase modulated waveforms require detectors with a large phase dynamic range, and thus careful calibration to insure that changes in the overall phase do not affect the diagnostic reliability.

Sample modulation waveforms for the complementary phase modulation method are shown in Figure 8-48. For this study, the desired SLED-II output pulse shape is a ramped increase in gradient from 40% to 100% over the 100-ns fill time, followed by a 150-ns flattop. (See Section 8.2.3 on Transient Beam Loading.) In this study, the klystron's output bandwidth is modeled by the convolution of a 10-ns gaussian with an ideal klystron's output in the time domain. While there are an infinite number of equivalent solutions satisfying the output requirements, the following simple model with five-fold repeated symmetry for each of the compressed sub-sections is proposed to reduce the diagnostic problems presented to the rf systems maintenance personnel. Amplitude is shown as a solid line, phase as dotted, and the fill/acceleration interval is delineated with vertical dashed lines.

The desired resultant SLED-II envelope and the mathematical basis for the waveform are introduced in Section 8.2.3, “Transient Beam Loading”. The operational goal for the NLC rf control system is to minimize beam energy errors; thus the controls must allow the development of models and algorithms for near real-time modification of the input waveforms to achieve the necessary energy-gain profile needed for the real current pulse shape and charge.

Knowledge of the Energy Gain

Control of the machine lattice (the magnetic optics which guide the beam) requires knowledge of the energy of the beam at all points in the accelerator, and of changes in the energy profile incidental to any significant bunch phasing, klystron phasing, or klystron population changes. Absolute calibration of the results is not as important as the consistency of the derived energy gain readings, since the total machine energy is ultimately determined by direct spectrometer measurements.

Modeling inputs use the nominal energy gain of the individual stations, beam current information, and other factors to compute the energy profile. Rf controls must be able to accurately report the approximate energy gain expected from the station as well as the current operational status of the station.

Experience with the SLC at SLAC has shown that the models for machine operation must be updated whenever the klystron population is varied, and that consistency in the predicted energy contribution of any individual station is more important than the absolute accuracy. Even so, each station's energy contribution has been individually calibrated using bolometric power meters and known coupling ratios, and the results have been compared with direct beam-based energy contribution measurements. Discrepancies in beam energy of greater than 5% have been observed in the two methods, and correction factors are generally placed in the database to insure that the instrumentation accurately reports each station's energy gain.

Phase Control

The DC phase of the output from each SLED-II compressor will be monitored and set to the nominal value using low-level phase shifters upstream of each klystron. Desired phase values will be derived by using a coaxial loop loosely coupled to the last cavity of an accelerator section and comparing a beam-induced rf signal in the accelerator section with the SLED-II rf signal.

Phase Stability

Stable beam operation is required over a range of 1 to 90 bunches and bunch charges varying between 10^9 and 10^{10} electrons per bunch. Uncorrected changes in either parameter vary the power deposited in the accelerator sections, with resultant changes in accelerator temperature of a degree Fahrenheit and average phase changes of 3° to 5°. Thermal characteristics of the accelerator structure are discussed in Section 8.2.5.

Corrections for changes in microbunch configuration will be done iteratively, allowing local controllers to modify energy profile appropriately.

Machine protection system (MPS) algorithms immediately reduce the bunch count to zero in the presence of beam loss. Operation resumes with a single-bunch, high-emittance probe beam at a low rate (1–10 pulse per second), followed by single bunch operation at full machine rate, with a ramped bunch count increase as machine losses permit. Multibunch operation is always initiated with a single bunch probe beam on the previous machine cycle. Bunches are always deleted from the tail of the pulse, and thus will not require real-time modifications to the rf shaping. (Machine protection issues are discussed in Chapter 16.)

Machine repetition rate will be stepped up from the lower to full bunch count, with stable machine operation demanded without any allowance for thermal corrections. Corrections for this effect must be incorporated into the design, reducing the related phase error to less than 0.1° maximum net phase excursion averaged over all klystrons. (See Section 8.6.2.)

Two control paradigms are being considered, both of which may be required to meet the phase-stability imposed by the above MPS requirement for the main linacs. They are:

- Complementary phase modulation of pairs of klystrons to reduce the excess power entering the accelerator by dissipating power in the fourth port of the primary power combiner. In the same manner as the power is ramped up for the fill time prior to injection of the beam, rf amplitude will be stepped down following the passage of the leading bunches to reduce average power to the beam-loaded level. (This requires that klystrons be separately driven by pairs of klystron drivers to allow fast control of power.)
- Feed-forward control of the thermal regulation system for the accelerator cooling water to directly compensate for anticipated thermal changes related to beam current changes. Control using such regulation could practically reduce the thermal excursion by factors of 10 to 30, bringing the net effect to within operational tolerances. Such a feedback system is an NLC-specific development project since the requirements are more stringent than those generally met by industrial control specialists. Experience at SLC also indicates that use of thermal controls of accelerator related water systems has very high installation and maintenance costs, and thus is not being considered as the preferred option.

Phasing of the Accelerator

Phasing of the accelerator will use the beam to establish the optimal phase for each rf station. Fields left in the accelerator following the passage of a short pulse (containing 1 to 10 bunches of 3 to 7×10^9 electrons per bunch) will be used to determine the phase of the beam via a coaxial loop loosely coupled to the last cavity of an accelerator section. Phase measurement equipment monitoring the phase of the rf exiting an accelerator section will measure: a) the rf in the absence of beam loading, b) the beam-induced rf in the absence of klystron rf, or c) a vector combination of klystron and beam-induced rf during normal operation. The optimal accelerating rf phase will be determined from the difference (modulo 180°) of the klystron-driven phase and the beam induced phase, with knowledge of the current beam offset phase (the "phase closure" value). The difference in the readings will be used to phase the klystron station during routine setting of the nominal station phase offset. The phase signal during normal operation (case c) is of limited interest, and is not used for station operation.

8.6.3 Systems Overview

Low-level Accelerator-wide RF Distribution

Phase-stable distribution of X-band rf over the entire length of the NLC by coax is infeasible with current technologies. The NLC rf distribution will be based on:

- A master X-band oscillator.
- Fiber-optic distribution of primary rf to satellite rf oscillators using temperature-stable fiber distribution line.
- Phase-locked ultra-stable satellite oscillators at each sector locked to a phase-shifted reference derived from fiber optic distribution.
- Local beam-induced rf reference cavity near each satellite oscillator.
- Phase feedback controlling phase offset between fiber reference and local oscillator to null phase detector monitoring beam reference cavity.

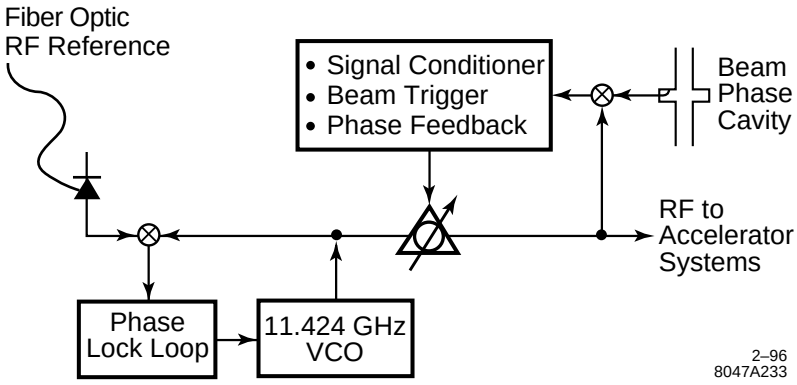


Figure 8-49. *Schematic of a local rf source for each sector. Sector rf is derived from a low-noise oscillator with phase locking to detected beam-induced phase reference. Fiber-optic reference using liquid-crystal polymer (LCP) time-stabilized fiber provides stable reference in absence of beam.*

Sector RF Source

Each sector will receive an amplitude modulated fiber frequency source, which will serve as the basis for the local oscillator's phase lock to master clock. The local oscillator serves as an extremely narrow-band filter, reducing the quantum (shot) noise and thermal noise power that would otherwise dominate the short term stability of an optical based distribution system. Distribution within each sector is achieved using low temperature coefficient copper coaxial lines [Andrew], which deliver a low power signal to each station. An additional cable will transmit the signal to the next sector, allowing a cross-check of each system's stability.

The local beam phase reference will consist of a tuned cavity, or of a few cells similar to the accelerator, that will provide a temperature-invariant reference signal or a high-bandwidth beam pickup followed by an appropriate bandpass circuit. The resultant phase information will be used to remove the local source's residual phase errors prior to distribution to the rf devices in the sector, as shown in Figure 8-49. The use of a beam pickup is considered a technical option which can eliminate an additional tuned cavity in each sector but will require the development of an specialized phase monitoring circuit. (Beam-phase pickups are used successfully at the SLC at SLAC for damping ring phase control at somewhat looser requirements.) Knowledge of the present beam parameters is required for this feedback loop, since the induced phase depends on the so-called “Phase closure” setting.

Long-haul distribution using fiber optic transmission lines has shown promise at SLAC and abroad. Special time of flight stabilized single mode fiber is manufactured with a proprietary coating system [Kakuta 1987, Kotseroglou, Sumitomo, Urakawa 1991], which substantially reduces the thermal drift inherent in glass-based transmission systems. Communication starts at a source laser using either direct intensity modulation or a narrow-band laser followed by an intensity modulator. Parallel fibers distribute signals to local sources in key sectors, where secondary fiber optic lines carry frequency information to each sector's rf source. Instrumentation at distribution points will consist of a local oscillator followed by a laser and modulator system as described above, and may share the oscillator with phase feedback of the standard local source oscillator. An additional fiber will transmit rf information to the next distribution point, allowing a cross-check of each system's stability. The long-haul fibers will be equipped with an interferometric length correction to remove the majority of the fiber's residual phase error. Similar fiber systems are used at SLAC to synchronize a laser for the E-144 experiment to the electron beam [Kotseroglou], with a residual fiber sourced phase error believed to be better than a fraction of a degree at X-band. The use of interferometric length correction has been used successfully for the primary rf distribution system for the two-mile-long SLC linac at SLAC, with distribution-based errors believed to be better than several degrees at X-band diurnally.

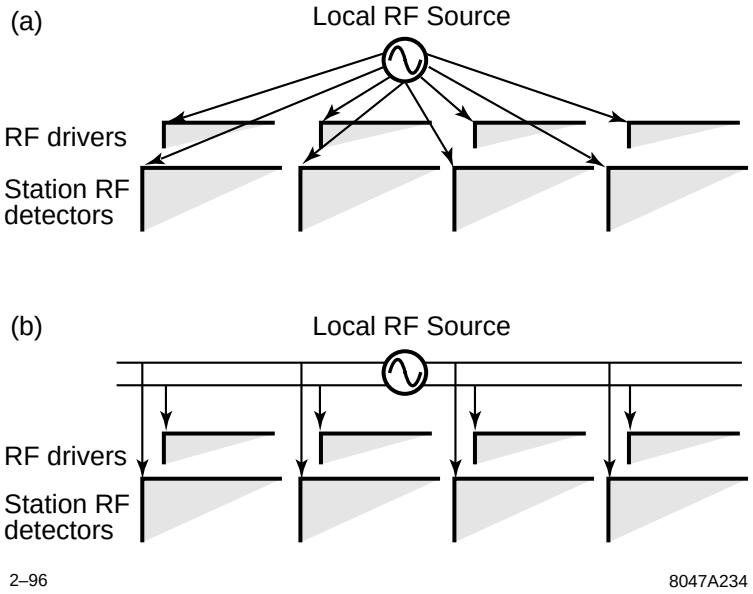


Figure 8-50. *Two methods of rf distribution within any given sector. The use of redundant distribution lines allows identification of distribution-related phase errors. The first schematic (a) uses sets of dedicated lines to each klystron/modulator and instrumentation station, while the second (b) uses common lines for each function, with signal extraction through the use of a directional coupler.*

Sector RF Distribution

Conventional coaxial distribution systems are proposed to transport rf drive and reference signals from the sector's rf source to each station. The choice of coaxial lines is based partially on the quantity of such connections and demonstrated reliability of the medium. Fiber-optic based transmission has been considered, but fails to meet the phase noise requirements of the klystron rf instrumentation without the use of an additional local oscillator at each station.

The cable plant consists of two identical rf cables from the sector rf source to each station: one cable is used for the amplification of the high-power rf by the linac klystrons, while the second cable is used solely for the rf instrumentation for a local phase reference (Figure 8-50). The use of a second reference helps for identifying distribution-related phase problems as anomalous station phase errors which can (with care) be identified and corrected. Experience at the SLC at SLAC has shown such redundancy is critical in identifying problems in the rf distribution system and has been specifically cited as a major factor in achieving the SLC's availability and reliability.

Station RF Components

The smallest unit for control of rf-related devices is the station. Each station consists of a pair of klystrons driven by a common modulator. Rf power is combined using a hybrid “T” with a load on the fourth port, and is compressed in time with a SLED-II compressor. Each station supplies power to four accelerator sections, as shown in Figures 8-47 and 8-51. Rf drive and rf reference signals come directly from the sector's low level rf distribution system on separate cables to allow diagnosing of rf distribution-related failures. The phase modulation generators will deliver different phase modulation when a klystron station is at standby (off-beam) time to maintain correct thermal equilibrium.

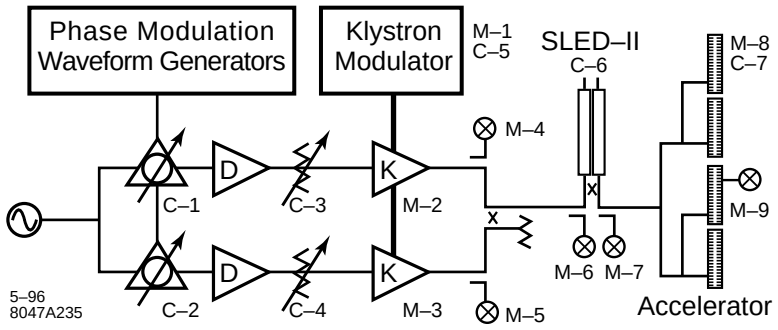


Figure 8-51. *Schematic of rf components and controls for each station. Klystron drive is derived from the sector's rf source, with an adjustable phase and drive control for each klystron. Major control and monitor points are shown.*

C-1 C-2	Klystron Phase Controls	Correction for slow and thermal drifts.
C-3 C-4	Klystron attenuator	Set for saturated operation. Part of klystron protection system.
M-1 C-5	Modulator monitor and controls	Triggering of klystron in response to control systems needs, readback of modulator status plus modulator (beam) voltage and current waveforms.
M-2 M-3	Monitor of klystron parameters	Signals such as klystron vacuum, body beam interception, etc.
M-4 M-5	Monitor of klystron rf forward and reflected signals	Klystron phase and amplitude waveforms. Primary monitor point for klystron performance. Reflected power for klystron protection systems.
M-6	Monitor of SLED-II input	SLED-II input phase and amplitude.
C-6	SLED-II cavity length tuning control	Motors or heaters as chosen to achieve modeled compression factor.
M-7	Monitor of SLED-II output and accelerator input	Phase monitor point for “trimming” phase of klystrons. Phase and amplitude waveforms of forward rf.
M-8 C-7	Accelerator cooling monitors and controls	Adjustment of water cooling set points and readback of temperatures.
M-9	Accelerator output monitor	Phase monitor allows independent measurement of phase of beam and phase of klystron system.

Table 8-16. *Monitor and control points in the rf system.*

Monitor and control points are summarized in Table 8-16.

Klystron RF Driver (TWT or Low-power Klystron)

There are various schemes for providing the approximately 1 kW of pulsed rf needed to drive each high-power klystron. It is useful to have individual drivers because the various phase manipulations involved in the operation of the rf compressors are best done at the milliwatt-level. Power-combining of the two klystrons on the single modulator is done by phase modulation of the two klystron inputs by equal amounts in opposite directions. This has the effect of sending the desired vectorial sum into the rf pulse compressor and dumping the vectorial difference into a load.

There exists extensive traveling-wave tube (TWT) technology at this frequency and power level as a product of military applications. The available TWTs are reliable, but fairly complex to manufacture and operate. They have already been extensively cost-optimized, and are still very expensive.

A small X-band driver klystron of fairly narrow bandwidth and modest gain could be designed. In construction, it would be relatively simple compared to a TWT. It is possible that after the initial engineering models are tested, a modest program of redesign for automatic manufacturing could reduce the unit cost significantly.

Low-level RF Station Controls

The rf for each klystron is appropriately phase modulated, at milliwatt power levels. Downstream from the kilowatt-driver amplifier, klystron attenuator controls are used to maintain saturation levels of individual klystrons. The attenuators will also facilitate soft run-up of stations as might be required following venting of vacuum systems. Interlocks are discussed in Section 8.7.3.

Klystron and Modulator RF-Related Control and Monitor Points

Each station has a single modulator station powering two klystrons. The modulator instrumentation and protection system receives triggers from the rf support system, as well as indications of excessive reflected power requiring a modulator trip.

When the station is not used for the current beam pulse, it is fired in a standby mode, where the trigger time is offset from the normal beam time by at least several microseconds, and the modulation phase patterns are altered to maintain thermal equilibrium in the absence of beam loading. Appropriate triggers and phase controls are supplied by the rf support systems. For the accelerator at SLAC, standby time offsets are used as part of the beam containment system to insure that even a dark-current beam is not generated coherently through the linac. These delays are set on a sector-based pattern and implemented in hardware. The NLC may require similar logic to ensure that the radiation safety considerations have been met, without burdening the operation with unreasonable thermal settling delays.

The modulator system will provide the rf controls status information as well as a few high quality video-bandwidth monitoring signals of klystron beam voltage and beam current for monitoring by the rf digitization system (three channels per station).

High-Power RF Monitor Points

Between the output of the klystrons and the input of the accelerator is a high-power combiner and a SLED-II pulse-compressor system. Forward and reflected rf phase and amplitude are monitored at each point as identified in Figure 8-51.

Klystron forward output couplers are used to diagnose the performance and stability of individual klystrons, while reflected output peak values are provided to the modulator protection system. Phase and amplitude waveforms for the forward rf signal are required.

Combiner output signals are used to identify the performance of the SLED-II system, while reflection signals measure SLED-II tuning accuracy. Phase and amplitude waveforms are required in the forward direction (combiner output), while amplitude is required for the reverse output (SLED-II reflection).

The SLED-II output coupler is used to routinely trim the phase of the two klystrons to meet accelerator phase requirements. Outputs at this point are expected to exactly reflect the modeled rf power profile, and differences in this waveform will be corrected by cautious modifications to the klystron input waveform. The phase stability of this monitor point must exceed the station's phase stability requirements. Phase and amplitude waveforms for the forward rf signal are required.

Peak values of each monitor for each pulse will be sampled and digitized. Waveform recording of each channel will occur periodically and as requested by the control program. Periodic analysis allows a historical monitor (jitter and values) to be stored over days and months, while on request waveform acquisition allows the control room and maintenance personnel to diagnose suspect devices.

Sample size and resolution must allow device bandwidth limited signal acquisition with sufficient granularity. Minimum sampling speeds of 100 MSample/second and buffer depths of 256 allow the entire pulse to be represented with 10-ns sample sizes. 0.1% amplitude and 0.1° phase stability requirements can be met with digitizers with resolution limits of 12 bits per sample ($\pm$2048). The actual instrumentation sampling speed and resolution will be determined to take best advantage of commercially available instrumentation.

There are several cost-containment decisions which must still be made regarding station instrumentation. Current costs for VXI-bus (computer interfaced) digital oscilloscopes of the required bandwidth are quite high, and suggest that dedicated signal acquisition hardware for each channel will be prohibitively expensive. Cost control may dictate multiplexing of the signal inputs into a common digitizer. If the reliability of multiplexers can be improved, their use can reduce instrumentation investments at the operational cost of limiting simultaneous monitoring of multiple channels within a station. Alternately, instrument development efforts combined with waiting for the growth of the related industries may allow cost effective solutions not currently available.

Instrumentation may be shared between adjacent klystron stations or between clusters of stations, resulting in further reductions in cost. The locations and granularity of instrumental and computational support in the NLC must be based on an evaluation of the cost and power of available instrumentation. Serious attention must be paid to the relative costs of shipping rf signals in phase stable cables *vs.* the local conversion of phase data to video and shipping video to the analysis points. The most probable configurations are to have instrumentation centrally located between pairs of stations, or individual instrumentation for each rf station.

SLED-II Delay-line Tuning

The SLED-II delay lines must be tuned to maintain both differential and common errors to a small fraction of a wavelength. This tolerance can only be met through the use of an active temperature or electrical length control.

Differential length errors will result in rf power reflections into the klystrons and reduced energy gain for the station. Differential length adjustment must be done by minimizing the rf power reflections back into the klystron power combiner (by tuning). Common and differential length errors result in a reduction of the station's energy gain. These errors can be removed mechanically by controlling the delay line length. The common length errors can be removed electronically by step-modulation of the input phase to compensate for the common phase-length error in the SLED-II delay lines. Common length errors can be directly measured from the SLED-II forward power coupler as follows: (1) measure the phase of the earliest delay-line fill (this is from the reflections off the coupling irises), and (2) measure the phase of a subsequent fill (reflections off the irises plus the emitted field from the cavity). The second measurement should be at a phase angle of 180° shifted from the first.

The SLED-II delay lines in the NLC will be tuned either by sliding choked shorts at the end of each line (as used in the NLC Test Accelerator), or by thermal control of delay lines. Either solution will work.

Accelerator Control and Monitor Points

The pickup loops at the downstream end of each accelerator structure will extract the beam-induced signal for phasing the accelerator, as discussed in Section 8.6.2. Temperature monitoring of the copper of the accelerator combined with measurements of the average rf power and beam current will be used to set the water system to the optimal control temperature. The pickup loops at the downstream end of the each accelerator structure will be used to optimize this control. The time-dependent phase angle of the output of the accelerator following the passage of a short (single or few bunch) beam will be used as a cross-check of the temperature settings for the accelerator structures.

Accelerator RF Beam Position Monitor

Signals from the damping manifold monitor couplers will be used to optimize the transverse positions of the accelerator structures to minimize steering due to dipole wakefields.

Trigger Timing

The timing system must deliver to each klystron station the appropriate triggers and machine-configuration information to allow the triggering of the modulator and the rf and protection sampling systems. Configuration information will include information identifying if the trigger should result in a beam acceleration, or if the station is on standby timing for the next pulse. Due to the Machine Protection System constraints, the triggers and configuration information must be delivered and sampled at least 100 μs prior to beamtime. The modulator trigger and rf shaping electronics will require timing that is adjustable in steps of about 5 ns, with short- and long term-stability a factor of two smaller.

Feedback and Related Systems

Active feedback systems, similar to those used in the SLC, are required to meet the NLC beam quality and machine performance requirements. Feedback systems, as defined below, which will directly interact with the rf systems are:

- Machine Energy Feedback
- Machine Energy Spread Feedback
- Bunch Energy Spread Feedback
- Pulse Energy Error Feedback
- Intensity Energy Feedforward
- Downstream Feedback Systems

Figure 8-52 illustrates the parameterization of energy and energy spread space for the NLC.

"Machine energy error" is the error in the energy of the bunch centroid. Errors are measured by the feedback system using beam position monitors in a dispersive region and computing the beam parameters (ΔE, X, X', Y, and Y') using the modeled machine lattice. Changes in desired accelerator energy gain are computed and transmitted to the effected sector controllers. The energy gain is controlled by deliberately misphasing sector pairs of klystrons with complimentary phase offsets, producing changes in the net energy gain without affecting the beam energy spread.

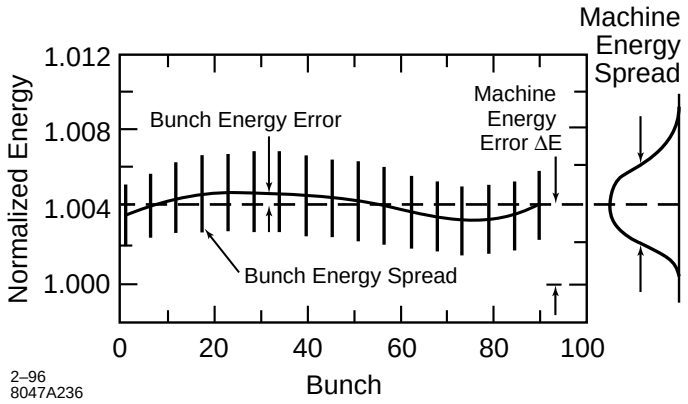


Figure 8-52. *Plot of selected bunches within a NLC pulse. Parameters Energy Error, Pulse Energy Error, and Bunch Energy Spread are shown in the central graph. The projection at right depicts total machine energy spread, with the Machine Energy Error as its mean, and the width representing both the Pulse Energy Error and the Bunch Energy Spread.*

Experience in the SLC suggests that at least $\pm 2\%$ of the total gradient be available for total energy correction, which will be used to correct for gradient errors in the accelerator, and for changes in centroid energy from energy spread and pulse energy corrections [Sheppard 1987]. Experience also suggests that the feedback correction rate for this system will require at least 20 energy corrections per second at 180-Hz machine repetition rate.

Machine energy spread is measured by a dedicated spectrometer which measures the composite energy spread of the entire bunch train. Energy spread errors are fixed by changing the injection phase of the bunch from the damping ring (the "phase closure"). Communication channels for this loop are similar to those for the ΔE correction with a system speed of several corrections per second.

Average beam energy errors and individual bunch energy spread are continuously measured by a dedicated spectrometer which can separately resolve the energy differences and the rms energy spread of the bunches. Energy errors are corrected by designing a modified SLED-II compressed rf profile, while energy spread errors in the individual bunches are corrected by controls in the damping ring extraction system and by the imposition of a phase modulation in the shaped SLED-II output. Communication channels for this loop are similar to those for the ΔE correction with a system speed of order a correction per hour with results available for history buffer logging every several minutes.

Intensity feedforward systems require the energy of the accelerator to be changed within a damping-ring store time. The measured stored beam current is used to predict an expected centroid energy shift ΔE due to off-normal beam loading effects. Communication channel speeds for this loop must be quite high, A dedicated channel to each sector may be necessary to maintain a rate of 180 corrections per second [Jobe 1991, Seeman 1991b].

Specific downstream feedback systems (such as steering systems, energy feedback systems, and detector noise minimization systems) may require knowledge of klystron errors or modulator protection cycles. Detector noise minimization systems, for example, will need to know if an rf station in the early section of the machine has faulted prior to computing orbit corrections to minimize beam scraper spray into the detector. Additionally, in the presence of extended error conditions (such as a vacuum fault), the machine protection system might abort beams to minimize detector backgrounds. In the SLC, each rf station announces its faults to the control system using a dedicated communication channel [Jobe].

Parameter	Resolution
Average Energy Measurement	$<0.5\%$
Machine Energy Spread	$<0.05\%$
Bunch Energy Spread	$<0.1\%$
Bunch Energy Error	$<0.05\%$

Table 8-17. *Spectrometer resolution.*

8.6.4 Spectrometers

Primary machine optimization for the rf phase settings for the individual stations will use beam-induced phasing; all other optimizations will require a beam line spectrometer. Resolution requirements are summarized in Table 8-17 and such measurements should be available at the machine end, and at several points along the accelerator.

In order to meet target resolutions for machine and bunch energy spread, the spectrometers must be located at points of high dispersion ($\eta \gg 2000\sigma$). For a horizontal beam size (σ) of 0.1 mm at the spectrometer, the horizontal dispersion (η) should be greater than 0.2 m. Backgrounds for the spectrometer detectors must be sufficiently low for the spectrometers to operate in the machine environment.

The spectrometer detectors could utilize a pulsed laser beam probing the density of the bunch at the high dispersion point. Compton scattered electrons and a scanned laser probe beam, would provide all the information required. Individual bunch analysis might not be necessary, if the detector is capable of reporting separate values for early, middle, and later bunches in the train with sufficient resolution.

8.7 RF Protection and Monitoring Systems

All protection systems are serviced by dedicated local computers or programmable logic controllers (PLCs) plus some additional faster protection circuitry. This protection system combines the logical protection and operating functions of the modulator, the klystron, and the waveguide and structure systems. It includes the water-cooling system, the temperature-monitoring system, the various electronic monitors of the klystron support power supplies and the modulator electronics, the related vacuum systems, and the high-power rf protection. While these various protection and interlock systems are serviced by one PLC, the functions will be discussed separately in the sections below. The PLC communicates with the central computer to allow remote data logging, error reporting, and configuration control.

8.7.1 Waveguide Protection and Monitoring

Protection includes the high-power X-band waveguide, the SLED-II energy compressors, and the disk-loaded waveguide. Other than accelerated electron beam interception damage over which this interlock system has no control, the primary causes of potential component damage are arcing brought on by poor vacuum and loss of cooling. Because of the limited pumping speed of small waveguide, the high-power rf distribution system contains many individual ion pumps. It is expected that these ion pumps will be powered by a small number of multiple-output ion-pump power supplies that will be interlocked with the rf system module.

There will be a limited number of reflected energy pickup points, especially in the output arms of the klystrons, that will detect reflected energy due to arcs or SLED-II mistuning, and will operate through the PLC to inhibit either the drive rf to the klystron, or the klystron pulse modulator.

8.7.2 Klystron Protection and Monitoring

Klystron protection is accomplished by providing monitoring and interlocks on the systems which provide power and cooling water to the klystron, the klystron vacuum and the rf systems including the system vacuum. These monitoring and interlock signals are both analog and digital.

Most of the interlocking is done with a PLC which provides the interlocking and sequencing logic for the klystron system. Critical interlocks such as Klystron Output Reflected Power, which require faster response time, may be hard-wired for speed, and redundant to the PLC interlock.

The water system is designed for two klystrons per station with separate water circuits for each klystron. Each klystron will have a water circuit for its anode, its body and window and its collector. This water manifold will provide the cooling water for the klystron and modulator oil tank. Water flows will be monitored by the PLC for interlock and general monitoring functions.

Each klystron is instrumented separately for window temperature and body temperature. Increases in the body temperature, which increases if beam interception increases, is fed to the PLC and compared against a predetermined level; exceeding this level will cause the PLC to turn off the klystron beam. Window temperature will cause the rf drive to be removed if a maximum temperature is exceeded.

The klystron ion pumps will be connected in parallel into one controller for each klystron. The controller will have an internal vacuum pressure limit interlock which will be manually set. An excessive vacuum pressure interlock will turn off the klystron beam.

The klystron heater circuit, separate for each klystron, is equipped with an interlock for minimum heater current, and a warm-up time delay is provided by the PLC once the minimum current is exceeded. The klystron beam is inhibited until these interlocks are cleared. Analog signals for heater voltage and current are read by the PLC.

Klystron peak-cathode voltage and peak-cathode current are digitized in the support electronics and read into the PLC for over-current and over-voltage interlocking and into the main computer system for data-taking and storage. Klystron heater and high voltage operating hours will be totaled in the PLC.

The rf output waveguide transport system, which is under vacuum, is connected to ion pump supplies with manually set high-vacuum pressure interlocks. These feed into the klystron and modulator PLC to remove the rf drive when there is an excess pressure condition.

The rf output waveguide transport system also has various forward and reflected monitor points. An example of one of these monitor points is at the output of the klystron. These signals are each transmitted through a four-port coupler. One of these ports feeds a crystal detector into a video amplifier interlock system which provides buffered video outputs, a peak and hold-detected output, and a comparator circuit. The comparator circuits with manual adjustment can be used to interlock any of these detected signals. The output of the interlock is sent to the PLC for action. Klystron-reflected power is one of the critical interlocks which will also be hard-wired from the video amplifier interlock to directly remove the rf drive.

8.7.3 Modulator and Support Electronics Protection and Monitoring

The modulator control, interlocking, and protection is also accomplished through the PLC. The main protection interlocks are excessive EOLC current (for a klystron arc) and high-voltage-power-supply-overcurrent (HVOC).

The modulator protection system contains discharge solenoids, barriers, and door interlocks for protection of both personnel and equipment. It may be connected to the overall accelerator Personnel Protection System.

Thyratron warm-up time delay will be provided in the PLC after the thyratron electrodes are energized. Thyratron heater hours will also be totaled in the PLC.

8.7.4 Klystron and Modulator Logic Controller

At the present time, PLCs are being used for monitor and interlock functions in the Klystron Test Lab. On the NLCTA, two PLCs are used at each klystron station to provide modulator interlocking and control, and to monitor the operation of the klystron rf system and the various interlocks that protect the rf station. It is expected to use this same technology for each NLC rf station, but the general PLC will be replaced by a dedicated logic controller that is designed and optimized for NLC rf station operation.

8.7.5 Modulator Interactions with the Machine Protection System

The Machine Protection System (MPS) requires that the beam quality be maintained at an energy of at least 85% the expected energy. This requirement is determined by the limited ability of the collimators and beam line transport systems to handle a wildly off-energy beam. Under certain failure modes, catastrophic system failures can result from a single errant pulse.

Because of the large number of klystron/modulator stations, the interactions between MPS and klystron controls are fairly weak. The MPS is expecting each station to deliver a readiness confirmation which will be summed using a majority logic system to insure that the energy profile of the machine will follow the modeled values within a small tolerance ($\pm$15% nominal). The implications and rules for the modulator and klystron systems are as follows:

1. The triggers for the klystron operation will be delivered at least 100 μs prior to beam time.
2. Following the receipt of the trigger, the modulator should be fully charged, and ready to fire.
3. The modulator system will promptly return a status indicator to the MPS controller indicating the readiness of an rf station and ability to fire its modulator.
4. Once confirmed, the modulator must fire at the designated time. Normal interlock conditions can not be allowed to disable modulator operation following the transmission of an MPS confirmation.

These MPS rules will insure that vacuum failure, AC power loss, trigger systems failure, and other larger system failures will not result in accelerator damage. Failures which are statistical in nature (such as thyratron triggering failure) are not expected to violate the MPS guidelines. (Special consideration also must be given to the readiness of kickers, dumpers, and perhaps certain critical klystron-driven systems.)

References

[SLAC 1993] SLAC, "Next Linear Collider Test Accelerator Conceptual Design Report", SLAC Report 411 (Stanford University 1993).

[Andrew] Andrew Corporation (Orland Park, Illinois), *Phase Stabilized Cables*, General Catalog No. 34, pp. 454–455.

[Adolphsen] C. Adolphsen, K. Bane, H. Higo, K. Kubo, R. Miller, R. Ruth, K. Thompson, and J. Wang, "Measurement of Wakefield Suppression in a Detuned X-Band Accelerator Structure", SLAC–PUB–6629, submitted to *Phys. Rev. Lett.*

[Adolphsen 1992] C. Adolphsen, K. Bane, G. Loew, R. Ruth, K. Thompson, J. Wang, "Measurement of Wakefields Generated in Accelerator Test Structures Using the SLC", SLAC–PUB–5941, 15th International Conference on High Energy Accelerators, Hamburg, Germany (1992).

[Adolphsen 1994] C. Adolphsen *et al.*, "Measurement of Wakefield Suppression in a Detuned X-Band Accelerator Structure", *Proc. 17th Int. Linear Acc. Conf.*, Tsukuba, Japan, submitted to *Phys. Rev. Lett.* (1994).

[Bane 1993a] K.L.F. Bane and R.L. Gluckstern, "The Transverse Wakefield of a Detuned X-Band Accelerating Structure", SLAC-PUB-5783, *Particle Acclerators* **42**, 123 (1993).

[Farkas 1974] Z.D. Farkas *et al.*, "SLED: A Method of Doubling SLAC's Energy", SLAC–PUB–1453, *Proc. 9th Int. Conf. on High Energy Accelerators*, Stanford, CA, 576–583 (1974).

[Farkas 1986] Z.D. Farkas, "Binary Peak Power Multiplier and Its Application to Linear Accelerator Design", *IEEE Trans. MTT*-**34**, 1036–1043 (1986).

[Farkas 1994] Z.D. Farkas and P.B. Wilson, "Ramping Profile for Exact Beam Loading Energy Compensation in the NLC Detuned Structure", NLC–Note 4 (Rev. 1), SLAC (December 1994).

[Farkas 1995] Z.D. Farkas and P.B. Wilson, "RF and Beam Loading Parameters for the NLC Detuned Structure with Damping Manifold", NLC–Note 15, SLAC (June 1995).

[Hoag 1993] H.A. Hoag *et al.*, "Flower-Petal Mode Converter for NLC", SLAC–PUB–6182, *Proc. 1993 Part. Acc. Conf.*, Washington, DC, 1121–1123 (1993).

[Jobe 1991] R. Keith Jobe *et al.*, "Energy Feed Forward at the SLC", SLAC–PUB–5541, May 1991, in *Proc. 1991 Part. Acc. Conf.*, San Francisco CA, 1464–1466 (1991),

[Jobe] R.K. Jobe, SLC Veto System, unpublished.

[Kakuta 1987] T. Kakuta and S. Tanaka, "LCP Coated Optical Fiber with Zero Thermal Coefficient of Transmission Delay Time", *Proceedings of the 36th International Wire and Cable Symposium*, Arlington, VA 234–240 (US Army Communications Electronics Command, Fort Monmouth, New Jersey, 1987).

[Kotseroglou] T. Kotseroglou, "Picosecond Timing of Terawatt Laser Pulses with the SLAC 46-GeV Electron Beam", SLAC–PUB–7130 (March 1996), submitted to *Nucl. Instr. and Methods*

[Kroll 1994] N.M. Kroll *et al.*, "Manifold Damping of the NLC Detuned Accelerating Structure", SLAC–PUB–6660, in *Proc. 6th Workshop on Advanced Accel. Concepts*, Lake Geneva, WI (1994).

[Kroll 1995] N. Kroll and K. Ko, private communication about work in progress (1995).

[Lanciani 1953] D.A. Lanciani, "H_{01} Mode Circular Waveguide Components", Conference on Millimeter Wave Research and Applications, Washington, DC (1953).

[Lavine 1991] T.L. Lavine *et al.*, "High-Power Radio-Frequency Binary Pulse Compression Experiment at SLAC", SLAC–PUB–5451, *Proc. 1991 Part. Acc. Conf.*, San Francisco, CA (IEEE 91CH3038–7), 652–655 (1991).

[Loew 1988a] G.A. Loew and J.W. Wang, "RF Breakdown Studies in Room Temperature Electron Linac Structures", SLAC–PUB–4647, presented at 13th Int. Symp. on Discharges and Electrical Insulation in Vacuum, Paris, France (1988).

[Matsumoto 1991] H. Matsumoto *et al.*, "Applications of Hot Isostatic Pressing (HIP) for High Gradient Accelerator Structure", *Proc. 1991 Part. Acc. Conf.*, 1008–1010, San Francisco, CA (1991).

[Matsumoto 1994] H. Matsumoto *et al.*, "High Power Test of a High Gradient S-Band Accelerator Unit for the Accelerator Test Facility", 17th Int'l Linac Conf., Tsukuba, Japan (1994).

[Nantista 1993] C. Nantista *et al.*,"High Power RF Pulse Compression with SLED-II at SLAC", SLAC-PUB-6145, in *Proc. 1993 Part. Acc. Conf.*, Washington, DC, 1196–1198 (1993).

[Nelson 1992a] E.M. Nelson, "A Finite Element Field Solver for Dipole Modes", SLAC–PUB–5881, in *Proc. 1992 Linear Accel. Conf.*, Ottawa, Canada (1992).

[OWGR 1993] "Technology and Costs for 2 Million Copper Discs of the CLIC Accelerator at CERN", Optische Werke G. Rodenstock (1993).

[Ramo 1984] S. Ramo, J.R. Whinnery, and T. Van Duzer, Fields and Waves in Communications Electronics, Second Edition (John Wiley & Sons, 1984). See Sections 7.8 and 7.10 for theoretical ohmic losses in waveguides.

[Read 1993] R.F.J. Read and W.J. Wills-Moren, Cranfield Precision Engineering, Ltd., Bedford, England. "Study Report for the Estimation of Unit Cost of Mass Producing Discs for the CLIC Accelerating Sections" (1993).

[Ruth 1993] R.D. Ruth *et al.*,"The Next Linear Collider Test Accelerator", SLAC–PUB–6252, in *Proc. 1993 Part. Accel. Conf.*, *loc. cit.*, 543–545 (1993).

[Ruth 1993a] R.Ruth *et al.*, "A Test Accelerator for the Next Linear Collider", SLAC–PUB–6293, in ECFA Workshop on e^+e^- Linear Colliders (LC92), Garmisch-Partenkirschen, Germany (1992).

[Seeman 1985] J. Seeman *et al.*, "RF Beam Deflection Measurements and Corrections in the SLC Linac", *IEEE NS32*, No. 5, in *Proc. 1985 US Part. Accel. Conf.*, Vancouver, Canada, 2629–2631 (1985).

[Seeman 1991] J. Seeman, "The Stanford Linear Collider", *Ann. Rev. Nucl. Part. Sci.* **41**, 389–428 (Annual Review, Inc., Palo Alto, CA, 1991).

[Seeman 1991b] J. Seeman *et al.*, "Multibunch Energy and Spectrum Control in the SLC High-Energy Linac", SLAC–PUB–5438, in *Proceedings of the 1991 IEEE Particle Accelerator Conference*, San Francisco, CA, 3210–3212 (1991).

[Seeman 1993] J. Seeman, "Accelerator Physics of the Stanford Linear Collider and SLC Accelerator Experiments Towards the Next Linear Collider", *Advances of Accelerator Physics and Technologies* (Vol. 12 of *The Advanced Series on Directions in High Energy Physics*), edited by Herwig Schopper, 219–248, (World Scientific, 1993).

[Sheppard 1987] J.C. Sheppard *et al.,* "Three Bunch Energy Stabilization for the SLC Injector", SLAC–PUB–4100, in *Proc. 1987 Part. Acc. Conf.,* Washington, DC (1987).

[Sumitomo] Sumitomo Electric Corporation (Yokohama, Japan), Liquid-Crystal Polymer (LCP) Fiber data.

[Takeda 1991] S. Takeda *et al.,* "High Gradient Experiments by the ATF", *Proc. 1991 Part. Acc. Conf.,* San Francisco, CA, 2061–2063 (1991).

[Tantawi 1993] S.G. Tantawi, K. Ko, and N. Kroll, "Numerical Design and Analysis of a Compact TE_{10} to TE_{01} Mode Transducer", SLAC–PUB–6085, *Proc. of the Computational Accel. Conf.,* Pleasanton, CA (1993).

[Tantawi 1995] S.G. Tantawi, private communication about work in progress (1995).

[Tantawi 1995a] S.G. Tantawi, A.E. Vlieks, and R.J. Loewen, "Performance Measurements of SLAC's X-band High-Power Pulse Compression System (SLED-II)", SLAC–PUB–95–6775, submitted for publication to *IEEE Microwave and Guided Wave Letters* (1995).

[Tantawi 1995b] S.G. Tantawi, R.D. Ruth, and A.E. Vlieks, "Active radio frequency pulse compression using switched resonant delay lines", SLAC–PUB–95–6748, in *Nucl. Instr. and Methods* **A370**, 297–302 (1996).

[Tantawi 1995c] S.G. Tantawi, T.G. Lee, R.D. Ruth, A.E. Vlieks, and M. Zolotorev, "Design of a multimegawatt X-band solid state microwave switch", SLAC–PUB–95–6827, in *Proc. 1995 Part. Acc. Conf.,* Dallas, TX (1995).

[Tantawi 1996] S.G. Tantawi, NLCTA test results, unpublished (1996).

[Thompson 1990a] K.A. Thompson and R.D. Ruth, "Controlling transverse multibunch instabilities in linacs of high energy linear colliders", SLAC–PUB–4801, in *Phys. Rev. D* **41**, 964 (1990);

[Thompson 1993] K.A. Thompson *et al.,* "Design and Simulation of Accelerating Structures for Future Linear Colliders", SLAC–PUB–6032, in *Particle Acclerators* **47**, 65–109 (1994).

[Thompson 1993a] K.A. Thompson and R.D. Ruth, "Simulation and Compensation of Multibunch Energy Variation in NLC", SLAC–PUB–6154, in *Proc. 1993 Part. Acc. Conf.,* Washington, DC (1993).

[Thompson 1994a] K.A. Thompson *et al.,* "Design and Simulation of Accelerating Structures for Future Linear Colliders", SLAC–PUB–6032, in *Particle Acclerators* **47**, 65 (1994).

[Urakawa 1991] J. Urakawa and T. Kawamoto, "The Improvement of TRISTAN Timing System", *Proc. 1991 Part. Acc. Conf.* (IEEE Catalog No. 91CH3038-7) 1555–1557 (1991).

[Vlieks 1993] A. E. Vlieks *et al.,*"Accelerator and RF System Development for NLC", SLAC–PUB–6148, in *Proc. 1993 Part. Accel. Conf., loc. cit.,* 620–622 (1993).

[Wang 1994] J.W. Wang *et al.,*"High Gradient Tests of SLAC Linear Collider Accelerator Structures", SLAC–PUB–6617, in *Proc. 17th Int'l Linac Conf., loc. cit.* (1994).

[Wang 1994a] J.W. Wang *et al.,* "High Gradient Tests of SLAC Linear Collider Accelerator Structures", SLAC–PUB–6617, presented at 17th International Linear Accelerator Conference (LINAC 94), Tsukuba, Japan, (1994).

[Wilson 1990] P. B. Wilson, Z.D. Farkas, and R.D. Ruth, "SLED-II: A New Method of RF Pulse Compression", SLAC–PUB–5330, in *Proc. 1990 Linear Accel. Conf.,* Albuquerque, NM (1990).

[Wilson 1992] P. B. Wilson *et al.*, "Progress at SLAC on High-Power RF Pulse Compression", SLAC–PUB–5866, in *Proc. of the 15th Int'l. Conf. on High Energy Accel.*, Hamburg, Germany (1992).

[Wright 1994] E. L. Wright *et al.*,"Design of a 50-MW Klystron at X-Band", Pulsed RF Sources for Linear Colliders: SLAC–PUB–6676, in *AIP Conference Proceedings 337*, Montauk, NY 1994, edited by Richard C. Fernow, 58–66 (AIP, New York, 1995).

Contributors

- George Caryotakis
- Dick Cassel
- Len Genova
- Saul Gold
- Harry Hoag
- Keith Jobe
- Ron Koontz
- Ted Lavine
- Greg Loew
- Bob Phillips
- Ron Ruth
- Sami Tantawi
- Kathy Thompson
- Juwen Wang
- Perry Wilson

9

Collimation Systems

Contents

9.1 Introduction to Beam Delivery Systems

Chapters 9 through 11 describe what are referred to as the beam delivery and removal systems. A schematic of these systems is shown in Figure 9-1. These systems begin at the end of the linac and terminate at a post-IP beam dump. They include a main post-linac collimation system (Section 9.2); an IP switch and big bend (Chapter 10); a pre-final-focus diagnostic and skew-correction section (Section 11.3); a final-focus system (Sections 11.4 to 11.7); and a post-IP beam line (Section 11.8). Chapter 9 also includes the description of the pre-linac horizontal and vertical collimation system (Section 9.3), and a pre-linac bunch-length collimation system (Section 9.4). The design of all of these systems has been strongly influenced by experience with similar systems at the SLC.

Figure 9-2 shows the horizontal and vertical beam envelope from the linac to the IP for the 1-TeV-c.m. beam line, and Figure 9-3 shows the corresponding β functions. The total length is 5.2 km. The first 50 meters contain a post-linac diagnostic chicane, which is then followed by a 2.5-km collimation system. The maximum horizontal points of the envelope in the collimation region, for the most part, correspond to the location of horizontal collimators and chromatic correction sextupoles; the maximum points of the vertical envelope are usually the location of vertical collimators and chromatic correction sextupoles. The next horizontal envelope peak, at 2.6 km, marks the location of the IP switch. After that, up to the 3.0-km marker, is a small envelope region that contains the big (10-mr) bend, and following this, up to the 3.4-km marker, is a small envelope region containing the pre-final-focus diagnostic region. In this region beam sizes in all phases and planes can be measured and the presence of coupling detected and corrected. The region from 3.4 km to 3.8 km contains the beta match into the final focus. Following the beta match, the first two peaks in the horizontal envelope are the positions of the horizontal chromaticity compensation sextupoles. The two large vertical envelope points following these, at about 4.4 km and 4.7 km, are the positions of the vertical chromatic compensation sextupoles. The last peak, in each plane, at the end of the beam line, is located at the position of the final-doublet elements.

Figure 9-4 shows the horizontal dispersion function for the same beam line. The peaks at the very beginning of the beam line are in a post-linac diagnostic region. These regions are similar to others along the linac and are described in Section 7.9.5. Following this and continuing to 2.4 km, one sees the dispersion function of the collimation system. The peak at 2.6 km is in the IP switch, and the small dispersion wiggles which follow are located in the big bend. The small bump at 3.4 km, at the beginning of the beta-match region in the final-focus system, is a short region containing two bends which provide an adjustable entry angle into the final-focus region. The net bend here will have three distinct values corresponding to three distinct final-focus systems that are required to cover the energy range from 350-GeV to 1.5-TeV c.m. Each has slightly different internal bend angles, but all reside on the same support bench and have the same IP location. The layouts are shown in Figure 11-9. The remaining dispersion is in the final-focus system.

We have studied all beam delivery systems for center-of-mass energies from 350 GeV to 1.5 TeV for a broad range of assumptions on beam and IP parameters, and have shown that it is possible to meet the specifications for this entire range. Above 1.0-TeV c.m. the elements of the collimation system also must be relocated. The tunnel length allotted to collimation is adequate to collimate energy and the horizontal and vertical planes at both phases only one time at 1.5-TeV c.m. Since the centroid orbit of the collimation system differs only slightly from a straight line, it would be possible to allocate length at the end of the linac tunnel for collimation at energies above 1-TeV c.m. And it should be sufficient to collimate the FD phase only one time in the collimation system. See Section 9.2.3 for details.

9.1.1 Introduction to Collimation Systems

The post-linac collimation section (Section 9.2) begins with a discussion of system specifications since the design is strongly influenced by assumptions on incoming beam conditions: Of the 10^{12} particles per bunch train, how many

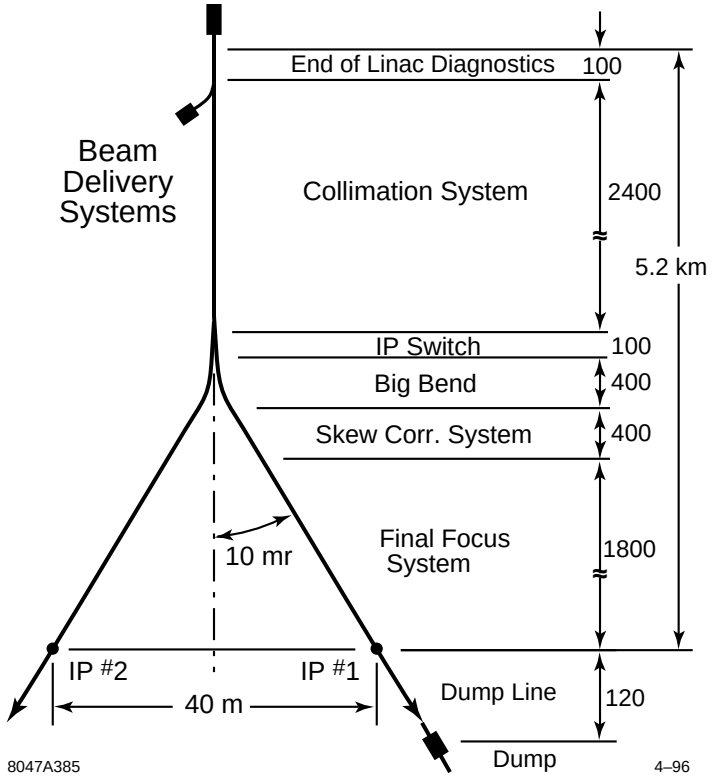


Figure 9-1. *A schematic layout of the beam delivery systems.*

particles must be routinely collimated, and how many particles can be safely collimated in the final-focus system? Because the number of particles needing to be collimated is difficult to predict precisely, since it depends on how well the upstream systems have been tuned, we have relied on SLC experience for this estimate. The number of particles that can be collimated in the final-focus system has been determined by edge-scattering and muon transport studies within the final-focus system. These studies are described in Chapter 12. The rough guideline that evolves from these considerations is that there may well be a few times 10^{10} particles in the beam tails at the end of the linac and this number needs to be reduced to a few times 10^{6} upon entry into the final-focus system.

Two types of collimation systems have been proposed: linear and nonlinear. In the former, the beam sizes at the collimator are achieved by traditional linear optics methods (quadrupoles); in the latter strong sextupoles are used to blow up the beam. Since the sextupoles are exceedingly strong and system lengths are not reduced in the specific system proposals we have studied, we have opted to look in depth at a linear collimation scheme. It is not precluded that a nonlinear (or combination linear and nonlinear) system could be found that would be operationally superior and have a lower total cost. Our primary objective is to show that at least one collimation system exists that fulfills all functional requirements.

Since a small perturbation in upstream conditions could cause a complete bunch train of 10^{12} particles to be incident on the collimators of the collimation system, it is necessary in both the linear and nonlinear systems to rely on a primary collimator that is a spoiler, followed by a secondary collimator which is the absorber. The spoilers must be thermally rugged and very thin ($\leq$1/4 radiation length). The best material we have found for spoilers is a titanium alloy plated with pure titanium or titanium nitride (TiN) for improved electrical conductivity. The function of the spoilers is to increase the angular divergence of the beam, so that when the beam arrives at the absorber it has a much larger size

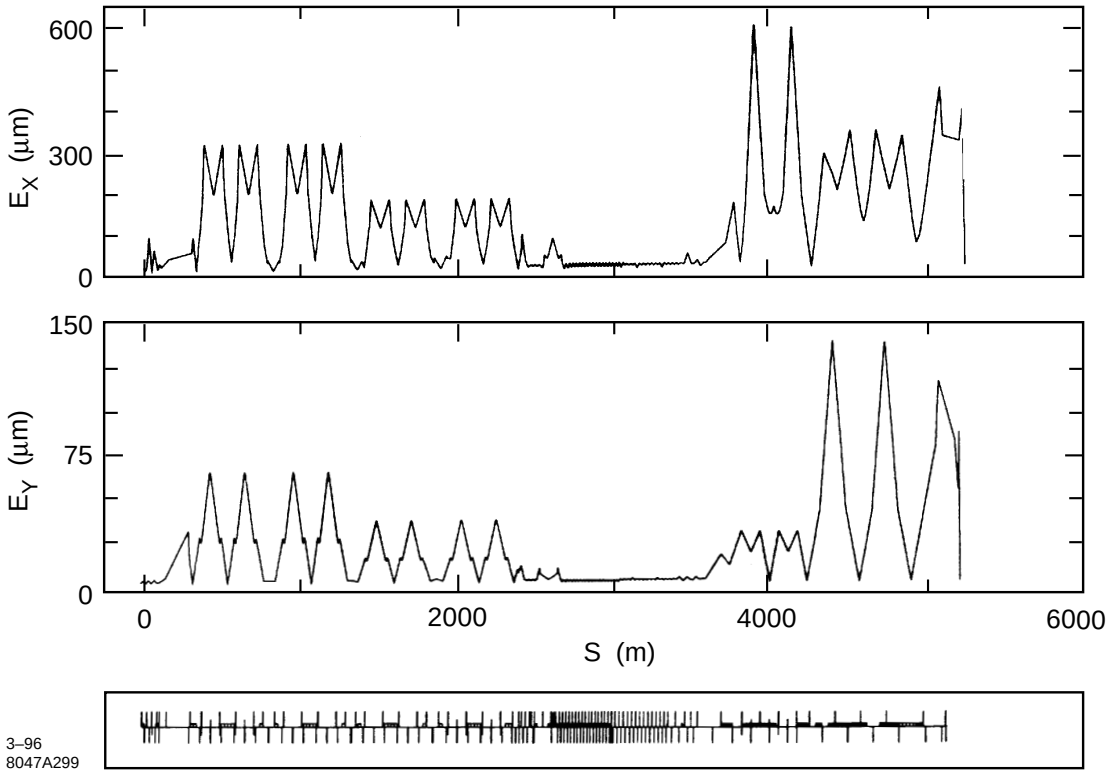


Figure 9-2. *The horizontal and vertical beam envelope from the linac to the IP for the 1-TeV-c.m. beam line.*

(millimeters). The absorber, on the other hand, must be able to routinely absorb and remove the energy in the tail of the beam. For the 1-TeV c.m. parameters, 1% of the time-averaged beam power is 84 kW. The preferred material for absorbers is copper.

The wakefields of the collimators can have a very deleterious effect on the beam core. To minimize the wakes, the beam pipe must be tapered before and after the collimator. Even for an on-axis beam core, a parallel-jaw collimator will have a quadrupole wake, which can influence focusing of the core and the trajectory of particles in the tails. And beams that have been mis-steered close to the wall can experience very large wake-induced kicks. All of these wake effects are described in Section 9.2.4.

We consider the geometric wakefields for tapered collimators to be uncertain. There are theoretical results which we will discuss below, but the small tapers and short bunch lengths have made these collimators very hard to simulate with existing numerical modeling codes. The geometric wake for the parallel-plate geometry has changed as this document was going to press. We have incorporated notes in the text describing the change. If the new result stands, because the geometric-part of the wake is now larger than the cylindrical wake by a factor of about 4, there would be an impetus to use the less-convenient cylindrical geometry collimators.

There has also been some doubt cast on the correctness of the resistive-wall wake formulae we have used for the tapered collimators. A recent measurement of collimator wakes at the SLC has given a wakefield kick that is stronger than expected. The situation indicates a need for clarifying experiments with tapered collimators.

Because of the large β-functions and strong focusing that arise when the beam is blown up with linear optics, there are important chromatic effects to compensate with sextupole pairs. And because it is necessary to collimate each

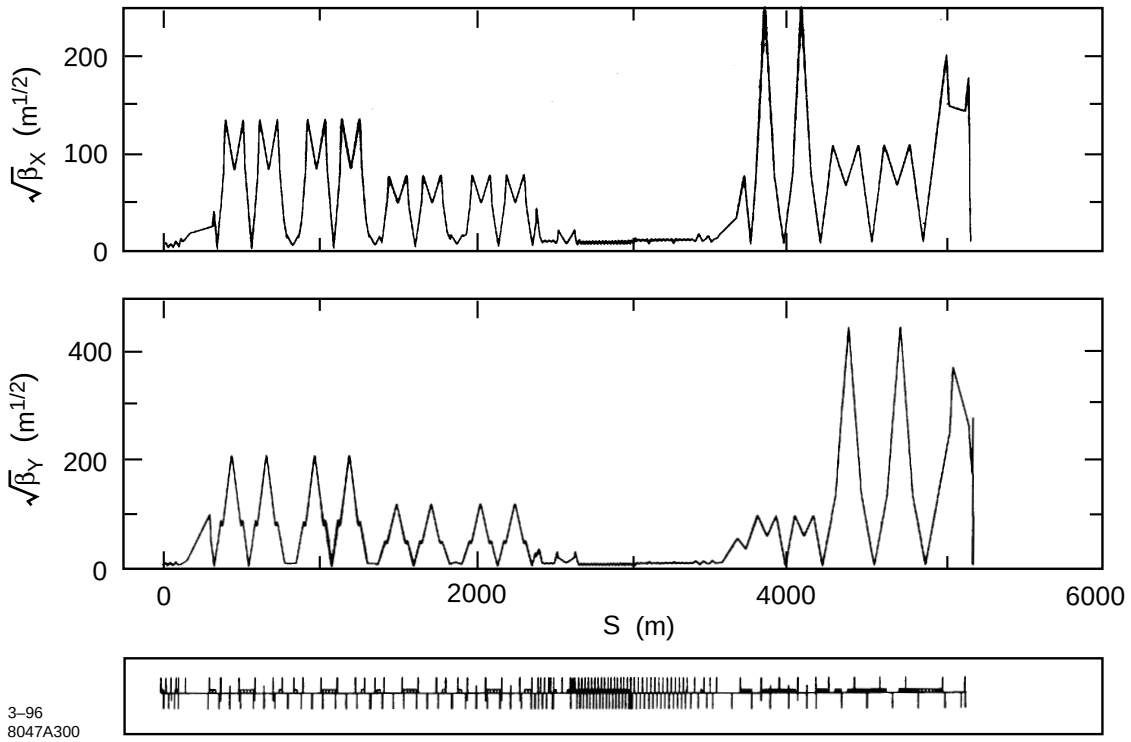


Figure 9-3. *The horizontal and vertical β function from the linac to the IP for the 1-TeV-c.m. beam line.*

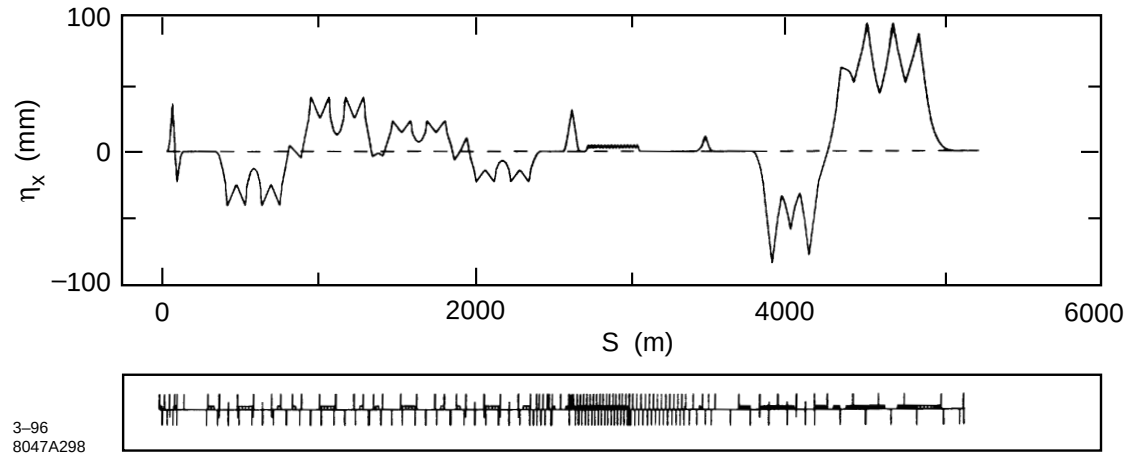


Figure 9-4. *The horizontal dispersion function from the linac to the IP for the 1-TeV-c.m. beam line.*

transverse phase at least one time, there are very large R_{12} and R_{34} functions within the system. Large R_{12} and R_{34} functions lead to important tolerances described in Section 9.2.5. Stability tolerances, that must be held between tunings of the waist knobs in the skew correction system, are looser than those within the final-focus system. However the particle backgrounds in the collimation system preclude the beam-based stabilization methods contemplated for the final-focus system. Alternatives are discussed in Section 9.2.6.

Vibration tolerances that must be met between orbit adjustments are greatly simplified by the ground motion studies described in Appendix C which, when applied to the collimation system, show that seismic ground motion has a negligible impact on beam collision offsets at the IP. As a result, vibration tolerances become tolerances between beam-line elements and the ground beneath them, or tolerances on ground motion coming from cultural sources.

9.2 Post-Linac Collimation

9.2.1 Specifications

Incoming Beam Parameter Range

The important parameters from the point of view of the collimation system are:

i) $Nn_b/\sqrt{(\varepsilon_x^N \varepsilon_y^N)}$, total bunch-train charge divided by the square root of the normalized emittance product, for single pulse-train spoiler survival;

ii) $N/(\varepsilon_x^N \sigma_z^{2/3})$ and $N/(\varepsilon_y^N \sigma_z^{2/3})$, for determination of minimum n_x^2 and n_y^2 that can be collimated in each plane (see Eq. 9.45); and

iii) the worst-condition population in tails, for average power loads on absorbers.

The strange units result from omission of a dependence on the surface resistivity of the collimators.

The range of IP parameters under consideration is given in Chapter 12, Table 12-1. Here one sees a maximum charge per bunch of $N = 1.25\ 10^{10}$, a maximum charge per bunch-train of $Nn_b = 1.125\ 10^{12}$, a minimum bunch length of 100 μm, a minimum horizontal normalized emittance of $\varepsilon_x^N = 4\ 10^{-6}$ m-rad, and vertical normalized emittances which vary from $\varepsilon_y^N = 8$ to $17\ 10^{-8}$ m-rad. The emittances at the end of the linac can be somewhat smaller and, depending on linac alignment conditions, could be as small as $\varepsilon_x^N = 4\ 10^{-6}$ m-rad and $\varepsilon_y^N = 5\ 10^{-8}$ m-rad.

Limit ii) implies that $n_x\sigma_x$ and $n_y\sigma_y$ are independent of energy and emittance. Thus if the downstream beta functions do not change, the physical apertures that may be collimated are independent of energy and incoming beam emittance.

The most sensitive parameter is that of item i), because the required value of the β functions at the spoilers depends on this ratio, and the system length increases as this β function becomes larger.

For evaluating worst conditions we assume the values $\varepsilon_x^N = 4\ 10^{-6}$ m-rad, $\varepsilon_y^N = 5\ 10^{-8}$ m-rad, and $N = 1.25\ 10^{10}$. The worst ratio $N/\sigma_z^{2/3}$, in the parameter range of Table 12-1, occurs when $N = 1.25\ 10^{10}$ and $\sigma_z = 150\ \mu$m. We also assume that a maximum of 1% of the beam may need to be collimated in any phase and any plane, or in energy. The total beam fraction collimated could be as high as 5%. This assumption is discussed in the next section.

Incoming Beam Halo Specification

Experience with the SLC has identified several sources of halo particles: those created

i) during extraction from the damping ring,

ii) in the bunch compressor,

iii) by wakefields within the linac,

iv) by mismatches, misalignment and steering errors in the linac,

v) from acceleration of dark current,

vi) from hard Coulomb and/or bremsstrahlung scattering within the linac, and

vii) from faulty multi-bunch energy compensation.

The first two items may be addressed by the pre-linac collimation system (Section 9.3). A bunch-length collimation system (Section 9.4) can reduce contributions from item iii). Linac diagnostics are intended to reduce iv). Since the injection energy into the linac is at 10 GeV, dark current will be considerably off-energy and presumably not accelerated far.

Item vi) cannot be eliminated, but the number of particles in the tail from this source can be derived [Yokoya 1991] beginning with the scattering angle

$$\Delta\theta = \frac{2Zr_e}{\gamma b} \tag{9.1}$$

of an electron passing at distance b from a nucleus with charge Z. For an azimuthal orientation of an angle ϕ specifying initial conditions, the kick in the horizontal direction would be $\Delta\theta_x = \Delta\theta\cos\phi$. For $\Delta\theta_x \geq n_x\sigma'_x$ we must have $b \leq b_n$ with

$$b_n(\phi) = \frac{2Zr_e}{\gamma n_x \sigma_{x'}}\cos\phi \quad . \tag{9.2}$$

Hence the cross-section for scattering beyond $n_x\sigma'_x$ is

$$\sigma_n = \int_0^{2\pi} d\phi \int_0^{b_n(\phi)} bdb = \pi b_n^2(0) \quad . \tag{9.3}$$

If ρ is the density of nuclei, then the fraction of particles scattered beyond $n_x\sigma_{x'}$ in length ds is given by

$$\frac{\Delta N}{N} = \pi\rho\int b_n^2(0)ds = \frac{2\pi r_e^2\rho Z^2}{n_x^2\varepsilon_x^N}\int\frac{\beta_x(s)ds}{\gamma(s)} \tag{9.4}$$

This can be integrated over the length of the linac. The β function oscillates between minimum and maximum values that scale roughly as the square root of energy (See Chapter 7, Figure 7-1 and 7-2. This oscillation is much more rapid than the change of γ so it may be replaced by an appropriate average, $\xi\beta_x^{\max}(s) = \xi\beta_x^{\max}(0)\sqrt{[}\gamma(s)/\gamma(0)]$. We may take $d\gamma/ds$ to be a constant and carry out an integral over γ to obtain

$$\frac{\Delta N}{N} = \frac{4\pi r_e^2\sum_i \rho_i Z_i^2}{n_x^2\varepsilon_x^N}\,\frac{\xi\beta_x^{\max}(0)L}{\sqrt{\gamma(L)\gamma(0)}} \quad . \tag{9.5}$$

We have included a sum over nuclear species, ρ_i being the density of nuclear species with charge Z_i. If the partial pressure associated with a nuclear species is P_{Gi} then the number of nuclei per unit volume is given by $\rho_i \approx 3\ 10^{22} n_i P_{Gi}\ m^{-3}$ where n_i is the number of nucleons per atom. If we take the major composition of gas within the beam pipe to be N_2 with nuclear charge $Z = 7$, and take $P_G = 10^{-8}$ Torr, $\gamma(0) = 2\ 10^4$, $d\gamma/ds = 10^2$, and $\beta_x^{\max}(0) = 10$ m, we obtain the estimate

$$\frac{\Delta N}{N} \approx \frac{1.4\ 10^{-7}}{n_x^2} \quad \text{or, in a similar manner} \quad \frac{\Delta N}{N} \approx \frac{1.4\ 10^{-5}}{n_x^2} \quad . \tag{9.6}$$

For $N = 10^{12}$, $\Delta N \approx 1.4\ 10^5/n_x^2$ or $\Delta N \approx 1.4\ 10^7/n_y^2$. For collimation at $n_x = 5$ or $n_y = 35$, $\Delta N < 10^4$. We will see below that this number of particles could be safely collimated in the final-focus system.

Bremsstrahlung interactions with beam gas can result in particle energy loss and create tails. The cross section for this process is given in Eq. 11.118. For more than 1% energy loss the cross section is estimated to be about $\sigma_{brems} \approx 6$ barn. This gives the estimate

$$\frac{\Delta N}{N} \approx 2\ 10^{-5} P_G L \tag{9.7}$$

where P_G is the gas pressure in Torr and L is the linac length in meters. For $N = 10^{12}$, $L = 10^4$, and $P_G = 10^{-8}$, we have $\Delta N \approx 2\ 10^3$, which is a smaller number than from Coulomb scattering.

Hence the principal source of concern is item iii), the tail particles generated by wakes in the linac. Compared to the SLC, the bunch-length collimation and pre-linac collimation should help reduce this number, but the number of particles in the longitudinal tails under typical (and worst!) operating conditions remains uncertain. Estimates indicate that tail populations should be less than 10^{-3} or even 10^{-4} [Raubenheimer 1995]. On the other hand, at the SLC it has been necessary on occasion to collimate 10% of the beam particles. When the systems are tuned this number is an order of magnitude smaller.

Assuming that the pre-linac collimation systems and other efforts to minimize tail particles will have a positive impact, we somewhat arbitrarily have adopted the specification that at most 1% of the beam will require collimation at any particular phase, transverse plane, or energy. The system to be described will actually have somewhat higher totals, up to 5%, depending on how beam tails are proportioned. The main impact of choosing a larger collimation percentage is in the size of the absorbers, and they could be designed to remove a higher (or lower) average power load if it was deemed necessary.

Required Collimation Depths

Final-doublet (FD) Phase Depth. The collimation depth at the final-doublet (FD) phase, $\pi/2$ from the interaction-point (IP) phase, is determined by beam emittances and the apertures and β functions in the final-doublet quadrupoles. The final-doublet β functions are determined by the IP β functions and the final-doublet design parameters: the free space from the last quadrupole to the IP and the strength of the quadrupoles.

Final-doublet designs get more compact and have smaller chromaticity for larger quadrupole gradients, so there is a preference for small apertures. But as we have mentioned in Section 9.2.1 above and shall show below in Section 9.2.4, given β in the doublet, there is a minimum beam aperture, $n_x\sqrt{(\varepsilon_x \beta_x)}$, that can be collimated which is determined by the ratio $N^{1/2}/\sigma_z^{1/3}$. Hence the final-doublet aperture is chosen as small as possible subject to the limits of the collimation system, the range of IP β functions which is to be accommodated, and resistive-wall and geometric wake considerations.

The IP β functions are adjusted to an optimum during operation. If the beam is free from tail particles, the optimum will be determined by the maximum luminosity, or equivalently the minimum IP spot size that can be obtained with

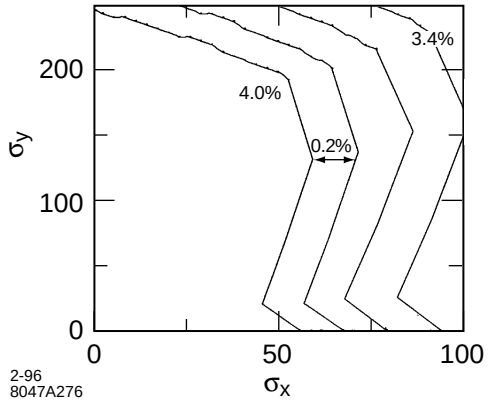


Figure 9-5. *The dynamic aperture of the final-focus system for the IP phase. The FD phase particles have been restricted to be within a diamond region in the $x - \delta$ plane defined by the collimation system, with $|x| \leq 7\sigma_x$ and $|\delta| \leq 0.04$.*

the given beam emittance. As the IP β functions are decreased, the linear spot size gets smaller, but the strengths of aberrations from nonlinearities in the final-focus system get larger. The values of the IP β functions shown in the parameter sets of Chapter 12, Table 12-1 will be close to the optimum. Synchrotron radiation, Oide effect and beamstrahlung control are also involved in finding the optimum.

Tail particles create problems in two ways (see Chapter 12 for details): i) by impacting an element of the final doublet, or ii) radiating photons which impact the final doublet. Studies of shower particles and their trajectories through the detector show that only tens of particles per bunch train can be allowed to impact the final doublet. Studies of synchrotron radiation lead to the guideline that the horizontal final-doublet aperture should be twice the collimated aperture, and the vertical aperture should be 1.5 times the collimated aperture.

The answer to the question, "What should be the final-doublet phase collimation depth?" turns out to be "As small as possible." As we shall see below, for the beam parameters under discussion, the minimum collimation apertures are $5\,\sigma_x$ and $35\,\sigma_y$. There needs to be some margin on this number, since the apertures must be gradually enlarged in successive collimation stages, and by the final doublet these minima may increase to $7\,\sigma_x$ and $40\,\sigma_y$. Following the synchrotron radiation guidelines, the final-doublets apertures should be chosen to be equal or greater than $14\,\sigma_x$ and $60\,\sigma_y$. Final-doublet designs are discussed in Section 11.6 and particle trajectories are shown in Section 11.6.5.

Interaction Point (IP) Phase Depth. The collimation depth for the interaction-point (IP) phase can be larger than the final-doublet (FD) phase since in the final-focus system the IP phase is demagnified, and apertures for this phase are more than $45\,\sigma_x$ and $200\,\sigma_y$. Figure 9-5 shows the dynamic aperture of the horizontal and vertical phases of the final-focus system determined from tracking.

There are several advantages of collimating the IP phase at a larger aperture than the FD phase.

i) The wakes will be smaller, and since these wakes give rise to jitter in the final-doublet phase, the beam jitter is reduced in the final-focus system, which is helpful for tuning.

ii) The β functions in the second stage IP-phase collimation can be chosen smaller. This means that the IP-phase collimation section can be shorter and the tolerances less severe.

The smaller β functions of item ii) can only be achieved in the first IP collimation stage if one abandons the principle that the spoiler should be able to survive the impact of a full bunch train. This needs to be carefully evaluated in the context of the machine protection system.

Just how large the collimation in the IP phase can be also depends on beam-line transport within the collimation system and from the collimation system to the IP. The beam line from the collimation system to the final-focus system must be achromatic, so that off-energy IP-phase particles do not migrate into the FD phase. It also requires that there not be significant residual beam-line aberrations from sextupoles or interleaved sextupoles, which distort the phase space of large amplitude particles. These question are addressed in Section 9.2.5.

There is a further constraint on IP-phase apertures that comes from machine-protection considerations. Mis-steered beams that come close to collimator apertures experience very large wake kicks, and can be steered into unprotected downstream absorbers. As we will see below, it is necessary to have two stages of IP collimation in the collimation system. The machine protection considerations alluded to will require that the second stage collimation have an aperture that is much larger than the first stage. These issues are discussed in detail in Sections 9.2.4, "Quadrupole and Large Amplitude Wakes of Parallel Plate Collimators", and "Consequences of Quadrupole and Near-Wall Wakes". We assume throughout that the collimators in the first IP stage are protected by spoilers that can withstand the impact of a full pulse train.

Energy Collimation Depth. The required energy collimation depth also depends on the properties of the lattice between the collimator and the final doublet. If an off-energy beam is able to pass through the spoilers, it should not be able to impact any downstream collimator in the collimation system or in the final-focus system. Figure 9-6 shows the result of tracking particles through the beam line from the end of the collimation system to the IP for an early beam-delivery lattice. The initial coordinates are chosen at the collimator edges of $6\,\sigma_x$ and $40\,\sigma_y$, in both the IP and FD phase, and tracked to the location of the horizontal and vertical collimators of the final-focus system. One sees in Figure 9-6 the image of these particles when they are off-energy by various amounts from 0 to -6%. Based on these and similar studies described in Section 9.2.5 we have chosen 4% as the energy collimation depth.

Because of the large β functions in the collimation system, there must be a chromatic correction scheme of sextupoles placed in dispersive regions [Brown 1979]. It is natural to take advantage of this dispersion function to collimate energy at the same locations where the horizontal plane is collimated. Though the magnitude of the dispersion function is somewhat flexible, 4% energy collimation turns out to be a possible and convenient energy depth.

Limits on Collimation in the Final-focus System

The CCX and CCY chromatic correction sections in the final-focus system provide excellent locations for collimators. However collimation here can create unacceptable muon backgrounds. Figure 9-7 shows the number of particles that may be collimated, and still have just one muon pass through the detector. If muon toroids are placed in the final-focus tunnel, one finds that 10^5 particles may be collimated in the CCY and 10^6 may be collimated in the CCX.

Figures 9-8 and 9-9 show the results of edge scattering within the final-focus system. These studies indicate that less than 1 particle in 10^5 of the particles that re-enter the beam at the collimator edge actually impact the final doublet. This simulation was done assuming the angular and energy edge-scattering profiles presented in Section 9.2.3. In fact no particles in this study impacted the final doublet. Either the particles impacted the beam line between the collimator and the doublet (at the latest in the bend in the final telescope) or they passed through the IP. Particles that are close to full energy will miss the doublet because the collimators are π from the final doublet and have a smaller aperture than the final doublet. About 10% of the particles are in this category. Particles that are off-energy don't make it through the bends. Since only one particle re-enters the beam for every 10 particles per micron incident on the edge of the collimator, this study would say that there is an allowed density of 10^6 particles per micron on these collimators. Since

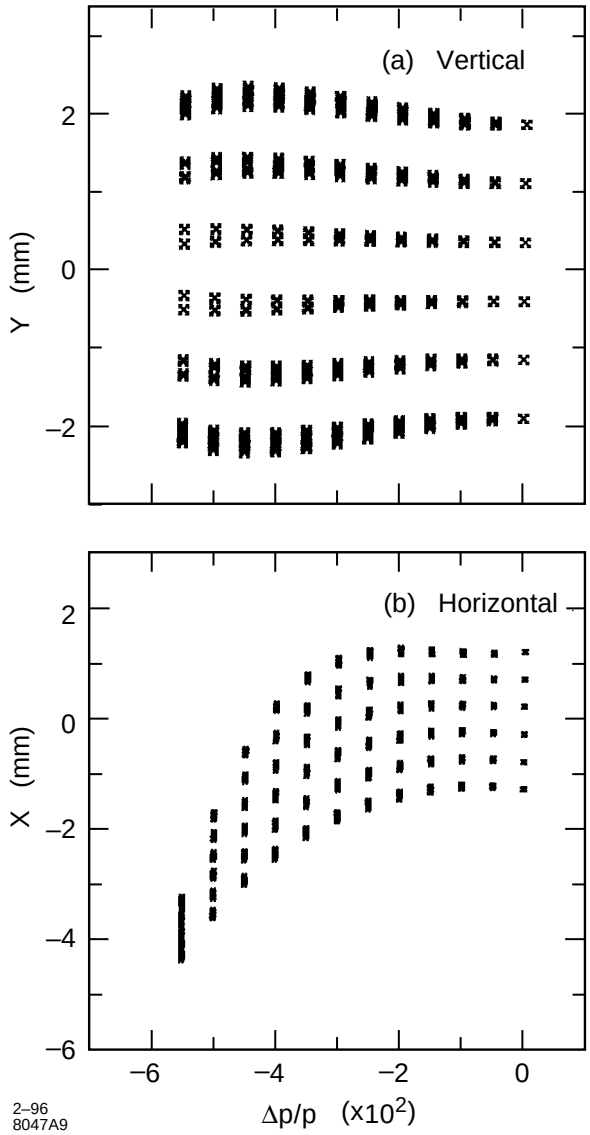


Figure 9-6. *The position of particles at the horizontal collimator in the final-focus system for particles launched in the collimation system on the boundary of the collimation aperture for several energy offsets.*

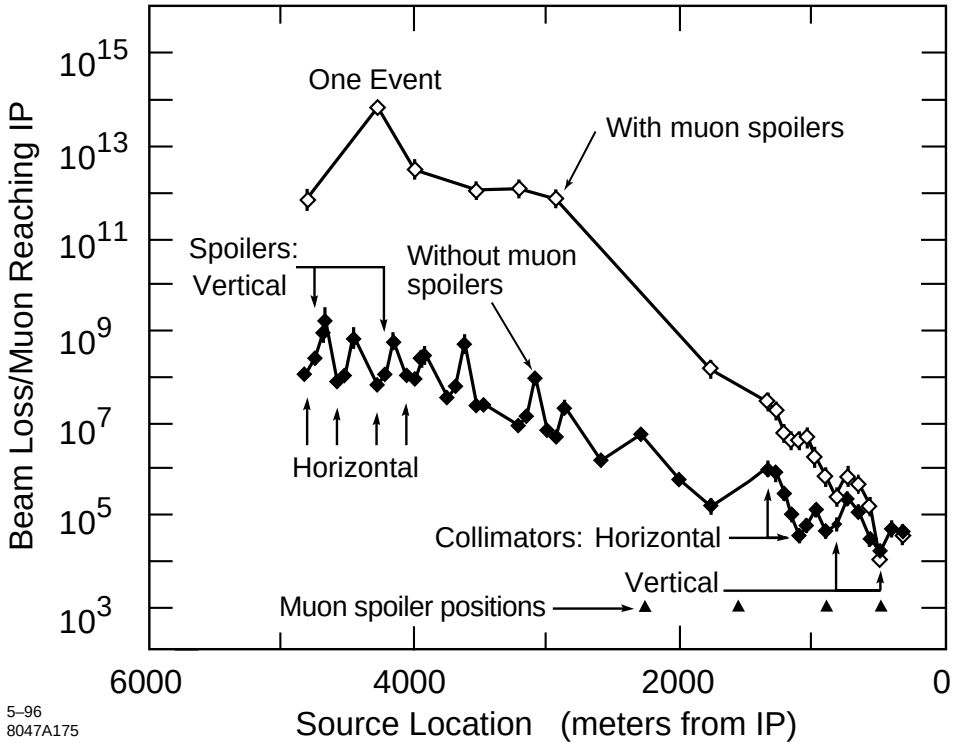


Figure 9-7. *Graph showing electrons that may be collimated if only 1 muon is to reach the detector.*

the muon studies conclude that only 10^5 to 10^6 particles may be incident on the whole collimator, the muon studies set the limits on the number of particles that may be collimated in the final-focus system.

Studies described in Chapter 12 indicate that tens of particles can be incident on the final doublet per bunch train without saturating the vertex detector. We conclude that the principal source of such particles will be from scattering on gas particles in the final-focus system. See Section 11.5.5, Eq 11.68 for an estimate of this source.

9.2.2 Materials Considerations

Machine Protection Issues

Single-Bunch Passive Protection. The machine protection strategy (see Chapter 16) consists of a two-level scheme. At level 1 the beam line should be able to passively survive the transport of a single bunch, no matter the state of any device in the beam line. At level 2, which assumes that a single bunch is known to successfully travel through the beam line to a beam dump, the emittance of the single bunch can be reduced and the number of bunches can be increased. It is now the responsibility of the machine protection system to detect any state changes in the system that could lead to the beam not being safely transported to a dump. Upon detection of such a state change it must be possible to turn off the beam before any component is damaged.

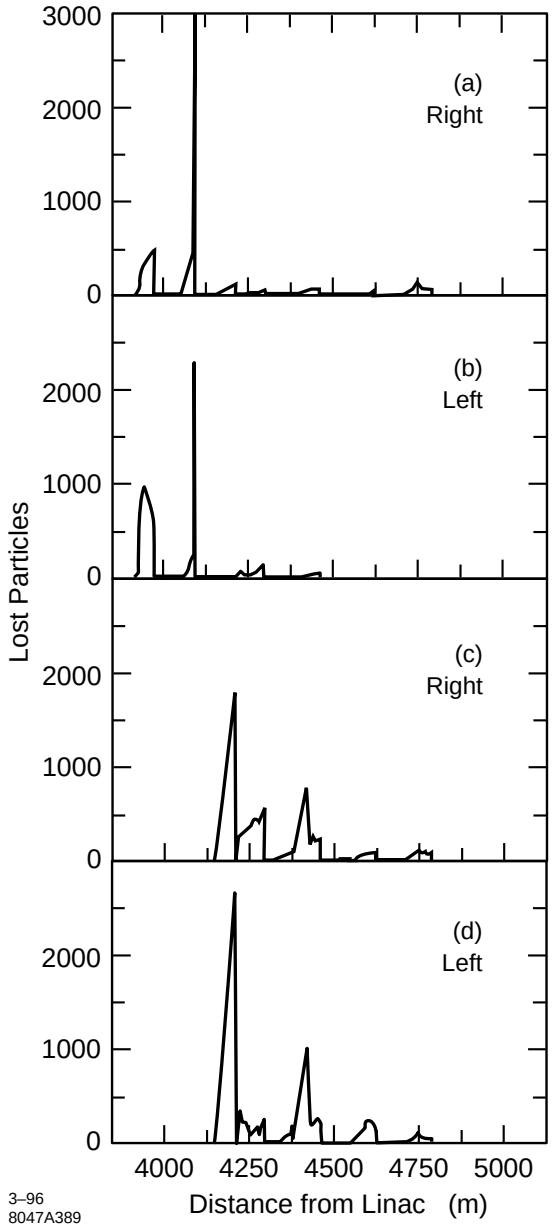


Figure 9-8. *Edge scattering results from the right and left edge of the two horizontal collimators in the CCX section of the final-focus system. In each case 100,000 particles were released from the edge according to angular and energy distributions determined from EGS studies (see Section 9.2.3). The curves give the number of particles which impact elements (or the beam pipe) along the beam line. In both cases no particles impact the final doublet or final-doublet region. About 10% of the particles pass through the IP.*

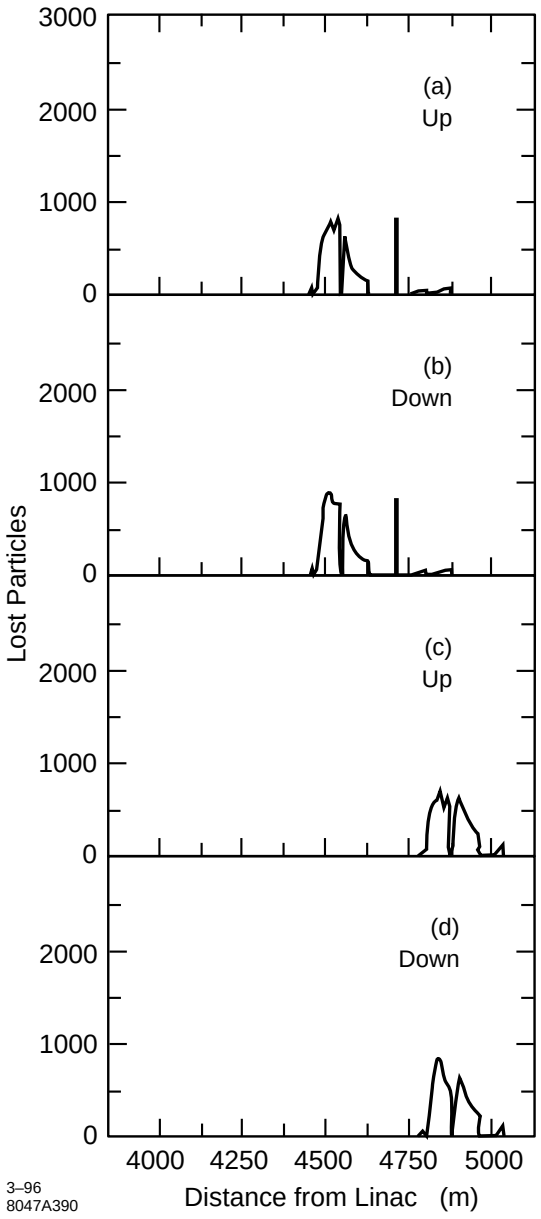


Figure 9-9. *The distribution of impact position for particles edge-scattered from the edge of the vertical collimators at the sextupoles in the CCY region of the final-focus system. A total of 100,000 particles were released from the collimator edge according to the angle and energy distributions shown in Figures 9-19–9-21. None of these particles impact the final doublet. About 10,000 pass through the IP. "Up" and "down" are from the top and bottom edges, respectively.*

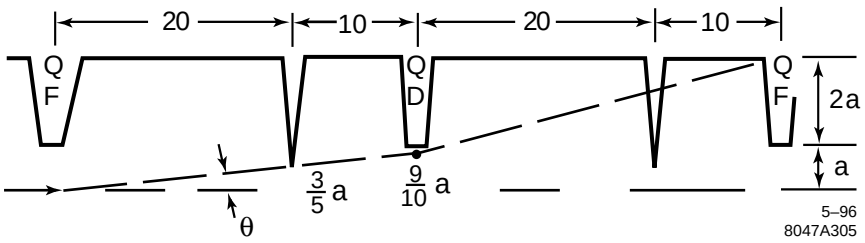


Figure 9-10. *Possible trajectories of a single bunch through a beam line as a result of completely or partially shorted quadrupole legs. Every trajectory encounters a spoiler at least 10 m before reaching the wall or another element. The quadrupoles are assumed to be spaced at 30 m and have a focal length of 20 m in this example.*

For a single bunch, the area of the beam when encountering a copper element must have a radius of about 100 μm (see Table 9-2). Since the divergence angle following a spoiler is about 15 μr (see Figure 9-3), the distance from the spoiler to any copper element must be equal or larger than 7 m or 10 m for a 500 or 750-GeV beam energy, respectively. The beam radius at a spoiler, assuming it is composed of 1/4 r.l. titanium, must satisfy $\sqrt{(\sigma_x\sigma_y)} \geq 10$ μm. Since there are many locations where $\sqrt{(\sigma_x\sigma_y)} < 10$ μm, the single-bunch beam emittance must be enlarged for level 1 beam-line check out. This is accomplished by blowing up the beam emittance in or at the exit of the damping ring.

Level 1 objectives are achieved by placing spoilers throughout the system, so that no matter what the trajectory of the bunch, it impacts a spoiler before any other element. Figure 9-10 shows a possible single-bunch trajectory, assuming a worst-case scenario of a shorted quadrupole with an on-axis field equal the design pole-tip field followed by a quadrupole which is defocusing in the plane of the kick. In this example the distance between quadrupoles is taken to be 30 m, and the focal length of the quadrupoles taken to be 20 m. If a spoiler, with a gap of 3/5 a, is placed 10 m before each element, then assuming that the distance between quadrupoles is $d = 3/2f$, where f is the focal length of the quadrupoles, the worst case trajectory has a displacement of 9/10 a at the second quadrupole and would have a displacement of 3 a at the third quadrupole. Thus if the beam pipe between quadrupoles has a radius greater than 3 a, all beam-line elements are protected at level 1.

We take as a guideline that the aperture at the beam-line elements is 2.5 times the gap of any nearby spoiler or absorber of the collimation system. Since the largest aperture of the spoilers or absorbers in the first collimation stage is 2 mm (see Section 9.2.5), it is consistent with this guideline to assume the element aperture is 5 mm. It follows that the beam-pipe inner radius between beam-line elements must be larger than 15 mm.

Under the conditions envisaged in Figure 9-10, when a bunch passes close to a spoiler edge it can receive a wake-field kick from the spoiler. The magnitude of these kicks is calculated in Section 9.2.4. The maximum kick goes like $1/\sigma^2$ where σ is the largest axis of the beam ellipse at the spoiler. Assuming the smallest major axis within the first stage of the collimation system is 100 μm, this maximum kick is about 150 nr. Between collimation sections, the beam size is much smaller and the kick can be 100 times larger, or 15 μr. This is still much smaller that the kick angles from the quadrupoles, which in Figure 9-10 are as large as 250 mr. Thus these wakefield kicks will not compromise the level 1 protection scheme we have described.

A complete treatment of this problem would involve a study of each beam-line segment for a complete list of assumptions on failure modes. These sections on machine protection are intended as a guide to the nature of a solution, and are not intended to represent a complete study.

Since the machine-protection spoilers have apertures that are not much larger than the system spoilers and absorbers, it will be important that they are tapered to minimize the wakes. (See Section 9.4.5.) The machine-protection spoilers will look very much like the system spoilers. Since the effect of the wakes from a taper go as $\beta/g^{3/2}$, where g is the spoiler gap, only spoilers at large β will have a significant impact. The wake from the largest-β machine-protection

spoiler, if constructed according to the above guidelines, would be roughly 0.3 of the wake from the main system spoiler. All machine-protection spoilers taken together can add an estimated additional 10% ($t'/t = 0.07$) to the jitter amplification budget (see Section 9.2.4). This is large enough that it should be reviewed carefully and taken into account when choosing the vertical collimation depth.

Multiple-Bunch Active Protection. For a full bunch train the machine must be turned off before the beams collide with any of the machine protection system (MPS) spoilers, since the full train would easily destroy them. Let us suppose as a worst-case scenario that a quadrupole leg is suddenly shorted. Let us suppose that the on-axis field has reached the fraction ξ of the pole tip field. This will give an angular deflection $\theta = \xi a/f$, where a is the quadrupole aperture and f is its focal length. The beam displacement downstream will be $\Delta x = \theta_x R_{12} \leq \xi R_{12} a/f$. Since the R_{12} values can be the order of $2\ 10^4$ m, for $f = 20$ m we would have $R_{12}/f = 10^3$. This means that the machine must turn off before $\xi = 10^{-3}$. Assuming the fastest rate at which the magnet can change field values is the order of 100 milli-seconds, it could change its value by 10^{-3} in 0.1 ms. Since the pulse-repetition rate is 2 ms, the magnets could change an unacceptable amount between beam pulses. Thus this error must be detected with a magnet-monitoring system, and can not be prevented by a BPM system monitoring the beam trajectory. We will assume such a magnet-monitoring system is in place.

If the magnet-monitoring system were to fail, a spoiler could be destroyed. In that case, the beam would not be sufficiently enlarged upon reaching the next beam-line element to protect it from damage. It may therefore be worthwhile to place sacrificial copper absorbers in the beam line in front of each element whose purpose would be to absorb the impact of the full beam. They would need to be replaced after a protection system failure event, but would limit damage to the system.

Multiple-Bunch Passive Protection for Collimators. There are many upstream system-state changes that could cause the beam in the collimation system to impact the main collimators. Major changes will be handled by upstream machine-protection systems, but because the β functions must be large enough to collimate particles that could impact elements in the final-focus system, the change in orbit will be larger in the collimation system than in any upstream system. And since everything in the final-focus system will be in the shadow of the collimation system, the collimation system will be the most vulnerable and important to protect.

The strategy of relying on monitoring all possible state changes in upstream systems becomes a very complex and potentially unreliable task. It seems unwise to assume the collimation system can be actively protected if collimation is to occur at apertures as small as $5\,\sigma_x$. Thus we have assumed that the collimation system collimators must be passively protected. (This decision needs to be carefully reviewed as the machine protection system becomes more mature. The cost impacts are large.)

Passive protection has major consequences, and implies that whatever the material, it must be able to withstand the full impact of at least one full bunch train. In the system to be described below, this is accomplished by having a sufficiently large linear beam size at the spoiler so that it can withstand this impact. Another scenario, which uses sextupoles to blow up the beam size, is discussed in Section 9.2.8. Other schemes, such as liquid-metal collimators and Compton collimators have been proposed. Though we have discussed such systems, we have not pursued them in great depth, preferring a more conservative solution. Since the system we describe is quite expensive, thorough investigations of alternative systems are encouraged.

Properties of Collimator Materials

For the 1-TeV c.m. parameters, each beam contains an average power of 8.4 MW. If routine collimation of 1% of the beam is adopted as a system specification, the energy intercepted at any phase could be as high as 84 kW. Half of this

Property	Unit	Carbon	Copper	Ti-6Al-4V	Tungsten-Rhenium
Density	(gm/cm^3)	2	8.95	4.5	19.3
Radiation length	(cm)	18.9	1.44	3.77	0.344
Resistivity	(μohm-cm)	1000	1.7	180	5.5
Specific heat	Joules/(gm°C)	0.97	0.385	0.54	0.134
Thermal conductivity	Watt/(cm°C)	2.0/0.025	3.9	0.17	1.26
Stress limit	(°C)	2500	180	770	700
Melting point	(°C)	3600	1080	1800	3100
Vaporization temperature	(°C)	4200	3000	3260	5000

Table 9-1. *Properties of materials considered for spoilers and absorbers.*

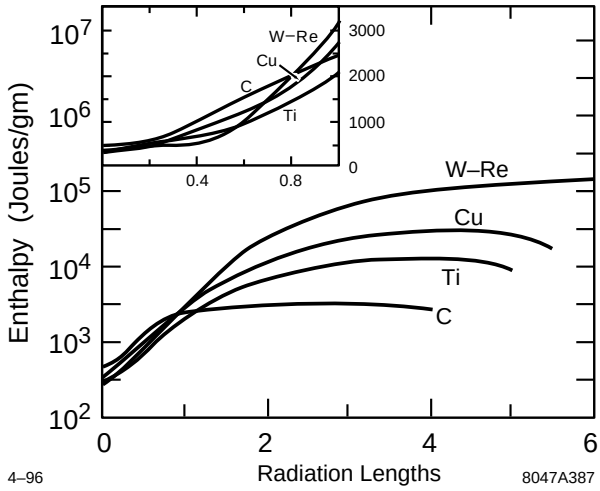


Figure 9-11. *Energy density deposition in four materials for an incident beam of 10^{12} particles with cross section given by $\sqrt{(\sigma_x\sigma_y)} = 100\,\mu$m.*

amount could be incident on a single absorber. Devices capable of absorbing and dispersing such heat loads are most often made from copper because of its high heat conductivity, compatibility with water, reasonable size, and general adaptability. See Table 9-1 for a list of some relevant physical properties for copper and some other materials we have considered for spoilers and absorbers.

Figure 9-11 gives the energy-density deposition in three of these materials from a beam of 10^{12} particles with initial cross-sectional area given by $\sqrt{(\sigma_x\sigma_y)} = 100\,\mu$m.

For other beam intensities and areas the curves of Figure 9-11 may be scaled by $N/(\sigma_x\sigma_y)$ because in the distances under consideration, showers have little chance to broaden beyond the initial spot size. The multiple-scattering formula (see Eq. 9.8) yields about 30 μr at $E = 500$ GeV and $t = 1$ r.l., so the initial spot size of 30 μm would not significantly change in several radiation lengths.

Since the energy is deposited instantaneously without any opportunity for the volume to change, the energy-density deposition represents a change in enthalpy. Figure 9-12 gives the temperature rise versus enthalpy for the same four materials, incorporating the changes of specific heat with temperature.

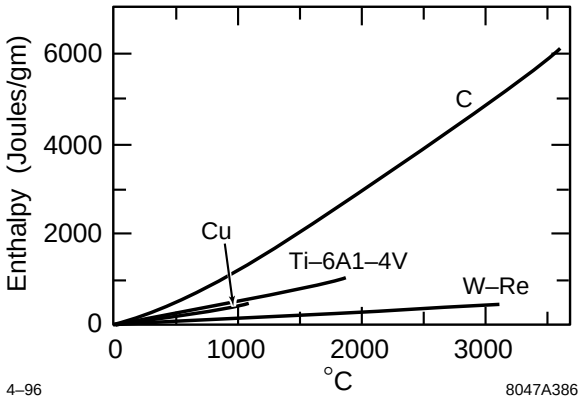


Figure 9-12. *Temperature rise versus enthalpy for four materials.*

	Fracture 2 cm from a Surface			Melt Anywhere within Volume	
Material	Enthalpy (Joules/ gm)	Incident Surface Minimum Sigma (mm)	Parallel Surface Minimum Sigma (mm)	Enthalpy (Joules/ gm)	Minimum Sigma (mm)
Carbon	3870	0.036	0.090	6000	0.072
Titanium alloy	405	0.14	0.56	960	0.36
Copper	63	1.0	2.2	470	0.82
Tungsten-rhenium	91	4.2	42.	410	2.0

Table 9-2. *Table of minimum $\sqrt{(\sigma_x \sigma_y)}$ to avoid fracture near the surface or melting of a collimator. It is assumed that there are 10^{12} particles per pulse train, and for the melt limit a particle energy of 500 GeV.*

For beams in the regime of 1-mm^2 cross section, it has been found experimentally that when a sudden local temperature rise creates local thermal stresses near the surface which exceed the tensile limit of the material, the material will probably fracture. If the temperature of copper rises suddenly by a mere 180°C, the tensile limit is exceeded (see Table 9-1). Hence for a single pulse of 10^{12} particles to give a temperature rise of less than 180°C anywhere within 2 cm of the surface, Figures 9-11 and 9-12 indicate the beam area must satisfy $\sqrt{(\sigma_x \sigma_y)} \geq 2.2$ mm. These limits for the remaining materials in Table 9-1 are shown in Table 9-2 [Walz 1973]. A beam incident on an absorber will always be within 2 cm of the collimation surface, so the temperature of the entire shower within the material must remain less than 180°C. The numbers for the incident surface apply only if the device is a dump. The fracture limit does not depend on beam energy. The melt limit is derived assuming a beam energy of 500 GeV.

To achieve 1-mm beam areas with the contemplated design emittances would require $\sqrt{(\beta_x \beta_y)} > 10^6$ m. Such β functions are impractical, and the usual strategy to achieve passive protection of collimator systems is to have a spoiler followed by an absorber [DeStaebler-Walz].

Material	Minimum Sigma (μm)
Carbon	50
Copper	280
Titanium	100
Tungsten-Rhenium	225

Table 9-3. *Table of minimum $\sqrt{(\sigma_x \sigma_y)}$ to insure survival of a 1/4 r.l. spoiler for a single pulse of a beam containing 10^{12} particles.*

Properties of Spoilers

Spoiler-Absorber Strategy. The spoiler is fabricated from a thermally-rugged, optically-thin low-Z material so that a full shower does not develop in the material. The requirement is that the temperature rise caused by a full bunch train incident on the spoiler not lead to fracture or melting. The spoiler increases the angular divergence of the incident beam so that down-beam copper absorbers can withstand the impact of one full bunch train. Additionally the absorber must be able to continuously absorb all particles hitting the spoiler even when this is as much as 1% of the beam.

Linear modules may be used to create the large β functions to achieve the required beam sizes at the spoilers, but β enlargement has two limitations: i) creating large β functions takes a long system length, and ii) wakefield kicks from the spoilers become large compared to decreasing angular beam divergences. We show below that because of ii) there is a limit to which a beam may be collimated with a linear β enlargement.

The large β functions create chromaticity which will increase the beam emittance unless locally compensated. This suggests the use of sextupoles in $-I$ configurations in each collimation phase and each transverse plane. Energy must be collimated first, since the bends employed to generate dispersion at the sextupoles could steer an off-energy beam into the wall of the beam pipe.

Single-Train Spoiler Survival. The minimum sigma required for a 0.25 r.l. spoiler are shown in Table 9-3. The best material we have found for use as a spoiler is pyrolytic graphite. However because it is porous and has a high resistivity, it must be plated. Even thin platings will separate from the plated material or melt if the beam size is not larger than the values which are derived from considerations of the energy deposition and temperature rise curves of Figure 9-11 and 9-12.

We have concluded that plated pyrolytic carbon spoilers are a bad idea for three reasons:

i) Carbon has not been used before, and there are questions about its performance under vacuum.

ii) Though it reportedly can be plated, details and experience are scarce.

iii) The plating, even if quite thin, will indeed be damaged unless the entry spot size is large enough to prevent melting and stress damage of the plating material.

From item iii) one concludes one might as well make the spoiler from the materials considered for the plating.

The next best material is titanium. Titanium also has a rather large resistivity when compared to good conductors (the resistivity of pure titanium is 45 $\mu\Omega$-cm, and the usual titanium alloy has a resistivity of 180 $\mu\Omega$-cm, compared to 1.7 $\mu\Omega$-cm for copper), but the wake of an angle-optimized taper goes as the 1/4th power of the resistivity (see

Section 9.2.4), so it is not out of the question to use titanium, especially if one uses the alloy plated with the pure titanium. TiN which has a resistivity of $22\mu\Omega$-cm may also prove useful as a plating material. The flat section of these spoilers will be very short since we want the longest path of particles through the material to be 0.25 r.l. (< 1 cm). In fact, as can be deduced from formulae presented in Section 9.2.4, the wake from one of these titanium spoilers is identical to the wake from a 20 r.l. Cu tapered absorber for a gap of 1 mm.

Because as indicated in Table 9-3, the entry spot size required for titanium must satisfy $\sqrt{(\sigma_x\sigma_y)} \geq 100\ \mu$m, whereas for carbon the requirement was $\sqrt{(\sigma_x\sigma_y)} \geq 50\ \mu$m, the β function product $\sqrt{(\beta_x\beta_y)}$ at the spoilers must be larger by a factor of four. The collimation system length increases by about 500 m for titanium spoilers as compared to carbon.

To keep the path length within a tapered spoiler to 0.25 r.l. radiation length, we propose fabricating it from 50-μm thickness rolled titanium which is then supported by a titanium honeycomb structure. Assuming a taper angle of about 20 mr, the path length in the incoming and outgoing taper will be 2.5 mm each. The rest of the 1-cm 0.25 r.l. thickness in titanium, can be in the honeycomb support structure. The honeycomb should be oriented at 45° to the line-of-flight of the particles.

The longest path length will occur at the tip of the spoiler. If the radius of curvature at the tip is 0.5 m, then the longest path in the titanium will be 1/4 r.l. Edge scattering is not an issue for spoilers, so the radius of curvature can be this small. A sketch of such a spoiler is shown in Figure 9-24(b).

To obtain the required spot size at the spoiler we must have $\sqrt{(\beta_x\beta_y)} \geq 10^{-8}/\sqrt{(\varepsilon_x\varepsilon_y)}$. At 500-GeV c.m., the smallest possible emittance product (see Section 9.2.1) at the end of the linac is $\sqrt{(\varepsilon_x\varepsilon_y)} = 10^{-12}$, and thus we must have $\sqrt{(\beta_x\beta_y)} \geq (100)^2$ m. For 1-TeV c.m., the smallest possible emittance product, $\sqrt{(\varepsilon_x\varepsilon_y)} = 0.5\ 10^{-12}$ m-rad, implies $\sqrt{(\beta_x\beta_y)} \geq (140)^2$ m. For 1.5-TeV c.m., the smallest possible emittance product, $\sqrt{(\varepsilon_x\varepsilon_y)} = 0.33\ 10^{-11}$, implies $\sqrt{(\beta_x\beta_y)} \geq (170)^2$ m.

Spoiler Heating. Each electron passing through the spoiler will deposit an energy of 1.8 MeV/cm^2. For the titanium alloy with a density of 4.5 g/cm^3, this amounts to an energy deposit of 8.1 MeV/ cm per electron. Since the path length of each electron in the titanium is in fact about 1 cm, this is an energy deposit of 8.1 MeV per electron. For 10^{10} electrons per train and 180 trains per second this amounts to 2.3 W. If we assume that the spot hitting the titanium has a cross-section of a semi-circle with a radius of 100 μm, then the surface of the spoiler, if straightened into a plane would be intersected in an area of width 200 μm and length equal to $d = 1\,\text{cm} + 2(100\ \mu\text{m}/\theta) \approx 3$ cm, where θ is the taper angle. Heat can flow both directions toward the walls, which we will assume is a heat sink at room temperature about 5-mm away. Since the length of the spot is six times larger than the distance to the sink, we assume a simple one-dimensional parallel heat flow. Using the thermal conductivity for titanium of 0.17 W/(cm-°C), we find a temperature elevation at the center of 230°C. If the 10^{10} electrons had been distributed between the two jaws, or the semi-circle had a radius of 0.2 mm, the temperature rise would be only 115°C. In any case, these are acceptable temperature rises in titanium.

Spoiler Transmission. Figure 9-13 shows an EGS simulation [Nelson 1993] of the angular distribution for several 1-GeV energy bins of particles that have passed through a 1/4 r.l. spoiler. The angular distributions are remarkably similar for all energies, peaking at about 14 μr. The solid line through the particles with energy in the 499–500-GeV bin is an expected distribution based on multiple scattering theory given by

$$\frac{dP}{d\,\Omega} = \frac{1}{2\pi\theta_0^2} e^{-\frac{\theta^2}{2\theta_0^2}} \qquad \text{with} \qquad \theta_0 = \frac{13.6}{E(\,\text{MeV})}\sqrt{t}\,(1 + 0.038\ln[t]) \quad . \tag{9.8}$$

$\theta_0 = 14\,\mu$r is, respectively, 10^3 or 10^4 times larger than the horizontal or vertical angular beam divergence at the spoilers. Therefore in order for particles coming from the spoilers to be incident on the edges of downstream collimators, they must have $\theta \ll \theta_0$.

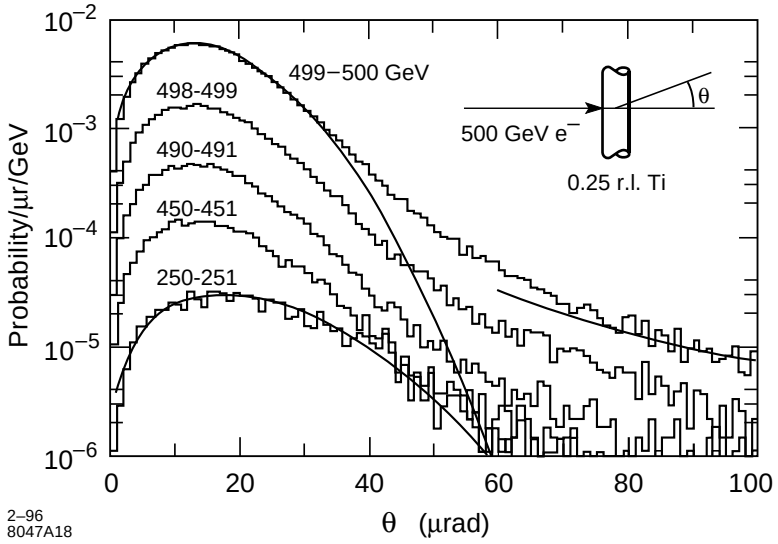


Figure 9-13. *Angular distribution of 500-GeV electrons that have passed through a 0.25 r.l. spoiler. The solid line is the theoretical prediction given in Eq. 9.8.*

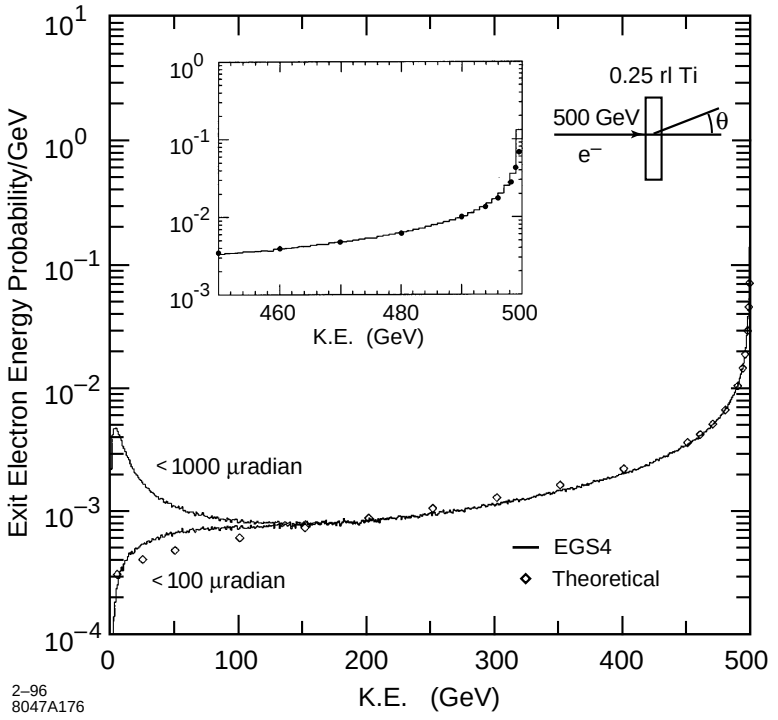


Figure 9-14. *The energy distribution of electrons from a 500-GeV beam that have passed through a 0.25 r.l. spoiler.*

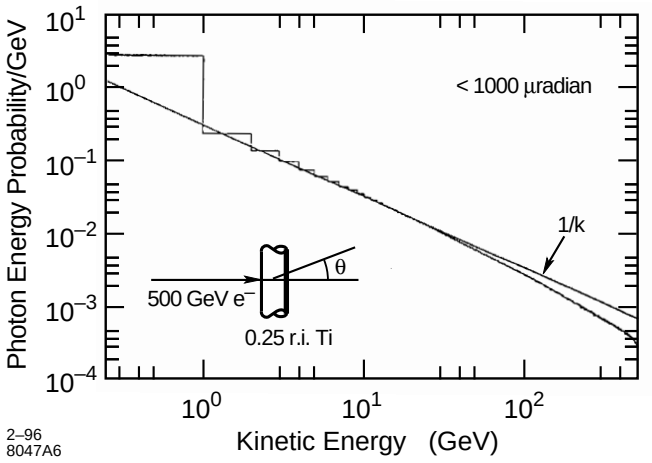


Figure 9-15. *The energy distribution of photons created by a 500-GeV beam of electrons passing through a 0.25 r.l. spoiler.*

Figures 9-14 and 9-15 show the energy distribution of exit electrons and exit photons, respectively, from 500-GeV electrons incident on a 1/4 r.l. spoiler. The energy distribution of the photons is fit very well by a 1/k law. The energy distribution of the electron is well fit by a formula based on single-scattering which is [Rossi 1952]

$$\frac{1}{N}\frac{dN}{dE} = \frac{1}{E_0\Gamma\left(\frac{t}{\ln 2}\right)}\left[\ln\left(\frac{E_0}{E}\right)\right]^{\left(\frac{t}{\ln 2}-1\right)} \quad . \tag{9.9}$$

The diamonds in Figure 9-14 are calculated using this equation. Figure 9-14 has an insert which is a blow-up of the region from 450 to 500 GeV. The dots in the insert are also calculated with the theoretical formula (Eq. 9.9). For small values of $\delta = \Delta E/E$, Eq. 9.9 becomes

$$\frac{1}{N}\frac{dN}{d\delta} \approx \frac{1}{\Gamma(1.44t)}\frac{1}{\delta^{1-1.44t}} \quad . \tag{9.10}$$

With the energy and angular distribution known, and the angular distribution independent of energy, it is straightforward to initialize distributions of particles that have passed through spoilers for the purpose of tracking them through the lattice. See Section 9.2.5, Figures 9-39 and 9-41.

Required Spoiler-Absorber Separation

If a beam is completely mis-steered and impacts a spoiler, the beam sigma product at the down-stream copper absorber should, according to the results of Section 9.2.2 and Table 9-2, be greater than $(2.2\,\text{mm})^2$. The shape of the image on a down-beam absorbers of the particles exiting a spoiler will be an ellipse. The horizontal and vertical axes will have Gaussian distributions with rms width equal to $\theta_0 R_{12}$ and $\theta_0 R_{34}$ m, respectively. Thus we have the condition $\theta_0^2 R_{12} R_{34} \geq (2.2\,\text{mm})^2$. For $\theta_0 = 14\,\mu$r we deduce $\sqrt{(R_{12}R_{34})} \geq 160\,\text{m}$.

This criteria can be eased somewhat by realizing that, on the average, the beam has lost one quarter of its energy when passing through the 1/4 r.l. spoiler, so that when passing off-axis through a quadrupole between spoiler and absorber the beam is spread out. This effect is not as large as the average energy loss would indicate since the distribution is

peaked at small energy loss. For example, only 25% of the particles have lost more than one-third of their energy. The exact magnitude of this effect will depend on lattice details, but if we assume that 25% of particles are not contributing to the core heating within the absorber, the condition of the last paragraph is modified to $\sqrt{(R_{12}R_{34})} \geq 140\,\mathrm{m}$.

If there is a quadrupole of inverse focal length f at a distance L_1 from the spoiler, followed by a drift of length L_2 to the copper absorber, then the $R_{12}R_{34}$ product is given by

$$R_{12}R_{34} = (L_1 + L_2)^2 - \left(\frac{L_1 L_2}{f}\right)^2 \quad . \tag{9.11}$$

Using Eq. 9.11 the current lattice does not satisfy $\sqrt{(R_{12}R_{34})} \geq 140\,\mathrm{m}$, since this lattice was designed for a lower beam intensity and a less stringent fracture criteria. A next-generation lattice under development does satisfy this condition. We will discuss future research directions in Section 9.2.9.

9.2.3 Tail Re-Population Estimates

Edge Scattering from Absorbers

Edge scattering from a spoiler is not a concern since very few particles will be incident on the edge (which, we will see below, has an effective width of about 1 μm) and, in any case, all the particles which pass through the spoiler are scattered at angles the order of 14 μr, which is similar to edge scattering. However at the absorbers, where one is hoping to clean up the beam, it is the edge scattering which re-populates the beam.

Figure 9-16 shows the results of an EGS simulation [Nelson 1993] to determine the fraction of particles that re-enter the beam as a function of distance form the edge of a flat titanium collimator edge (Note: We are expecting to use copper, but the curve can be readily scaled for all materials as shown in Figure 9-17). Only particles losing less than 4% of their energy have been counted because particles losing more energy will be collimated elsewhere. Later we present distributions with energy loss up to 20%. There is an abrupt edge in penetration at about 0.1 μm (note logarithmic scale) establishing this as the relevant edge-depth scale for titanium. To insure that the surface of an absorber with a length of a several cm be placed parallel to the beam with an accuracy of better than 0.1 μm would require an unlikely angular-placement accuracy of 1 μr. Without this placement accuracy particles in the incident beam would hit only a corner of the absorber and a larger number would be scattered back into the beam.

To avoid this effect, it has been usual to use collimators with curved surfaces [von Holtey]. Figure 9-16 also contains a curve showing the penetration probability for a curved titanium spoiler with a 100-m radius of curvature, which would require an easier 0.2-mr orientation accuracy. The transmission of the 100-m curved surface is increased by about a factor of 7, with a transmission edge at 0.7 μm. A third curve on this graph shows the edge scattering for a surface with a radius of curvature of 20 m, which would require an orientation accuracy of 1 mr. This has a transmission edge at about 3 μm, with a transmission that is a factor of 4 worse than the 100-m curvature. In the SLC collimator radii are about 10 m. The eventual flattening of these transmission curves at large offsets occurs because the EGS study was performed for a 2 r.l. absorber. If it had been thicker the curves would continue to fall.

Figure 9-17 compares the penetration probability of a curved titanium collimator with a curved carbon collimator, with the abscissa now chosen to be micro-radiation lengths (which we will denote by μrl) rather than microns. In terms of micro-radiation lengths the collimators perform identically. Since the radiation length of titanium is 3.77 cm, whereas copper's radiation length is 1.44 cm, the knee for a 100-m curved copper absorber occurs at 0.3 μm. For a 20-m curved copper absorber the knee would be at 1 μm and for a tungsten absorber it would be 0.3 μm. It is interesting to note that one can use the results of Figure 9-17 to take into account the roughness of an absorber by degrading its radiation length [DeStaebler 1994].

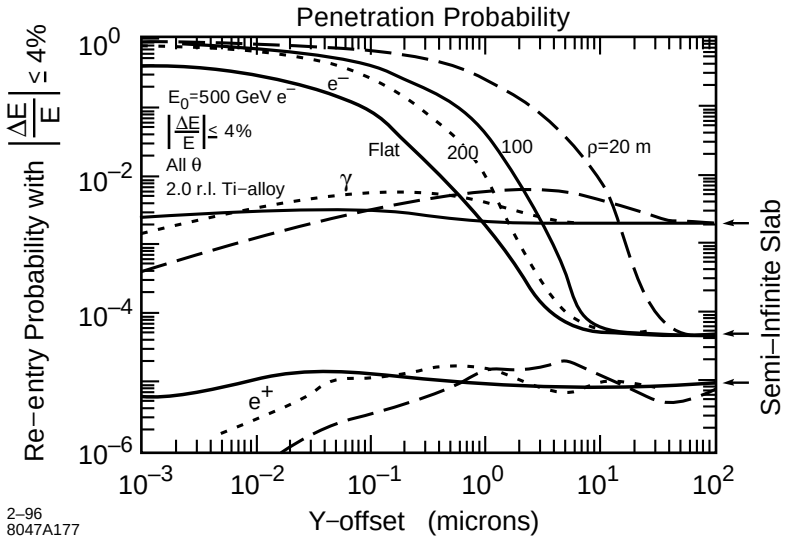


Figure 9-16. *The fraction of particles that re-enter the beam as a function of distance from the edge of a flat and curved (200-m radius of curvature) titanium scraper.*

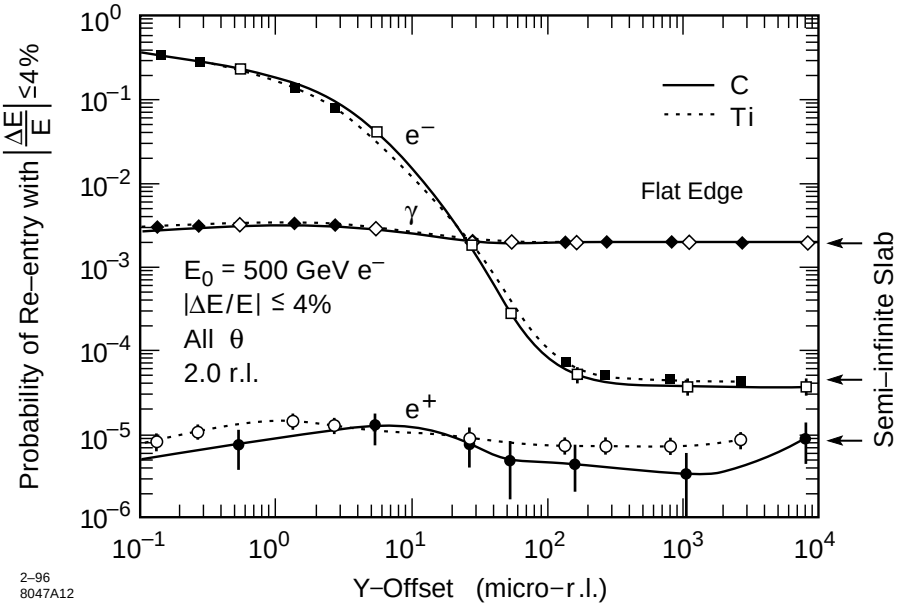


Figure 9-17. *The fraction of particles that re-enter the beam as a function of distance from the edge of a curved (200-m radius of curvature) titanium and carbon collimator. Abscissa is in micro-radiation lengths.*

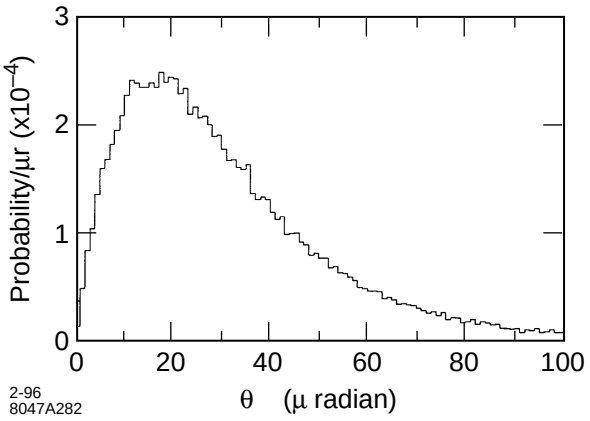


Figure 9-18. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the cone angle measured from the initial particle direction. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

Figure 9-18 shows the angular distributions for edge-scattered particles re-entering the beam with energy losses of less than 20%. The angle in this distribution is the cone angle measured from the particle direction. The particles incident on the absorber were uniformly distributed from 0 to 10 μm from the edge of a Cu absorber with a radius of curvature of 100 m. Note the remarkable similarity to the angular distributions from spoiler transmission shown in Figure 9-13. The total number of particles scattered back into the beam (with energy loss less than 20%) is 0.1 times the number of particles incident on the edge per micron. In other words the number re-entering the beam is equal to the number that fall on the edge at a distance of 0.1 microns or less. The theoretical distribution of Eq. 9.8 fits this graph rather well for an rms angle of $\theta_0 = 20\ \mu$r. The effective thickness of the material (using Eq. 9.8) would be $t_{eff} \approx 0.5$ r.l.

Figure 9-19 shows the angular distribution when the angle is chosen to be the angle from a plane of the collimator surface, for the same condition as those in Figure 9-18. This distribution is rather surprising in that the shape is still Gaussian. A quick calculation shows that the rms angle of this distribution is achieved in a distance of $20\ \mu\text{r} \times 100\ \text{m} = 2$ mm, which is about one fifth of a radiation length for Cu, so indeed the particles can be headed back toward the plane if they are close enough to the edge that they pass through two-fifths of a radiation length or less in the collimator. We note however that there is a mean of this distribution away from the surface of about 12 μr.

Figure 9-20 shows the angular distribution when the angle is chosen to be the angle from a plane containing the line of flight of the particles and perpendicular to the surface of the absorber. If the absorber where a vertical collimator, then this angle is the horizontal angle. This distribution should be symmetric because there is no physical way to distinguish left from right.

Figure 9-21 shows the energy distribution for the particles which re-enter the beam. The particles with energy loss less than 4% is similar to Figure 9-14. The distribution for small energy losses fits a power law as in Eq. 9.9 with $t_{eff} = 0.4$. For energy losses greater than 4% the distribution is quite flat. A number which will be important for us is the total number of particles that re-enter the beam. For a uniform density distribution of 1 particle per micron incident on the edge of this 100-m curved copper absorber there are a total of 0.1 particles re-entering the beam with an energy loss of less than 20%. Of this number 30% (or 0.03 particles) have an energy loss less than 4%, and 0.07 have energy loss between 4% and 20%.

Figure 9-22 shows the distribution of Figure 9-19 for four energy bins: the top 1-GeV bin (0.2%), the next 1-GeV bin, a 1-GeV bin with energy 2% below the maximum energy, and a 1-GeV bin with energy 10% below the maximum.

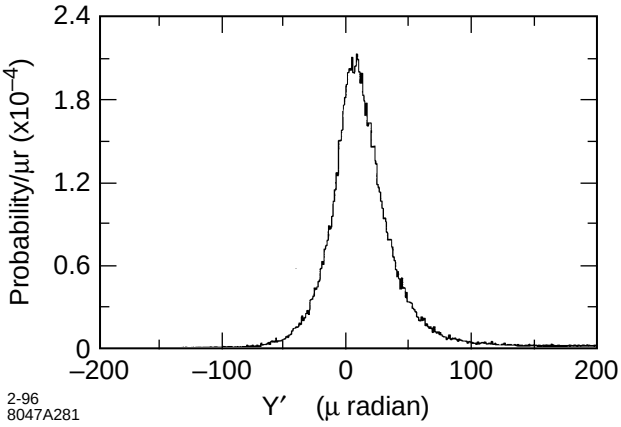


Figure 9-19. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber surface. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

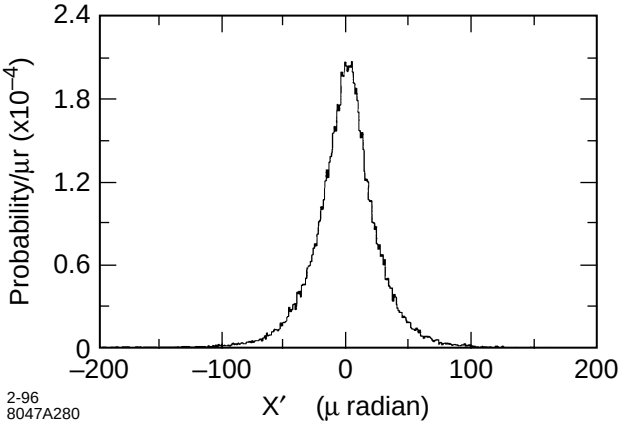


Figure 9-20. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from a plane which is perpendicular to the absorber surface containing the particle line-of-flight. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

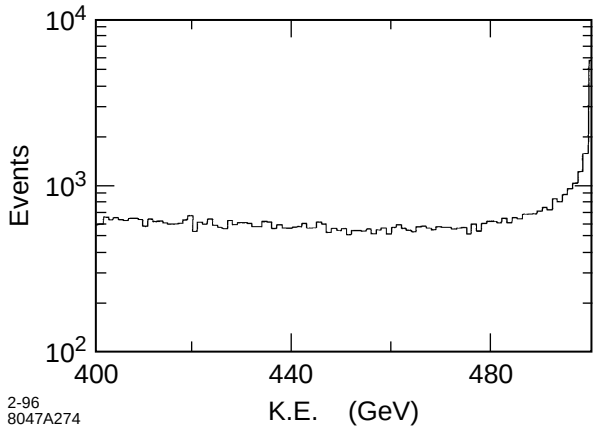


Figure 9-21. *The energy distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

There is a tendency for the larger energy loss particles to have a wider angular distribution. Figure 9-22b contains the same distributions as shown in Figure 9-22a but with a radius of curvature equal to 20 m rather than 100 m.

From the distributions of this section we can arrive at the following conclusions:

- It is primarily particles very close to the edge (≤ 0.3 μm) that are scattered back into the beam, so the distributions will not depend on the incoming particle distribution (as long as it is constant over a 1-μm distance). The total number re-entering the beam will depend only on the number of particles per μm incident on the absorber at its edge.
- The angular distributions are very similar to the angular distribution of particles transmitted through a 0.5 r.l. spoiler. The tails of the distribution will be spread due to low energy particles transmitted with larger angular kicks.
- The energy distribution, for particles with less than 4% energy loss, follows a power law distribution characterized by an effective thickness of 0.4 r.l.
- The total number of particles with energy loss less than 20% that reenter the beam is 0.1 times the number of incident particles per micron. The total number of particles with energy loss less than 4% is 0.03 times the number of incident particles per micron.

These distributions can be used to track re-scattered particles through the beam line, and the results can be used to estimate the number of particles that will be present in the particle tails as the beam travels along the beam line from the end of the linac to the IP.

Same-Section Absorbers. After the spoilers there are absorbers in the beam line to actually absorb the bulk of the energy of the tail particles. One percent of the average beam power at 1-TeV c.m. is 84 kW, so this function is substantial.

We have argued above that we must have $\sqrt{(R_{12}R_{34})} \geq 160$ m (or the equivalent when allowing for energy spread) so that the beam particles that have passed through the spoiler are spread over an area with $\sqrt{(\sigma_x \sigma_y)} \geq 2.2$ mm. This implies that at the edge of the absorber there will be a particle density of about $n_\mu \approx 10^{10}/\,\text{mm} = 10^7/\,\mu\text{m}$.

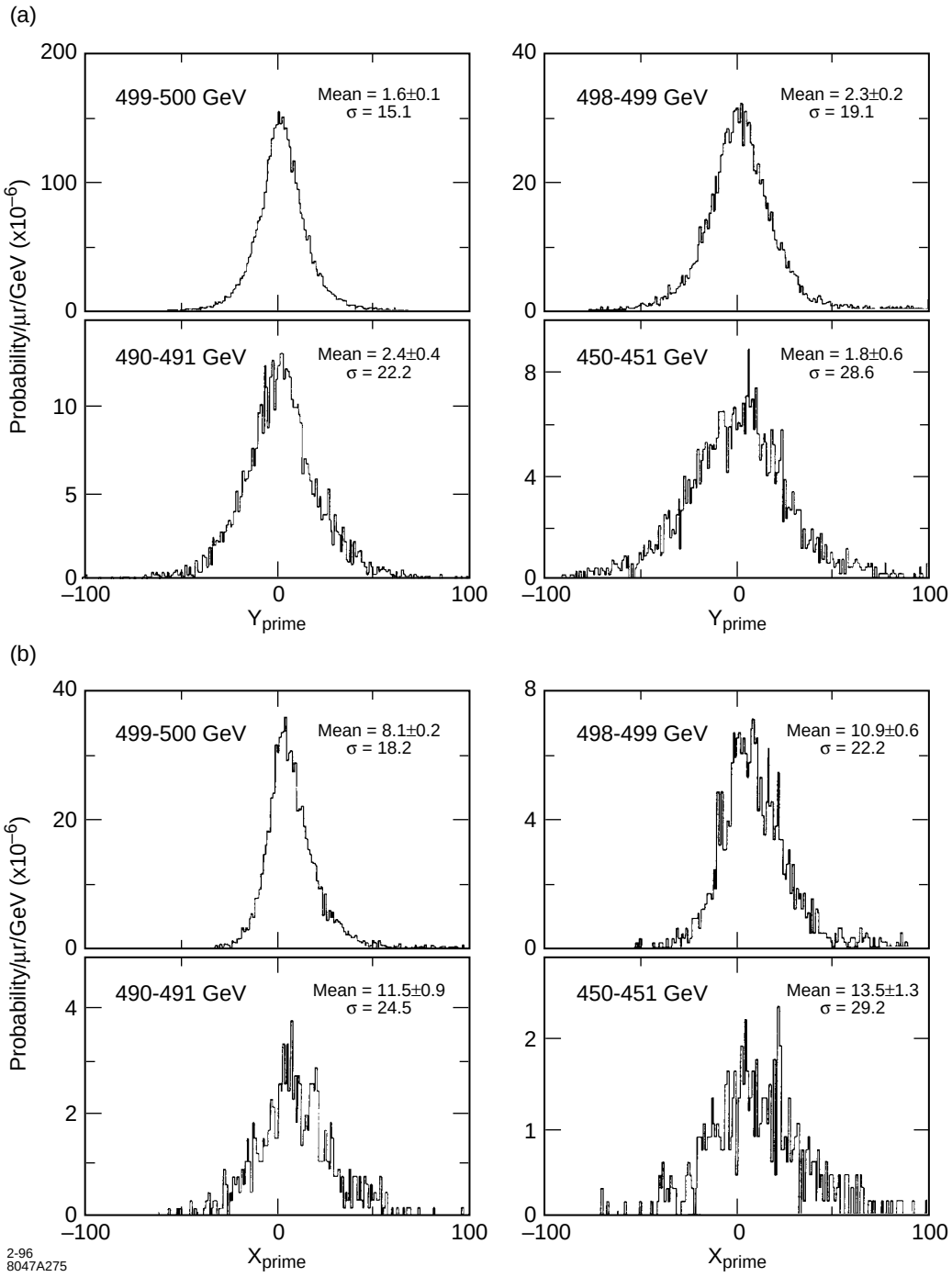


Figure 9-22. *a. The angular distribution of particles that re-enter the beam for four different energy losses for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber. The incident beam is uniformly distributed at the edge for 0 to 10 cm. b. The angular distribution of particles that re-enter the beam for four different energy losses for a Cu collimator with a 20-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

Since the first spoilers are located at a β maximum, they are adjacent to a quadrupole which is focusing in the plane being collimated. The next quadrupole is invariably defocusing in that plane, so at the absorber, which is located just prior to the second downstream quadrupole, the beam distribution will be elliptical with the major axis in the plane being collimated. In other words, the estimate of $10^7/\ \mu$m can be taken as an upper limit.

In the previous section it was determined that the number of particles re-scattered from a copper absorber with energy loss less than 4% is about 0.03 times the number of particles incident per micron at the edge. Using this fact and the above estimate of $n_\mu = 10^7/\ \mu$m we have the estimate for the number re-scattered (with less than 4% energy loss) as $N_{sc} \leq 0.3\ 10^6$.

Of the 10^{10} tail particles coming from the spoiler only 1/2 will impact the first absorber, since 1/2 of the particles have an angle away from the collimated plane.

We have established in Section 9.2.2 that the β functions at the spoiler must satisfy $\sqrt{(\beta_x \beta_y)} \geq 100^2$ m. From this we can conclude that the rms angle ($\theta_0 \approx 14\,\mu$r) of particles emerging from the spoiler is much larger than the beam divergence. Indeed we have $\hat{\sigma}_{x'} = \sqrt{(\varepsilon_x/\beta_x)} \leq 25$ nr and $\hat{\sigma}_{y'} = \sqrt{(\varepsilon_y/\beta_y)} \leq 2.5$ nr. [Note: There can be, and often is, a large non-zero $\alpha = -1/2\ d\beta/ds$. The relevant angular spread within the beam is $\sqrt{(\varepsilon/\beta)}$, not the rms of the angular distribution in phase space, $\sigma' = \sqrt{[(1+\alpha^2)(\varepsilon/\beta)]}$. Therefore we have introduced the notation $\hat{\sigma}' = \sqrt{(\varepsilon/\beta)}$]. So $\theta_0/\hat{\sigma}_{x'} \geq 600$ and $\theta_0/\hat{\sigma}_{y'} \geq 6,000$. However, in the transport system from the collimation system to the IP, the β functions are about 50 m, so that beam-stay-clears are also large, approximately 300 σ_x and 3,000 σ_y. Collimation must be inserted somewhere in this beam line, since 10^9 particles can not be collimated in the final focus. Though they are originally in the IP phase, because of beam-line nonlinearities, chromaticity and high order dispersion, they are sure to migrate to other phases because of their extremely large amplitudes. (See the tracking studies in Section 9.2.5. Actually only 3.5 10^8 particles reach the final focus, but of these 10^8 impact elements in the final-focus system, which is 100 times the acceptable number.)

Suppose collimation is introduced so that only 10^6 particles remain in the tails. In Section 9.2.2 we integrated the distribution to find the number of particles with angles less than some small angle θ_1 to be $\Delta N/N = 1/2(\theta_1/\theta_0)^2$. For this ratio to be 2 10^{-4}, we must have $\theta_1/\theta_0 \leq 2\ 10^{-2}$. Hence we have the estimates $\theta_1/\hat{\sigma}_{x'} \sim 10$ and $\theta_1/\hat{\sigma}_{y'} \sim 100$. We do not have to collimate to apertures that are this small, since these particles are at the IP phase, and such small amplitude particles could be transported through the IP to the dump. However we can conclude that after spoiling (and absorbing) the final-doublet phase, we must collimate the IP phase at least to an amplitude that will be transported safely through the IP. According to the dynamic aperture studies of the final-focus system (see Section 9.2.1 and Figure 9-5) this must be less than 45 σ_x and 200 σ_y.

Considerations of this section lead us to conclude that at a minimum we must either collimate the IP phase at least two times, once before the FD-phase collimation (so that the FD-phase collimation can clean up tail particles generated in the IP-phase collimation) and once after the FD-phase collimation, or collimate the FD phase twice and the IP phase once. The latter option seems attractive because the effect of gas scattering from the end of the collimation system to the FD phase of the final focus is smaller (R_{12} and R_{34} are smaller.) However the gas scattering with the IP phase last is acceptable, and it has the advantage that the last IP phase can have a larger aperture. This turns out to be important when we consider machine protection issues that arise from wake effects on badly mis-steered beams.

Though in this chapter we describe an option in which there is a second FD phase as well as a second IP phase collimation, our present opinion is that the second FD phase is unnecessary.

Next-Phase Absorbers. The R_{12} and R_{34} between phases is given by $R_{12} = \sqrt{(\beta_{x1}\beta_{x2})}$ and $R_{34} = \sqrt{(\beta_{y1}\beta_{y2})}$. Between the first IP-phase and FD-phase collimation (assuming they are collimated to the same depth), we have, from our estimates on the minimum size of the β functions, that both R_{12} and R_{34} are greater than 8 km. This implies that the size of the distribution coming from the first IP-phase spoiler will be about $R\theta_0 \approx 10$ cm when reaching the FD collimators, 100 times larger than the 1-mm size at the absorber immediately following the spoiler.

These estimates can be made more precise by writing the distribution from the spoiler as (see Eq. 9.8)

$$\frac{dN}{N} = \frac{1}{2\pi\theta_0^2} e^{-\frac{\theta_x^2}{2\theta_0^2}} e^{-\frac{\theta_y^2}{2\theta_0^2}} d\theta_x d\theta_y \quad , \tag{9.12}$$

then integrating over one of the planes to find the distribution in the other plane as

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}\,\theta_0} e^{-\frac{\theta_x^2}{2\theta_0^2}} d\theta_x \quad . \tag{9.13}$$

For small θ_x the particle density is given by

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}} \frac{d\theta_x}{\theta_0} \quad . \tag{9.14}$$

The distribution of Eq. 9.14 coming from the spoiler implies that the spatial distribution of particles hitting the downstream absorber edge is

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}} \frac{dx}{R_{12}\theta_0} = \frac{1}{\sqrt{2\pi}} \frac{n_{1,x} n_{2,x}}{g_1 g_2} \frac{\varepsilon_x}{\theta_0} dx \quad . \tag{9.15}$$

We have put the expression in this form because, if the later phases are collimated at a different beam aperture, still the gaps g will be mostly between 1 or 2 mm, whereas β-functions are varied to achieve the appropriate collimator size. If the beam aperture (number of sigma) of the second stage is larger, we see that the estimate for the number of particles at the edge increases linearly with the aperture.

Figure 9-23 shows the spoilers and absorbers of a collimation system having two IP stages and a single FD stage and indicates the re-scattering that is expected at each stage.

Tail Re-population From Gas Scattering

Gas scattering must also be considered as a potential source of particles that will hit downstream collimators. The cross section for Coulomb scattering is given by

$$d\sigma = \left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 d\phi \frac{dq}{q^3} \quad . \tag{9.16}$$

Depending on the nature of the downstream aperture, one can find an expression for the minimum transverse momentum $q_{\min}$ which an electron must receive in order to impact that aperture. If the aperture is round the expression is given by

$$\frac{q_{\min}(s,\phi)}{p_0} \left(R_{12}(s)^2 \cos^2\phi + R_{34}(s)^2 \sin^2\phi\right)^{\frac{1}{2}} = a \tag{9.17}$$

where p_0 is the beam momentum, a is the aperture radius, and ϕ is the azimuthal angle at which the particle is scattered. If the aperture is a pair of horizontal planes, with half gap g, the expression for $q_{\min}$ will be

$$\frac{q_{\min}(s,\phi)}{p_0} R_{12}(s) \cos\phi = g \quad . \tag{9.18}$$

The cross section for a scattered particle to hit the final doublet is

$$\frac{d\sigma}{d\phi} = \left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 \int_{q_{\min(s,\phi)}}^{\infty} \frac{dq}{q^3} = \frac{1}{2}\left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 \frac{1}{q_{\min}^2(s,\phi)}$$

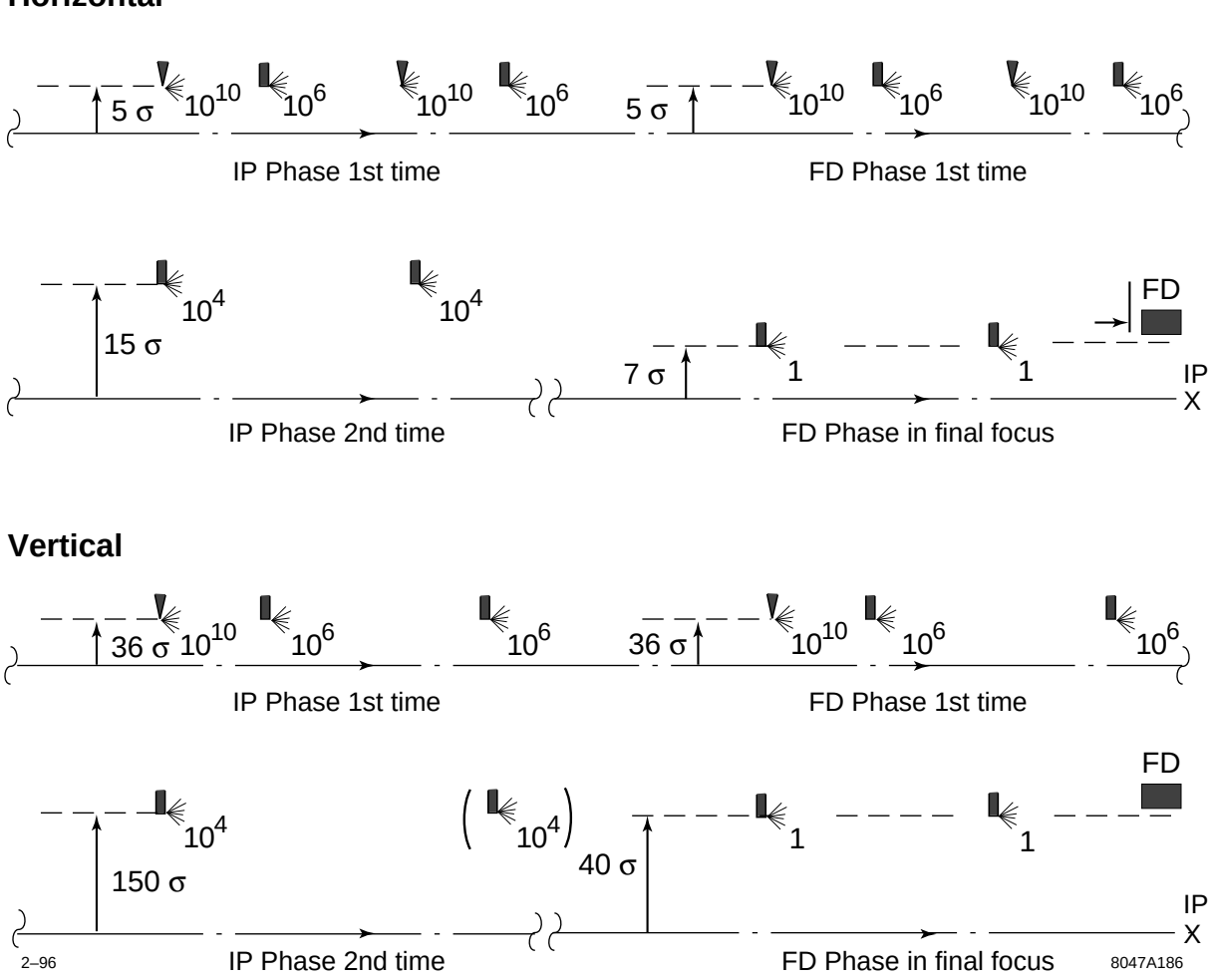


Figure 9-23. *A diagram showing the spoilers and absorbers of a collimation system that contains a single FD phase and two IP phases. The estimated particle numbers emerging from each collimator is shown.*

$$= 2\left(\frac{r_e Z}{\gamma a}\right)^2 \left(R_{12}(s)^2 \cos^2\phi + R_{34}(s)^2 \sin^2\phi\right) \ , \quad \text{circular aperture, or} \tag{9.19}$$

$$= 2\left(\frac{r_e Z}{\gamma g}\right)^2 R_{12}(s)^2 \cos^2\phi \quad \text{for a flat aperture.}$$

This may now be integrated over ϕ. The probability of a scatter in distance ds is given by $\rho\sigma ds$. Thus the number scattered along the beam line which will exceed the aperture limit is given by

$$\frac{\Delta N}{N} = 2\pi\rho_N \left(\frac{r_e Z}{\gamma a}\right)^2 \int ds \left(R_{12}(s)^2 + R_{34}(s)^2\right) \ , \quad \text{or}$$

$$= 2\pi\rho_N \left(\frac{r_e Z}{\gamma g}\right)^2 \int ds R_{12}(s)^2 \quad . \tag{9.20}$$

If we assume that the major composition of gas within the beam pipe is a diatomic molecule like N_2, the number of nuclei per unit volume is approximately $\rho_N \approx 6\ 10^{22} P_G m^{-3}$ where P_G is the pressure measured in Torr. If we assume a value of $Z = 7$ and a gas pressure of 10^{-8} Torr, these equations become (for 1-TeV-c.m. energies)

$$\begin{aligned} \frac{\Delta N}{N} &= 1.5\ 10^{-24} \frac{1}{a^2} \int ds \left(R_{12}(s)^2 + R_{34}(s)^2 \right) \quad , \qquad \text{or} \\ &= 1.5\ 10^{-24} \frac{1}{g^2} \int ds R_{12}(s)^2 \quad . \end{aligned} \tag{9.21}$$

Also of interest is the number density at the edge. This is given by

$$\begin{aligned} \frac{dN}{da} &= 3\ 10^{-12} \frac{1}{a^3} \int ds \left(R_{12}(s)^2 + R_{34}(s)^2 \right) \quad , \qquad \text{or} \\ \frac{dN}{dg} &= 3\ 10^{-12} \frac{1}{g^3} \int ds R_{12}(s)^2 \end{aligned} \tag{9.22}$$

where we have set $N = 10^{12}$.

First we can make an estimate of how many particles will hit the absorber in the first FD phase from gas scattered particles in the first IP phase. Since the two beam-line sections are separated by $\pi/2$ in phase we can set $R_{12}^2(s) = \beta_{x1}(s)\beta_{x2}$. We get the estimate

$$\frac{dN}{dg} = 3\ 10^{-12}\ \frac{\beta_{x,2}}{g^3} \int ds \beta_{x,1}(s) = 3\ 10^{-12}\ \frac{\beta_{x,2}}{g^3} L \left\langle \beta_{x,1} \right\rangle \tag{9.23}$$

where L is the length of the first phase collimation section. If we take $L = 400\,\text{m}$ (see the lattices in Section 9.2.5), $< \beta >= 1/2\ \beta_{\max}$, and $g = 2\,\text{mm}$, we get $dN/dg = 5\ 10^6$. The number per micron would be $n_\mu = 5$, and hence negligible. The total number hitting the collimator is estimated to be $N \approx 2\ 10^4$. These numbers are small compared to other backgrounds within the first stage of the collimation system.

The same approach can be used to find an estimate for number of particles that are gas-scattered in the last IP phase of the collimation system (assuming it is an IP phase) that are incident on the collimators of the final-focus system. The $\beta_{x2} = 4\ 10^4$ is larger by a factor of about 2, but g is larger by a factor of about 3. Thus estimates are $dN/dg \approx 3\ 10^6 (n_\mu = 3)$ and $N \approx 6\ 10^3$. These numbers are now not negligible, but they are safely within the number of particles that can be collimated there.

There are two other important estimates to make: the number of tail particles gas scattered in the entire beam line from the collimation system to the final-focus system, and the number of particles incident on the final doublet from particles that are gas scattered within the final telescope of the final-focus system.

To estimate the number from the beam line, we will assume the value of $\sin^2 \phi_{12}(s)$ can be taken to be 1/2. This should be true since the phase advances of these beam-line modules are not locked into the final-focus system phases. We can thus use the same formula derived for the IP-to-FD phase multiplying by a factor of 1/2. Now $L = 1.3\,\text{km}$, and $< \beta_{1,max} > \approx 40\,\text{m}$. Taking $\beta_{x,2} = 4\ 10^4$ and $g = 4\,\text{mm}$, we get the estimates $dN/dg \approx 5\ 10^4 (n_\mu = .05)$ and $N \approx 10^2$. These are much smaller than the particles from the last collimation section.

Within the final telescope it is important to use the correct R_{12} function and the round aperture formula. The numbers are still quite acceptable. See Chapter 11, Eq. 11.68.

9.2.4 Wakefield Considerations

Wakefields of Tapered Collimators

The wakefield kick from a spoiler or absorber consists of a geometrical part and a resistive-wall part. To reduce the geometrical wakefield, the spoiler and/or absorber may be tapered as shown in Figure 9-24.

The geometric wake from a tapered spoiler was first studied by K. Yokoya for a cylindrical beam pipe geometry [Yokoya 1988]. The resulting formula may be written as

$$\Delta y_G'^T = A\,[2 I_1 y_0] \quad \text{where} \quad A = \frac{N r_e}{\gamma \sigma_z} f_G(\tau)\ , \quad I_1 = \int \frac{b'^2}{b^2} ds\ , \quad \tau = \frac{z}{\sigma_z} \quad , \quad \text{and} \quad f_G(\tau) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} \quad . \tag{9.24}$$

Here $b(s)$ is the half-height of the beam pipe as a function of longitudinal position s. y_0 is the displacement of the source particle from the axis. There is no dependence on the position of the test particle for small displacements. The integral I_1 is especially easy to evaluate for a constant b', by taking one factor of b' outside the integral and writing $b' ds = db$ inside the integral. The integral evaluates to $(1/g - 1/b)$, where b is the maximum half-height and g is the collimator half-gap.

The geometric wake for a parallel-plate taper has been studied by G. Stupakov [Stupakov 1996]. The initial result, which is the formula used in many of the computations in this chapter, was

$$\Delta y_G'^T = A\,[1.85\ I_1 y_0 + 1.43\ I_1 y] \tag{9.25}$$

where there is now a dependence on y, the position of the test particle. This term is called the quadrupole wake which is present even for on-axis $y_0 = 0$ source particles. For $y = y_0$ this wake is 64% larger than the cylindrical geometry wake.

The current theoretical result, which was deduced by Stupakov as this text was going to press, may be written

$$\Delta y_G'^T = A\,[(2\pi\, w\, I_2 - 2\, I_1)\, y_0 + 2\, I_1\, y] \quad \text{where} \quad I_2 = \int \frac{b'^2}{b^3} ds \quad . \tag{9.26}$$

Now a dependence on w, the width of the collimator, is present. For a constant b', the integral I_2 evaluates to $1/2(1/g^2 - 1/b^2)$. For $y = y_0$ the I_1 terms cancel, and the dependence on gap width changes to $1/g^2$. The resulting wake is larger than the one used in this chapter by a factor of $(\pi/3.82)(w/g)$, which for $w = 5g$ is about 4.

Since the result of the last paragraph is new and requires confirmation, we have decided to use a modified cylindrical collimator wake in the following sections. For the near-wall and quadrupole-wake effects we will use the original parallel-plate results. This is unfinished and somewhat unsatisfactory, but is a compromise which allows us to illustrate the range of physical effects that are important.

Evaluating the integral I_1 and inserting $b' = (b - g)/L_T$ where L_T is the length of the taper, the taper will have the wake equation

$$\Delta y_G'^T(\tau) = \alpha_G\, \frac{N r_e}{\gamma \sigma_z}\, \frac{2(b-g)^2}{b g L_T} f_G(\tau) \Delta y \qquad \text{where}$$
$$\tau = \frac{z}{\sigma_z} \qquad \text{and} \qquad f_G(\tau) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} \quad . \tag{9.27}$$

α_G is a geometric factor that is unity for a cylindrical beam pipe (Eq. 9.24), equal $3.28/2 = 1.64$ for parallel plate geometry if Eq. 9.25 is valid, and under conditions where w was assumed to scale with g (large g case), would be $(\pi/2)\ w/g$ where Eq. 9.26 holds.

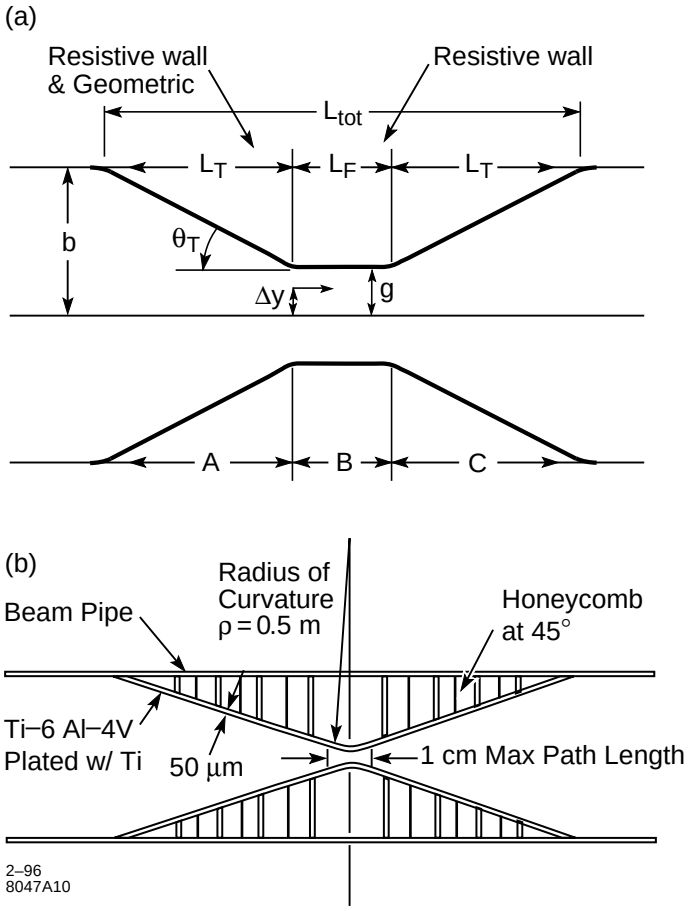


Figure 9-24. *a. Cross-sectional profile of a tapered absorber. b. Cross-sectional profile of a tapered spoiler.*

To get the resistive-wall kick from a taper we have integrated parallel-plate resistive-wall wake formula [Chao 1992].

$$\Delta y_R'^T(\tau) = \alpha_R \frac{N r_e}{\gamma \sigma_z} \frac{2}{r^3(s)} ds \sqrt{\lambda \sigma_z} f_R(\tau) \Delta y \qquad \text{where}$$

$$\tau = \frac{z}{\sigma_z} \qquad \text{and} \qquad f_R(\tau) = \frac{\sqrt{2}}{\pi} \int_0^\infty \frac{d\tau'}{\sqrt{\tau'}} e^{-\frac{(\tau'+\tau)^2}{2}} \quad . \tag{9.28}$$

The result is

$$\Delta y_R'^T(\tau) = \alpha_R \frac{N r_e}{\gamma \sigma_z} \frac{(b+g)}{b^2 g^2} L_T \sqrt{\lambda \sigma_z} f_R(\tau) \Delta y \quad . \tag{9.29}$$

The function f_R is shown in Figure 9-25. α_R is a geometric factor that is 1 for a cylindrical beam pipe and equals $\pi^2/8$ for parallel plate geometry. Note: This formula is also somewhat uncertain, and SLC measurements are indicating a larger wake [SLC].

$\lambda \equiv \rho/(120\pi)$ is referred to as the resistive depth (ρ is the resistivity in Ω-m). This formula can be taken as valid over the entire bunch if $\sqrt{(\sigma_z \lambda)} \ll g \ll (\sigma_z/\lambda)\sqrt{(\sigma_z \lambda)}$. These inequalities are true for all combinations of λ, σ_z, and g

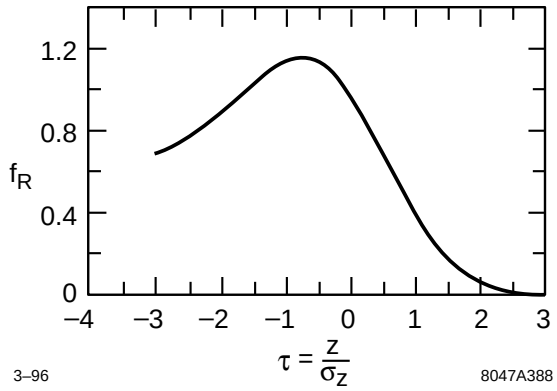


Figure 9-25. *The resistive-wall wake strength as a function of longitudinal position within the bunch.*

we will consider. For Cu, with $\lambda = 0.045\,\text{nm}$, this translates to $.08\,\mu\text{m} \ll g \ll 0.2\,\text{m}$, and for Ti, with $\lambda = 1.2\,\text{nm}$, this becomes $4\,\mu\text{m} \ll g \ll 0.4\,\text{m}$.

We will need averages over the bunch length of f_R, f_R^2, f_G, f_G^2, and $f_G f_R$, whereby the average of any function, h, over the bunch length we mean

$$\langle h \rangle \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\tau'\, e^{-\frac{\tau'^2}{2}}\, h(\tau') \tag{9.30}$$

These averages are $< f_G >= .282$, $< f_G^2 >= .092$, $\sigma_{f_G} = .111$, $< f_R >= .816$, $< f_R^2 >= .774$, $\sigma_{f_R} = .330$, and $< f_G f_R >= .25$ where by σ_h is meant: $\sigma_h \equiv \sqrt{(< h^2 > - < h >^2)}$. Note that both for f_G and f_R, $\sigma_f / < f >\approx 0.4$. This suggests that emittance growth effects will be smaller than centroid change effects. We also note for later reference that for both f_R and f_G the maximum values $f_{\max} = 1.4 < f >$.

To find the taper length L_T for which the total kick to the centroid is a minimum we must find the minimum of

$$\frac{A\,\langle f_G \rangle}{L_T} + B\,\langle f_R \rangle\, L_T \tag{9.31}$$

which occurs at $L_T = \sqrt{(A < f_G > /(B < f_R >))} = .58\sqrt{(A/B)}$ where both kicks are equal, and the total centroid kick is $< \Delta y'^T >= 2\sqrt{(< f_G >< f_R > AB)} \approx \sqrt{(AB)}$. Plugging in the expressions for A and B we have a minimum kick equal to (now for both incoming and outgoing taper)

$$\begin{aligned} \Delta y'^T &= 2\sqrt{\alpha_G \alpha_R}\, \frac{N r_e}{\gamma \sigma_z} \left[2\,\langle f_G \rangle\, \langle f_R \rangle\, \frac{(b^2 - g^2)\,(b - g)}{b^3} \right]^{\frac{1}{2}} \left(\frac{\lambda \sigma_z}{g^6} \right)^{\frac{1}{4}} \Delta y \\ &\approx 2\sqrt{2}\, \sqrt{\alpha_G \alpha_R}\, \frac{N r_e}{\gamma \sigma_z} \left(\frac{\lambda \sigma_z}{g^6} \right)^{\frac{1}{4}} \Delta y \qquad \text{for} \qquad g \ll b \end{aligned} \tag{9.32}$$

and

$$\begin{aligned} L_T^{opt} &= (b - g) \left(\frac{2b\,\langle f_G \rangle}{(b + g)\,\langle f_R \rangle} \right)^{\frac{1}{2}} \left(\frac{g^2}{\lambda \sigma_z} \right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \\ &\approx 0.8(b - g) \left(\frac{g^2}{\lambda \sigma_z} \right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \qquad \text{for} \qquad g \ll b \quad . \end{aligned} \tag{9.33}$$

For the usual case where $g \ll b$ the optimum taper angle is given by

$$\Theta_T^{opt} \approx 1.1 \left(\frac{\lambda \sigma_z}{g^2}\right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \tag{9.34}$$

Taking $g = 1\,\text{mm}$, $\sigma_z = 100\,\mu\text{m}$, $\alpha_G = \alpha_R$ and $\lambda = \{0.045, 0.15, 1.2\}\,\text{nm}$ for {Cu, W, Ti} respectively yields $\Theta_T^{opt} = \{9, 12, 21\}\,\text{mr}$. For $b = 5\,\text{mm}$ $L_T^{opt} = \{55, 42, 24\}\,\text{cm}$. The tungsten-rhenium alloy referred to here can be plated with copper to improve its surface conductivity. The titanium, because it will be used as a spoiler, cannot be plated, except with a material that has survival characteristics as good or better than itself.

The emittance of the beam is also enlarged by the kicks. Since these kicks are in one direction in phase space, they cause the normalized phase-space distribution to depart from a circle into something resembling an ellipse. We will introduce the notation that the IP- and FD-phase axis lengths of this ellipse are $r_{y,\text{IP}}$ and $r_{y,\text{FD}}$. For kicks that are small compared to the original radius of this distribution we would have the estimate for the final radius: $\tilde{r}_{y,\text{IP}}^2 = r_{y,\text{IP}}^2 + \left\langle \Delta r_{y,\text{IP}}{}^2 \right\rangle$, or equivalently

$$\tilde{r}_{y,\text{IP}} \approx r_{y,\text{IP}} \left(1 + \frac{1}{2} \frac{\left\langle \Delta r_{y,\text{IP}}{}^2 \right\rangle}{r_{y,\text{IP}}^2}\right) \quad . \tag{9.35}$$

The ratio in the parenthesis way be determined from

$$\frac{\left\langle \Delta r_{y,\text{IP}}{}^2 \right\rangle}{r_{y,\text{IP}}^2} = \frac{\left\langle \Delta y'^2 \right\rangle - \left\langle \Delta y' \right\rangle^2}{\hat{\sigma}_{y'}^2} \quad . \tag{9.36}$$

Hence to minimize the emittance growth we must minimize

$$\begin{aligned} &\left\langle \left(\frac{Af_G}{L_T} + Bf_R L_T\right)^2 \right\rangle - \left\langle \left(\frac{Af_G}{L_T} + Bf_R L_T\right) \right\rangle^2 \\ &= \frac{A^2 \sigma_{fG}^2}{L_T^2} + B^2 \sigma_{fR}^2 L_T^2 + 2AB \left(\left\langle f_G f_R \right\rangle - \left\langle f_G \right\rangle \left\langle f_R \right\rangle\right) \end{aligned} \tag{9.37}$$

The minimum emittance growth occurs at $L_T = \sqrt{(A\sigma_{fG}/(B\sigma_{fR}))} \approx 0.53\sqrt{(A/B)}$, (compare $0.58\sqrt{(A/B)}$ for the minimum centroid kick), and the emittance growth at this minimum is given by

$$\frac{\left\langle \Delta r_{y,\text{IP}}^2 \right\rangle}{r_{y,\text{IP}}^2} = 2\left(\sigma_{f_g}\sigma_{f_R} + \left\langle f_G f_R \right\rangle - \left\langle f_G \right\rangle \left\langle f_R \right\rangle\right) \frac{AB}{\sigma_{y'}^2} = 0.11 \frac{AB}{\sigma_{y'}^2} \approx 0.11 \frac{\left\langle \Delta y'_T \right\rangle^2}{\sigma_{y'}^2} \quad . \tag{9.38}$$

For a resistive-wall kick or geometric kick alone we would have

$$\frac{\left\langle \Delta r_{y,\text{IP}}{}^2 \right\rangle}{r_{y,\text{IP}}^2} = \frac{\sigma_{\Delta y'_R}^2}{\sigma_{y'}^2} = 0.16 \frac{\left\langle \Delta y'_T \right\rangle^2}{\sigma_{y'}^2} \quad . \tag{9.39}$$

Since $\frac{\Delta\sigma_{\text{IP}}}{\sigma_{\text{IP}}} = \frac{\Delta r_{y,\text{IP}}}{r_{y,\text{IP}}} = \frac{1}{2} \frac{\left\langle \Delta r_{y,\text{IP}}^2 \right\rangle}{r_{y,\text{IP}}^2}$, the luminosity loss from the taper emittance growth is

$$\frac{\Delta L}{L} = -\frac{\Delta\sigma_{\text{IP}}}{\sigma_{\text{IP}}} = 0.055 \frac{\left\langle \Delta y'_T \right\rangle^2}{\sigma_{y'}^2} \quad . \tag{9.40}$$

The luminosity loss from centroid kicks can be derived from the expression $(\Delta y_R - \Delta y_L)^2/(4\sigma^2)$. Taking the right and left kick to be statistically independent, this becomes $\Delta y_R^2/(2\sigma^2)$. Disruption, even at 1/2 design current, allows for twice the mis-steering for the same luminosity loss. Hence with disruption we have the estimate:

$$\frac{\Delta L}{L} = -\frac{1}{8} \frac{\Delta y_R^2}{\sigma_{\text{IP}}^2} = -\frac{1}{8} \frac{\left\langle \Delta y'_T \right\rangle^2}{\sigma_{y'}^2} \quad . \tag{9.41}$$

We see that for an optimized taper the emittance growth effect is 2.3 times smaller than the centroid kick effect, even when including disruption. In the horizontal plane, without disruption, it is 9.2 times smaller.

We are expecting an incoming vertical jitter from the linac of 0.2 σ. If we assume an $\frac{\langle \Delta y'_T \rangle}{\sigma_{y'}} \leq 0.7$ for the kicks from both the collimation and final-focus system, the luminosity loss from additional jitter would be about $(1/8)[(0.2)(0.7)]^2 = 0.25\%$, and the luminosity loss due to emittance growth would be 0.1%.

System Wake Equation

If the beam-centroid displacement Δy is due to beam jitter, then the wakefield $\Delta y'$ contributes to beam jitter 90° out of phase from and proportional to the source. The kick from the collimator `i' will give a kick to the beam, which measured in terms of $\hat{\sigma}_{y'} = \sqrt{\varepsilon_y/\beta_y}$ is

$$t'_i = \frac{\Delta y_i'^T + \Delta y_i'^F}{\hat{\sigma}_{y',i}} = 2\alpha_R \frac{N r_e}{\gamma \sigma_z} \left[\sqrt{2} \frac{(\lambda_i \sigma_z)^{1/4}}{g_i^{3/2}} \sqrt{\frac{\alpha_G}{\alpha_R}} + \langle f_R \rangle \frac{L_{F,i} (\lambda_i \sigma_z)^{1/2}}{g_i^3} \right] \frac{\Delta y_i}{\hat{\sigma}_{y',i}} \quad . \tag{9.42}$$

If we write the incoming jitter $\Delta y_i = t_i \sigma_{y,i}$ and use $\beta_{y,i} = \sigma_{y,i}/\hat{\sigma}_{y',i}$ then the above equation contains a dependence on the β function at the collimator, the collimator gap, and two parameters which depend on the material and dimensions of the collimator, λ and L_F. The β function can be written in terms of the gap and the number of σ collimated, through $\beta_y = g^2/(\varepsilon_y n_y^2)$, to obtain

$$\frac{t'_i}{t_i} = 2\alpha_R \frac{N r_e}{\varepsilon_y^N \sigma_z} \left[\sqrt{2}\, g_i^{1/2} (\lambda_i \sigma_z)^{1/4} \sqrt{\frac{\alpha_G}{\alpha_R}} + \langle f_R \rangle \frac{L_{F,i} (\lambda_i \sigma_z)^{1/2}}{g_i} \right] \frac{1}{n_{y,i}^2} \quad . \tag{9.43}$$

There is an optimal gap, $g = g_0$, which minimizes the bracket in the above equation,

$$g_{0i} = 0.54 (\lambda_i \sigma_z)^{1/6} L_{F,i}^{2/3} (\alpha_R/\alpha_G)^{1/3} \tag{9.44}$$

which depends only on σ_z and the collimator parameters. If the gap is taken to be optimal, then $t'_i = (\overline{n}_i/n_i)^2 t_i$ where $\overline{n}_i$ is equal to

$$\overline{n}_{y,i}^2 \equiv 2\alpha_R \frac{N r_e}{\varepsilon_y^N \sigma_z} [\ldots]_{opt} \approx 3.2 \frac{N r_e}{\varepsilon_y^N} \frac{(\lambda_i L_{F,i})^{1/3}}{\sigma_z^{2/3}} \alpha_G^{1/3} \alpha_R^{2/3} \tag{9.45}$$

When the collimation aperture is set to $n_i = \overline{n}_i$ and g is chosen optimally, then the ratio $t'/t = 1$. In terms of the parameters g_{0i} and $\overline{n}_i$, Eq. 9.43 can be written

$$\frac{t'_i}{t_i} = \frac{\overline{n}_{y,i}^2}{n_{y,i}^2} \left[\frac{2}{3} \sqrt{\frac{g_i}{g_{oi}}} + \frac{1}{3} \left(\frac{g_{oi}}{g_i} \right) \right] = \frac{\overline{n}_{y,i}^2}{n_{y,i}^2} \phi_i \tag{9.46}$$

For convenience we have introduced the function

$$\phi(x) \equiv \frac{1}{3} \left(2\sqrt{x} + \frac{1}{x} \right) \quad \text{and defined} \quad \phi_i \equiv \phi \left(\frac{g}{g_{0i}} \right) \quad . \tag{9.47}$$

It has the values $\phi(1) = 1$, $\phi(2) = 1.1$, $\phi(4) = 1.42$, $\phi(8) = 1.93$. We see that g can be a factor of 8 larger than g_0 before the kick from a collimator has doubled.

Equation 9.46 has the advantage that the parameters have an intuitive meaning and in addition the dependence on g will be quite weak if it is chosen near g_{0i}. The optimal gaps g_{0i} and the $\overline{n}_i$ parameters for the three types of collimators

Attribute	Symbol	Titanium Plated Titanium Alloy Spoiler (1/4 r.l.)	Copper Absorber (20 r.l.)	Copper-plated Tungsten-Rhenium Absorber (20 r.l.)
Resistive Depth	$\lambda = \rho/(120\pi)$ (nm)	1.2	0.045	0.045
Flat length	L_F (cm)	1.0	30.	7.0
Optimum gap	g_o (mm)	0.23	1.3	0.5
Horizontal $\overline{n}$	$\overline{n}_x$	1.42	1.45	1.14
Vertical $\overline{n}$	$\overline{n}_y$	11.2	11.5	9.0

Table 9-4. *Important parameters for spoilers and absorbers. The parameter depends on the emittance, bunch length and particle number of a bunch. This table uses the values $\alpha_R = \alpha_G = \pi 2/8$. π varies as $\alpha_G^{1/6}$.*

we will use are shown in Table 9-4. For the titanium spoiler, we take the length of the curved section at its apex ($L_F = 1$ cm) to be the appropriate length of the flat section.

The half-gaps, g, cannot be too large, or else the apertures of nearby beam-line elements become uncomfortably large, and they can not be too small, or else the collimator becomes impossible to control, align and adjust. It is remarkable that the optimum g, depending only on properties of materials and the bunch length, are actually reasonable values for collimator gaps.

Now we can write the equation for a system of collimators. The kicks from each collimator at any particular phase will add monotonically. If we assume that throughout the system the jitter in the phase being collimated does not increase appreciably, then setting $t' = \Sigma t'_i$ and all $t_i = t$ we have the result

$$\frac{t^i}{t} = \sum_i \left(\frac{t'_i}{t_i}\right) = \sum_i \left(\frac{\overline{n}_i}{n_{y,i}}\right)^2 \phi_i \quad , \tag{9.48}$$

where $n_{y,i}$ is the setting of the ith collimator. Because of machine protection and off-axis wake requirements to be discussed below, the collimator apertures $n_{y,i}$ will increase through the system. Since the resultant jitter is perpendicular in phase to the source jitter, statistically jitter from the collimation system will add in quadrature to the incoming jitter. A ratio of $t'/t = 0.7$, which would statistically produce a 25% increase in average beam jitter, is taken as an upper limit on acceptable jitter amplification.

Equation 9.48 must be solved iteratively for the required $n_{y,i}$ because the β functions may not be chosen freely to optimize the gap. The first iteration consists of making a list of all collimators that are planned for the system, taking $g = 1$ mm for all gaps, and assuming all $n_{y,i}$ are equal. Equation 9.48 can then be solved for a first guess n_y. With this value of n_y and an estimate of the β functions required at each collimator, a table of estimated gap sizes can be calculated. Any anticipated growth in n_y through the system can also be included. Now Eq. 9.48 can be used to obtain the values of t'/t for this n_y estimate. An iteration in the $n_{y,i}$ may be required to exactly meet the system amplification (t'/t) budget. This procedure is carried out in Section 9.2.5, and the results are listed in Table 9-5.

Quadrupole and Large Amplitude Wakes of Parallel Plate Collimators

Large Amplitude Resistive-Wall Wakes. The potential for the resistive-wall wake of parallel-plate collimators is given by [Piwinski]

$$V(x,y;x_0,y_0) = -\kappa f_R(\tau)\left[\frac{-x_-\sinh x_- + y_+\sin y_+}{\cosh x_- + \cos y_+} + \frac{x_-\sinh x_- + y_-\sin y_-}{\cosh x_- + \cos y_-}\right]$$
$$\text{where} \quad y_+ = \frac{\pi}{2g}(y+y_0), \qquad y_- = \frac{\pi}{2g}(y-y_0), \qquad x_- = \frac{\pi}{2g}(x-x_0), \tag{9.49}$$
$$\text{and} \quad \kappa = \frac{1}{2}\,\frac{Nr_e}{\gamma\sigma_z}\,\frac{L}{g}\,\sqrt{\lambda\sigma_x} \quad .$$

Here x and y are the horizontal and vertical position of the particle experiencing the wake, and x_0 and y_0 are the coordinates of the source of the wake. The collimator planes are set to collimate in y, and are separated by $2g$.

There are several limits of interest. One important limit ignores the difference between the horizontal position of the source and the test particles. If we are only interested in the vertical kicks we can set $x_- = 0$. Then

$$V(x_0,y;x_0,y_0) = -\kappa f_R(\tau)\left[\frac{y_+\sin y_+}{1+\cos y_+} + \frac{y_-\sin y_-}{1-\cos y_-}\right] \quad . \tag{9.50}$$

We can take the derivative of this equation to get the kick anywhere in the beam pipe

$$\Delta y' = \frac{\pi}{2}\,\frac{\kappa f_R}{g}\left[\frac{y_+ + \sin y_+}{1+\cos y_+} - \frac{y_- - \sin y_-}{1-\cos y_-}\right] \tag{9.51}$$

The kick received by particles in the core of the beam can be found by setting $y = y_0$ in Eq. 9.51.

$$\Delta y' = \frac{\pi}{2}\,\frac{\kappa f_R}{g}\left[\frac{\sin\hat{y}+\hat{y}}{1+\cos\hat{y}}\right] \qquad \text{where} \qquad \hat{y} = \pi\frac{y_0}{g} \tag{9.52}$$

The wake kick for a mis-steered beam corresponding to Eq. 9.52 is shown in Figure 9-26b.

For small y and y_0 we can expand V in a power series

$$V(x_0,y;x_0,y_0) \approx -\frac{\pi^2}{3}\,\frac{\kappa f_R}{g^2}\left[y_0 y + \frac{1}{4}y^2 + \ldots\right] \quad , \tag{9.53}$$

which produces the kick

$$\Delta y' \approx \frac{\pi^2}{3}\,\frac{\kappa f_R}{g^2}\left[y_0 + \frac{1}{2}y + \ldots\right] \quad . \tag{9.54}$$

For particles ($y = y_0$ in the beam)

$$\Delta y' \approx \frac{\pi^2}{2}\,\frac{\kappa f_R}{g}\,\frac{y}{g} \quad . \tag{9.55}$$

The coefficient of y/g in this equation would be the extrapolation to the wall of the small amplitude kick. We define this kick at the wall to be

$$\Delta y'_{ref} \equiv \frac{\pi^2}{2}\,\frac{\kappa f_R}{g} \tag{9.56}$$

For the flat part of a 20 r.l. Cu absorber with $g = 1.8\,$mm, $N = 10^{10}$, $\sigma_z = 125\,\mu$m, and $\gamma = 10^6$, $< \Delta y'_{ref} >= 11$ nr. For the flat part of a 20 r.l. copper -plated tungsten-rhenium absorber, $< \Delta y'_{ref} >= 2.6$ nr.

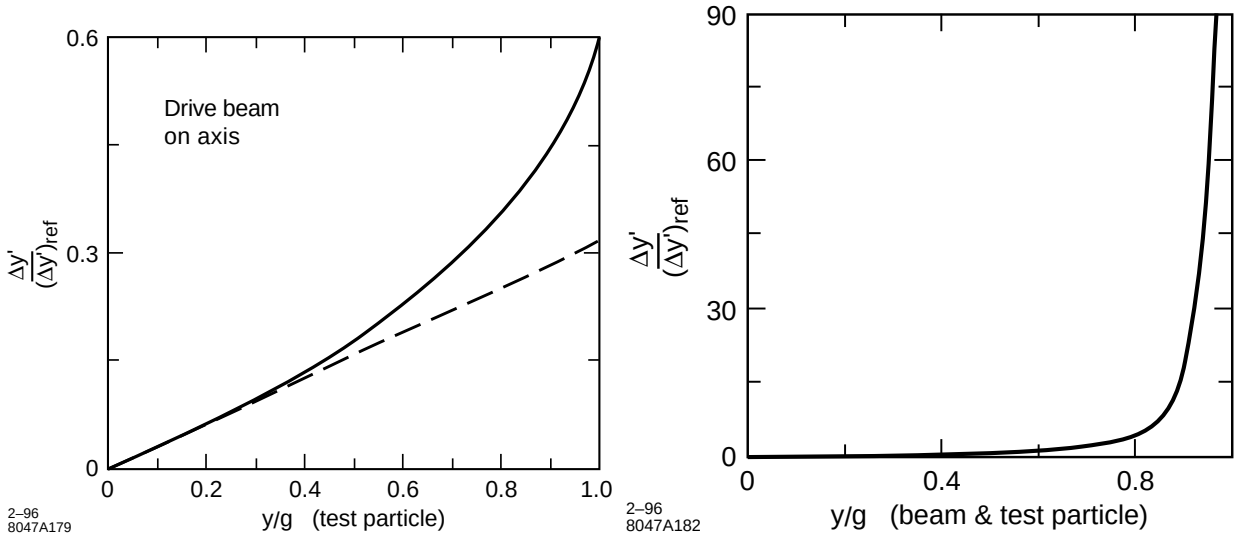


Figure 9-26. *(a) The resistive-wall wake from parallel plate collimators for an on-axis beam. (b) The resistive-wall wake from parallel plate collimators for a mis-steered beam. The wake-kick diverges as the beam approaches the collimator edge.*

Another limit of interest are wakes for on-axis beams. Keeping $x_- = 0$ and setting $y_0 = 0$, we have

$$V(x_0\,, y; x_0\,, 0) = \kappa f_R \left[\frac{2\tilde{y}}{\sin\tilde{y}}\right] \quad \text{where} \quad \tilde{y} = \frac{\pi}{2}\,\frac{y}{g} \quad . \tag{9.57}$$

Taking the derivative of this equation we get the kick

$$\Delta y'_{y_0=0} = \pi\frac{\kappa f_R}{g}\left[\frac{\sin\tilde{y} - \tilde{y}\cos\tilde{y}}{\sin^2\tilde{y}}\right] \quad . \tag{9.58}$$

As $y \to g$, $\tilde{y} \to \pi/2$, hence $\Delta y'_{y_0=0} \to \pi\kappa f_R/g$. The quadrupole term in the multipole expansion of Eq. 9.54 extrapolated to the wall would have given a kick $\Delta y' \to 1/3\Delta y'_{ref}$, which is smaller than the value of $\Delta y'_{y_0=0}$ at the wall by a factor of $\pi/6 \approx 1.9$. Figure 9-26(a) is a graph of the wake for an on-axis beam (Eq. 9.58) from $y = 0$ to $y = g$.

A third limit of interest is when both y and y_0 are close to g. Here it is useful to introduce the variable $\Delta_+ = \pi - y_+$ into Eq. 9.50 to obtain

$$V(x_0\,, y; x_0\,, y_0) = -\kappa f_R\left[\frac{\pi\sin\Delta_+}{1-\cos\Delta_+} - \frac{\Delta_+\sin\Delta_+}{1-\cos\Delta_+} + \frac{y_-\sin y_-}{1-\cos 6_-}\right] \xrightarrow{\Delta_+ \to 0} -2\pi\frac{\kappa f_R}{\Delta_+} \quad . \tag{9.59}$$

As $y_0 \to g$ we have the result that

$$\Delta y' \approx \kappa g f_R \frac{1}{(g-y_0)^2} \quad . \tag{9.60}$$

When the entire beam passes near the edge of a collimator the wake kick can be very large. However, as a beam approaches the wall, when $g - y_0 \le \sigma_x$, the singularity of the potential found in Eq. 9.53 becomes modified by the fact that $x_- \neq 0$. We must return to Eq. 9.49 and expand in x_- , Δ_+, and y_-. The result is

$$V(x\,, y; x_0\,, y_0) \quad = \quad -\kappa f_R\left[\frac{\pi\sin\Delta_+}{\frac{x_-^2}{2} + (1-\cos\Delta_+)} - \frac{x_-^2 + \Delta_+\sin\Delta_+}{\frac{x_-^2}{2} + (1-\cos\Delta_+)} + \frac{x_-^2 + y_-\sin y_-}{\frac{x_-^2}{2} + (1-\cos y_-)}\right]$$

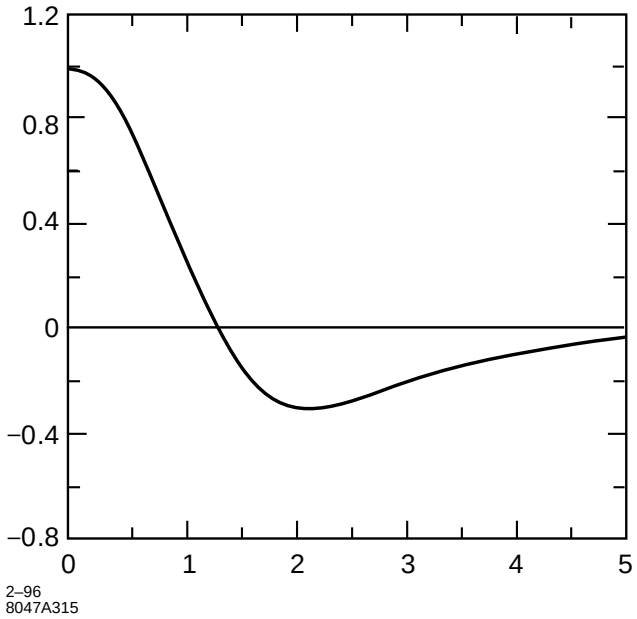


Figure 9-27. *The coefficient of the first power of λ in Eq. 9.65.*

$$\xrightarrow[\Delta_+\to 0]{} -2\pi\,\kappa f_R \frac{\Delta_+}{(x_-^2+\Delta_+^2)} \tag{9.61}$$

It is an interesting feature of this potential that for $\Delta_+ > x_-$ the force is toward the wall, while for $\Delta_+ < x_-$, the force is away from the wall. For $x_- = 0$ the force is always toward the wall (positive), and indeed the potential diverges as $1/\Delta_+$ as in Eq. 9.59.

Assuming the beam has a Gaussian shape, we can integrate Eq. 9.61 over the horizontal distribution. We have the result [Stupakov 1995].

$$\lim_{\Delta_+\to 0}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{dx_0}{\sigma_{x_0}}e^{-\frac{x_0^2}{2\sigma_{x_0}^2}}V(x,y;x_0,y_0) = -\frac{\sqrt{2\pi}\kappa f_R}{\sigma_{x_-}}\left[\pi e^{-\frac{\xi^2}{2}} - \frac{\Delta_+}{\sigma_{x_-}}\sqrt{2\pi}\,h_1(\xi)\right]$$
$$\text{where}\quad \sigma_{x_-} = \frac{\pi}{2g}\sigma_{x_0}\,,\quad \text{and}\quad \xi = \frac{x}{\sigma_{x_0}}\quad . \tag{9.62}$$

A plot of $h_1(\xi)$ is shown in Figure 9-27. From Eq. 9.62 we may obtain the horizontal and vertical kicks:

$$\begin{aligned}\Delta y'(x,y',y_0) &\approx \frac{4\kappa g f_R}{\sigma_{x_0}^2}\,h_1(\xi)\\ \Delta x'(x,y',y_0) &\approx -\sqrt{\frac{\pi}{2}}\,\frac{4\kappa g f_R}{\sigma_{x_0}^2}\,\xi e^{-\frac{\xi^2}{2}}\quad .\end{aligned} \tag{9.63}$$

Note that these kicks do not depend on the distance to the wall of either the source or the test particle, as would be expected for a plane geometry. Comparing Eq. 9.63 with 9.60 we see, for $\xi = 0$, the kick from Eq. 9.60 limits out at $g - y_0 = \sigma_{x0}/2$.

Geometric Wake Formulae. For the quadrupole term from the geometric wake we will use Eq. 9.24.

When the beam is close enough to the wall that it is appropriate to neglect the contribution of the opposite wall, the expression for the wake is [Stupakov 1996]

$$\Delta y'^T_G(\tau) = 4A \int \frac{a'^2}{a} ds = 4A\theta \ln\left[\frac{0.4g}{a_g}\right] \tag{9.64}$$

where a is the distance from the beam to the wall as a function of s, and a_g is the nearest approach of the beam to the taper.

If the beam is very close to the wall, the dimension of the beam parallel to the wall be comes important. Then the appropriate formula is

$$\Delta y'^T_G = 4A \int \frac{a a'^2}{x_2 + a^2} ds = 2A\theta \ln\left[\frac{x^2 + (0.4g)^2}{x^2 + a_g^2}\right] \tag{9.65}$$

where $\theta = b - g/L_T$.

This must now be integrated over the horizontal distribution. This can be performed in the limit of $a_g < \sigma_x$ to give (independent of the exact value of a_g)

$$\Delta y'^T_G = 2A\theta \ln\left[1.9\left(\frac{0.4g}{\sigma_x}\right)^2\right] \tag{9.66}$$

A table of numerical results for the collimators of the beam delivery lattice is given in Section 9.2.5.

If the Equation 9.26 is valid, the dipole term will be longer than this term, and is the current near-wall limit.

Consequences of Quadrupole and Near-Wall Wakes

Quadrupole-Wake Focusing. We see in Eq. 9.54 that for the resistive-wall wake the quadrupole term is 1/3 of the total wake kick. If the quadrupole wake were written as $\Delta y' = k_R \Delta y$, then $\Delta y'/\hat{\sigma}_{y'} = (k_R\beta)\Delta y/\sigma_y$. Hence we have

$$\text{“}k_{R,i}\beta_i\text{”} = 0.4\,\frac{1}{3}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}} + \frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}} + \frac{g_{0i}}{g_i}}\right] \quad . \tag{9.67}$$

The factor of 1/3 comes from the fact that, for the resistive-wall wake, 1/3rd of the wake comes from the quadrupole wake. t'_i/t_i is the number of σ' from a 1-σ displacement at the ith collimator. The bracket gives the fraction of the wake that is a resistive-wall wake. One-half of the taper wake is resistive-wall. At $g_i = g_{0i}$ the second bracket has the value 2/3 which is just the fraction of the total wake that is resistive-wall at the optimum gap. The value of the bracket lies between 1/2 and 1.

The initial factor of 0.4 comes from the fact that $\sigma_{fR}/ < f > \approx 0.4$. The average quadrupole effect of the wakes will be tuned out when the waist is tuned. It is the spread in focusing that occurs longitudinally along the bunch that is untunable, and will result in a luminosity loss.

As we have seen above in Eq. 9.65, the geometric parallel plate wakefield also has a quadrupole term. In other words, even for on-axis beams there will be a wakefield kick for off-axis particles. For the geometric wake the fraction of the total wake that is a quadrupole wake is $1.43/3.28 = .44$. This is somewhat larger than for the resistive-wall wake where this fraction is equal to 0.33. Multiplying this result together with the fraction of the wake from a collimator

that is geometric we have

$$\text{“}k_G\beta_i\text{”} = 0.4\left(\frac{t_i'}{t_i}\right)\left[\frac{.16\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_0 iu}{g_i}}\right] \quad . \tag{9.68}$$

The bracket has a value less than 1/2. Adding the resistive-wall and geometric together we have

$$\text{“}k_{R+G,i}\beta_i\text{”} = 0.4\left(\frac{t_i'}{t_i}\right)\left[\frac{0.77\sqrt{\frac{g_i}{g_{0i}}}+0.33\frac{g_0 i}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \tag{9.69}$$

The value of the bracket lies between 0.39 and 0.33.

We can sum Eq. 9.69 over all collimators, taking the worst case of 0.39 for the bracket to get

$$\text{“}k_{R+G}\beta''_{\rm Total} \leq (0.4)\left(\frac{t'}{t}\right)[0.39] \leq 0.11 \tag{9.70}$$

The luminosity is decreased by 2% when $k_Q\beta = 0.2$, hence the luminosity loss expected from the quadrupole focusing of collimators is less than 0.5%.

Near-Wall Wakes for On-Axis Beams. For parallel-plate collimators there is a wakefield even when the beam is on-axis. This raises the concern that particles passing near the collimator could be deflected into the tails and cause a problem downstream [Yokoya 1995].

We have shown above that for the resistive-wall wake the value of the wake at the wall for an on-axis beam is 1.9 times the linear extrapolation of the quadrupole wake of the on-axis beam. It is possible, by integrating the resistive-wall wake over the taper, to calculate the wake for a particle passing at the taper minimum. One finds the wake at the taper minimum is 1.4 times the linear extrapolation of the quadrupole wake. Putting these results together we get the following expression for the kick at the wall from the resistive-wall part of the wake:

$$\frac{\Delta y'_{RWX}}{\sigma_{y'}} = \frac{1}{3}n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{1.4\sqrt{\frac{g_i}{g_{0i}}}+1.9\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.71}$$

For the geometric wake we have not formally found an answer, but will assume that it is a factor of 2 times the linear extrapolation of the quadrupole wake. Hence

$$\frac{\Delta y'_{GWX}}{\sigma_{y'}} \cong (0.44)n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{2\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.72}$$

Adding Eq. 9.71 and 9.72 we obtain

$$\frac{\Delta y'_{R+G,WX}}{\sigma_{y'}} = n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{1.34\sqrt{\frac{g_i}{g_{0i}}}+0.63\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.73}$$

The bracket has a value between 0.67 and 0.63. When we sum over i, it should be over the upstream collimators only. Let us assume that 3/5 of the collimation budget for t'/t is in the first stage. Then Eq. 9.73 gives

$$\left.\frac{\Delta y'_{R+G,WX}}{\sigma_{y'}}\right|_{stage1} \leq \frac{3}{5}n_{y1}\left(\frac{t'}{t}\right)[0.67] \leq 0.28\, n_{y1} \tag{9.74}$$

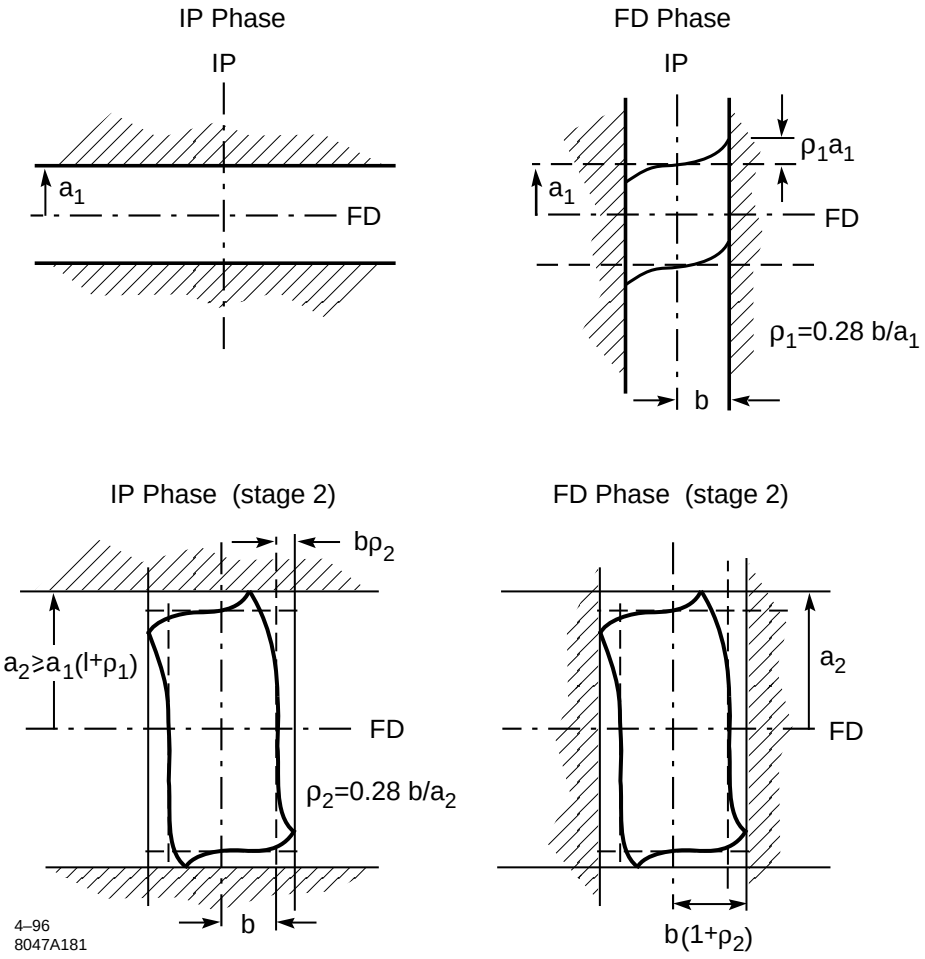


Figure 9-28. *When phase #2 is collimated, the on-axis beam wake which is present for parallel plate collimators, moves particles into a phase space region that was previously collimated in phase #1.*

The consequences of near-wall wakes for on-axis beams can be clarified by considering the particle motion in a sequence of normalized phase-space diagrams. See Figure 9-28. Let us suppose that the IP phase is first collimated to some aperture "$a_1\sigma$" and then the FD phase is collimated to an aperture "$b\sigma$". Because of the wake at the wall, the IP boundary will be distorted, as shown in Figure 9-26a, and particles will extend into the previously collimated region by an amount $0.28b\sigma$ (see Eq. 9.74). When the IP phase is collimated again, at a larger aperture "$a_2\sigma$" the distortion will be $0.28a_2(b/a_2)^2\sigma = (0.28b/a_2)b\sigma$. The factor $(b/a_2)^2$ arises from the assumption that b was the appropriate aperture for 3/5 of the 0.7 jitter amplification budget, and the effect of $a_2 > b$ can be determined by using the fact that the kick scales as the aperture squared. In Figure 9-28 we have set $\rho_2 = 0.28\ b/a_2$.

Because of near-wall wake for mis-steered beam, to be discussed in the next section, $a_2/b \approx 3.5$, hence $\rho_2 = 0.28b/a_2 = 0.08$. The stage-2 FD aperture must be enlarged by only 8%: from 36 to 39 in the vertical plane, from 5 to 5.4 in the horizontal plane.

Near-Wall Wakes for Mis-steered Beams. We would like to estimate the near-wall kick. The limit from the resistive-wall part is $(\kappa f_R/g)(2g/\sigma_x)^2 < h_1 > \approx 1/\pi^2 \Delta y'_{ref}(2g/\sigma_x)^2$. Hence we have

$$\frac{\Delta y'_{RWW}}{\sigma_{y'}} = \frac{4}{\pi^2}\left(\frac{g}{\sigma_x}\right)^2 n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.75}$$

The geometric-wake limit can be found from Eq. 9.66, putting $A\theta$ as the extrapolation of the small-amplitude wake to the wall:

$$\frac{\Delta y'_{GWW}}{\sigma_{y'}} = 2\,\ln\left[1.9\left(\frac{0.4g}{\sigma_x}\right)^2\right] n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.76}$$

We have not included any term for the resistive-wall part of the taper. Let us assume that only that length of taper contributes for which the surface is still within $\sigma_x/2$ of the taper minimum, since the wake falls off very sharply beyond that distance. When we integrate the resistive-wall-wake formula over this distance we find that the contribution to the small-amplitude wake for this region as compared to the whole taper is the fraction σ_x/g. Thus we have the result

$$\frac{\Delta y'_{RWW}}{\sigma_{y'}} = \frac{4}{\pi^2}\left(\frac{g}{\sigma_x}\right)^2 n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_01}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.77}$$

Combining these results we get (setting $g/\sigma_x = 10$)

$$\frac{\Delta y'_{R+GWW}}{\sigma_{y'}} = n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{7.4\sqrt{\frac{g_i}{g_{0i}}}+40\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_01}}+\frac{g_0i}{g_i}}\right] \quad . \tag{9.78}$$

The bracket lies between 1.9 and 40. The resistive-wall contribution is much larger than the taper. If we assume g/g_{0i} is about 2, we get a value for the bracket of about 10. This gives an estimate for the first stage of

$$\left.\frac{\Delta y'_{R+G,WW}}{\sigma_{y'}}\right|_{stage1} \leq \frac{3}{5}\,n_{y1}\left(\frac{t'}{t}\right)10 \leq 4.2\,n_{y1} \tag{9.79}$$

For $n = 36$ this gives the estimate for the near-wall kick of $150\sigma'$. The detailed calculation with the exact apertures of the lattice gives $118\sigma'$. A plot of this kick with distance from the wall is shown in Figure 9-29(a). Only the resistive-wall terms, adding up to $90\sigma'$ are included.

While these kicks are large, they are still in the beam pipe and smaller than the dynamic aperture of the final-focus system. Also because of the τ dependence, the spread of these kick is at least $0.4(120)\sigma_{y'} = 48\sigma_{y'}$. Additionally, there is a dependence on x, making the spread about equal to the magnitude of the kick. This can be deduced from the following relationships:

$$\begin{aligned}
<\Delta y'> &= \frac{4\kappa g}{\sigma_{x_0}^2}\langle f_R\rangle\langle h_1(\xi)\rangle = 0.41\frac{4\kappa g}{\sigma_{x_0}^2} \\
\sigma_{\Delta y'} &= \frac{4\kappa g}{\sigma_{x_0}^2}\sqrt{\langle f_R^2\rangle\langle h_1(\xi)^2\rangle - \langle f_R\rangle^2\langle h_1(\xi)\rangle^2} = 0.44\,\frac{4\kappa g f_R}{\sigma_{x_0}^2} \\
\langle\Delta x'\rangle &= 0 \\
\sigma_{\Delta x'} &= \sqrt{\frac{\pi}{2}}\,\frac{4\kappa g}{\sigma_{x_0}^2}\sqrt{\langle f_R^2\rangle\langle h_2(\xi)^2\rangle - \langle f_R\rangle^2\langle h_2(\xi)\rangle^2} = 0.49\,\frac{4\kappa g f_R}{\sigma_{x_0}^2} \\
\text{where} \quad & h_2(\xi) \cong \xi e^{-\frac{\xi^2}{2}}, \langle h_2(\xi)^2\rangle = 3^{-3/2}, \langle h_1(\xi)\rangle = 0.5, \text{and}\,\langle h_1(\xi)^2\rangle = 0.44 \quad .
\end{aligned} \tag{9.80}$$

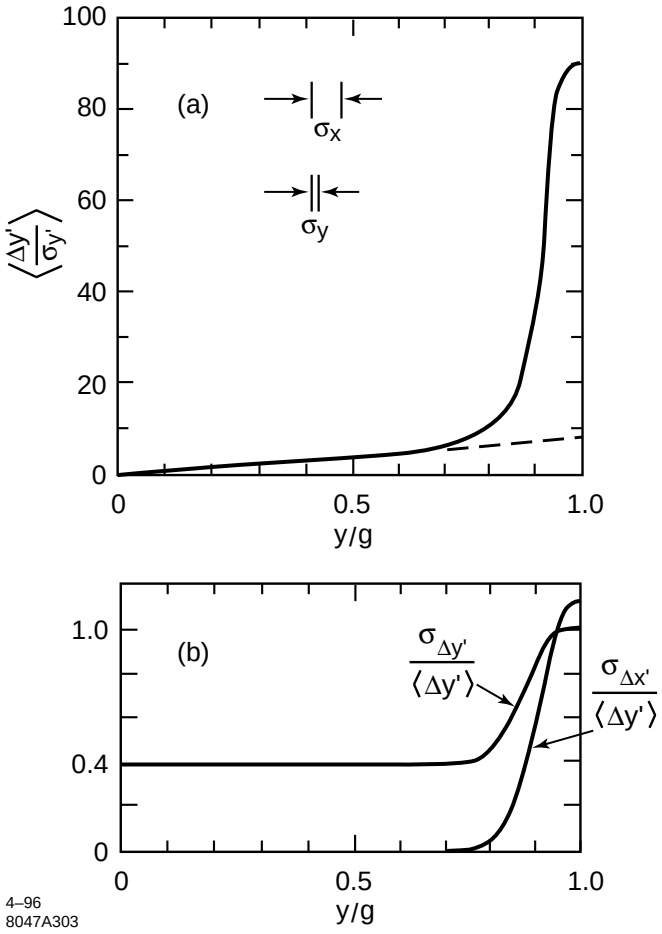


Figure 9-29. *(a) The wake kick as a function of distance to the wall showing the cut-off at the wall when the distance approaches the width of the bunch. (b) An estimate for the variation of the σ of the wake-kicks of (a).*

It is remarkable that both the sigmas of the horizontal and vertical kick are a bit larger than the average vertical kick. An estimate for the variation of these σ with amplitude is shown in Figure 9-29(b).

We have gone in to some detail here because these kicks are large. Since there can be no spoiler in the second stage of the IP collimation, either the collimators at the next stage need to be set back this far, or it could be that the beam is blown up by the kicks and has a sufficiently large area that it is permissible for it to hit the collimator. As mentioned above, the sigma of the kick near the wall is very close to the magnitude of the kick itself. And both the vertical and horizontal beam sizes are blown up. However the beam size must be very large for copper. We shall see that a kick of 135σ is required to blow up the beam so that the area of the beam is larger than the mandatory $(2.2\,\mathrm{mm})^2$.

To get an estimate for the enlarged horizontal σ we use the product

$$\begin{aligned} \text{“}\sigma_{y_2}\text{”} &= R_{34}\sigma_{\Delta y_1'} = \frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle}\frac{\langle \Delta y_1' \rangle}{\sigma_{y_1'}} R_{34}\sigma_{y_1'} = \frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}\sigma_{y_2} \\ \text{“}\sigma_{x_2}\text{”} &= R_{12}\sigma_{\Delta x_1'} = \left\langle \frac{R_{12}}{R_{34}} \right\rangle \frac{\sigma_{\Delta x_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}\sigma_{y_2} \end{aligned} \tag{9.81}$$

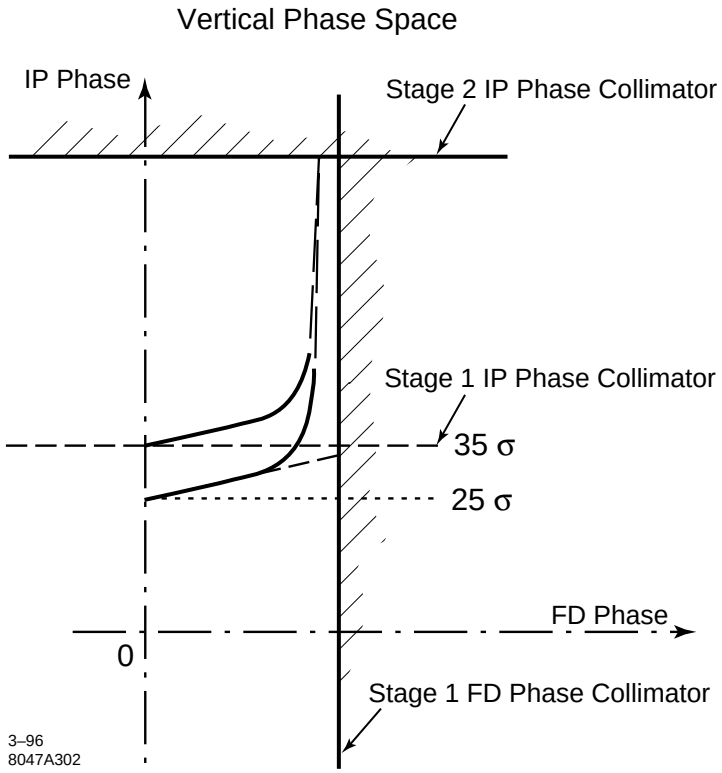


Figure 9-30. *This figure illustrates the enlargement of the aperture that is necessary to avoid impact of a mis-steered beam which narrowly misses upstream spoilers.*

$$\text{“}\sigma_{x_2}\text{”}\,\text{“}\sigma_{y_2}\text{”} \quad = \quad \left\langle \frac{R_{12}}{R_{34}} \right\rangle \frac{\sigma_{\Delta x_1'}}{\langle \Delta y_1' \rangle} \frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}^2 \sigma_{y_2}^2$$

The average $< R_{12}/R_{34} >$ is an rms average weighted by the relative importance of the three collimators in stage 1. For the lattice described in Section 9.2.5 it has a value of about 1/2. Also for this lattice $\sigma_{y2} = 23\,\mu$m. To get the sigma product to be 4.8 mm^2 we will need $n_{y'} = 135$. This is larger than the 118-$\sigma_{y'}$ maximum kick, and will not occur. Figure 9-30 shows the phase space distortions due to these large kicks when the beam is off-axis.

We might conclude that the aperture of vertical IP stage 2 collimation must be (118+35) $\sigma_{y'} = 153\sigma_{y'}$ because the beam could start out near the collimator edge. But if the beam was already out at 35 σ before receiving the wake kick, it would have gotten a rather large kick in the first stage of the IP collimation. That implies that all of the beam could not have been be near the edge. This effect dies off rapidly though, and by an amplitude of 30 σ the kick was 15 σ' with a spread of about 9 σ'. For an amplitude of 28 σ, the spread is 4.5 σ'. The conclusion is that the required aperture is about $(118 + 30)\sigma_{y'} \approx 150\sigma_{y'}$, not much less.

Fortunately the dynamic aperture of the final-focus system is large enough to accommodate these large amplitude particles. See Figure 9-5.

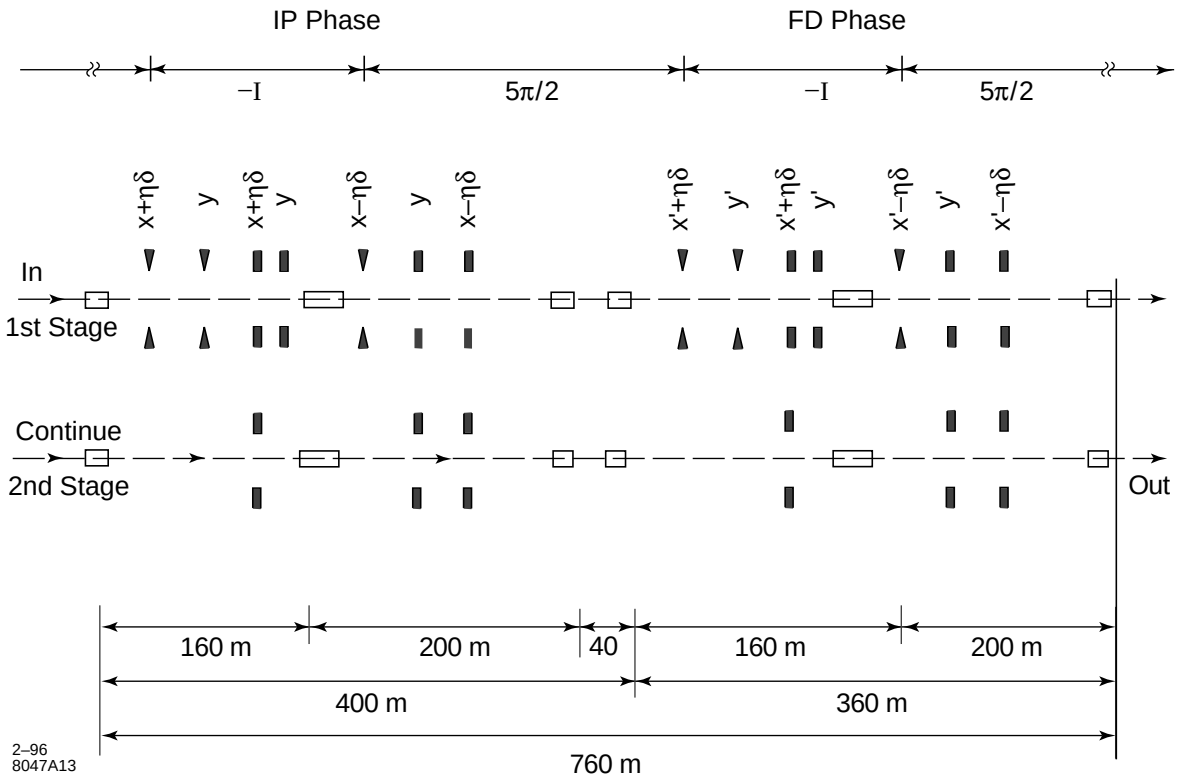


Figure 9-31. *Schematic of the collimation system spoilers and absorbers.*

9.2.5 Lattice Description and Analysis

Collimation System Schematic

Figure 9-31 is a schematic layout of the collimation-system spoilers and absorbers according to the guidelines arrived at in the preceding sections. In this system we have added a second stage for the FD-phase collimation. This is not necessary according to our estimates. The inclusion of a second FD-phase collimation section should be considered tentative. In the diagram of Figure 9-31 we have let a spike represent a titanium spoiler and a rectangular block represent a copper absorber. Further,

- $x + \eta\delta$ indicates collimation outside of the lines $x + \eta\delta = \pm n_x\sigma$ in an $(x\,,\delta)$ plane, with x at the IP phase,
- $x' + \eta\delta$ indicates collimation outside of $x' + \eta\delta = \pm n_x\sigma$ in an $(x'\,,\delta)$ plane, with x' at the FD phase, and
- y and y' indicate collimation at $\pm n_y\sigma$ in the IP and FD phase, respectively.

Figure 9-32 shows the volume in $(x\,,x'\,,\delta)$ space collimated by this arrangement. The horizontal and vertical collimation can be interleaved since on the one hand we need large βs in both planes to achieve sufficient spot area, and on the other hand, the chromaticity can be compensated by interleaved sextupoles without problems arising from octupole aberrations.

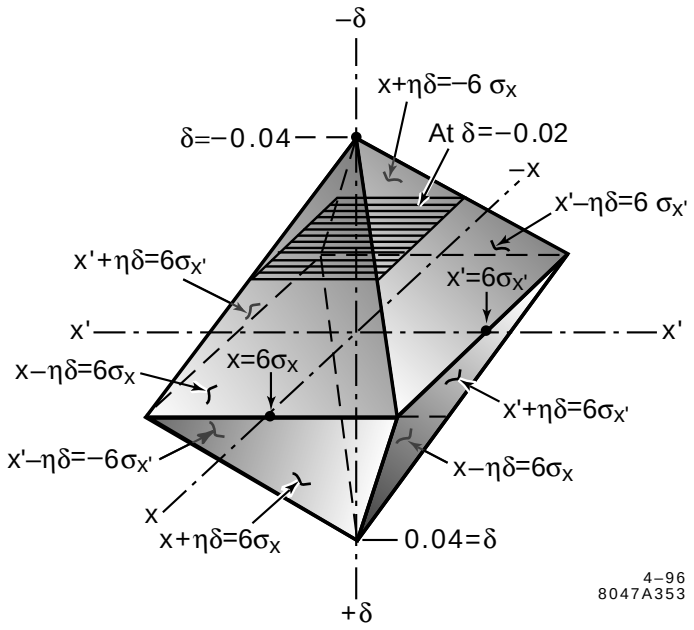


Figure 9-32. *The 3-D volume in x, x', δ space that remains after the beam passes through the spoilers of the collimation system defined in Figure 9-31.*

Each element in the first spoiler and absorber set is located at $-I$ from a corresponding element in the second set. Absorbers with lesser absorption requirements, that will be located downstream of spoilers and absorbers to protect various components and absorb remaining scatter, have not been specifically indicated.

The Lattice

Lattice Functions. Lattice functions for an optical system which satisfies the requirements we have outlined are shown in Figure 9-33 [Helm]. Sextupoles are located at each of the maximum β points.

System Bandpass. Figure 9-34 shows the exit spot size as a function of incident energy for the 1-TeV-c.m. collimation system. At $\delta = \pm 0.5\%$ the spot size has grown by about 4%. This is satisfactory, when one considers that the net result for a beam will be an average over the energy distribution.

Collimator Apertures and Wakes

Collimator Apertures and Small Amplitude Wakes. We will now turn to calculating the minimum values of n_x and n_y that may be collimated. The results are presented in Table 9-5.

For this discussion we will concentrate on the FD phase, and assume that the system consists of two stages, one in the collimation system denoted in Table 9-5 by FD_V_1 and FD_H_1 for the vertical and horizontal plane, respectively, and one in the final-focus system, labeled FD_V_3 and FD_H_3. The data on the second stage of FD collimation is also included but the results are not summed into the totals. In the vertical plane we assume there are one Ti spoiler and one Cu absorber in the first stage and two copper-plated W-Re absorbers in the second stage. In the horizontal plane

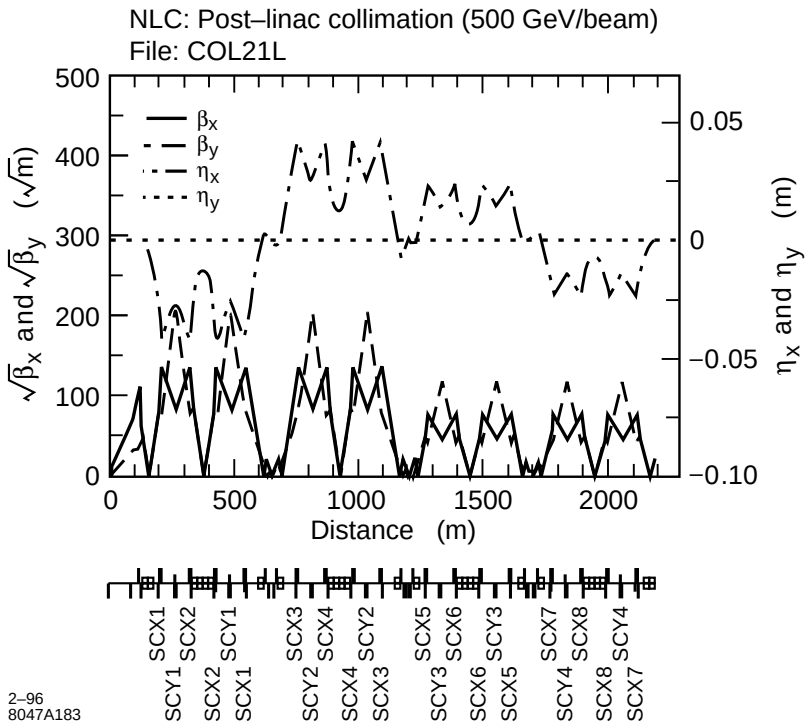


Figure 9-33. *The 1-TeV-c.m. collimation system lattice functions.*

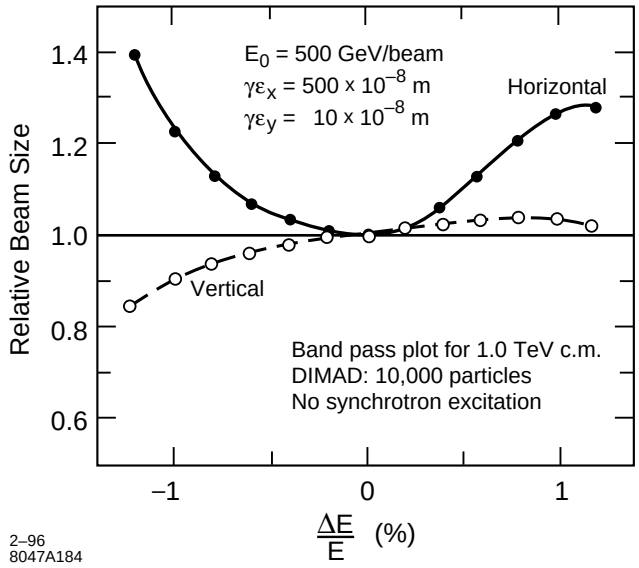


Figure 9-34. *The spot size increase in the 1-TeV-c.m. collimation system as a function of mono-energetic energy offset.*

we will assume there are two Ti spoilers and two Cu absorbers in the first stage and two copper-plated W-Re absorbers in the second stage. The spoilers and absorbers are doubled in the first stage of the horizontal plane because energy is also collimated in this plane.

The n_i have been chosen to produce a t'/t value less than 0.7. They have been expanded one stage to the next to allow for near-wall wake effects and alignment inaccuracies.

In the final-focus system the gaps are determined by a complicated set of relationships that determine optimum βs required for minimizing higher order aberrations, tolerances and system length. At 1-TeV c.m. we have (for the 1.5-TeV design operated at 1.0 TeV, see Figure 11-12) $\beta_{x,max} = 10^5$ and $\beta_{y,max} = 3\ 10^5$. βs are a factor of three smaller for the 0.5-TeV-c.m. design, Figure 11-8. They can be taken to scale approximately linearly with energy. Hence we have chosen the worst-case 1.5-TeV-c.m. β values in Table 9-5.

In the first stage the gaps are determined by the requirements that the βs be large enough so that the impact of a full beam train will not destroy the spoiler. For $\varepsilon_x = 5\ 10^{-12}$ m-rad and $\varepsilon_y = 8\ 10^{-14}$ m-rad, we have the condition that at the spoilers, $\sqrt{(\beta_x \beta_y)} \geq 1.6\ 10^4$, where we used the condition from Table 9-3 that $\sqrt{(\sigma_x \sigma_y)} \geq 100\ \mu$m. For Table 9-5 we have used the values from the lattice of Figure 9-33. Because the parameters have changed during the course of design, this does not quite meet the beta-product specification at the horizontal spoilers. For this lattice $\sqrt{(\beta_x \beta_y)} = 1.1\ 10^4$ at the horizontal spoilers and $\sqrt{(\beta_x \beta_y)} = 1.7\ 10^4$ at the vertical spoilers . The $\sqrt{\beta}$s must be 25% larger at the horizontal spoilers. Another change in this lattice must be made to increase the R_{12} and R_{34} between spoiler and absorber. See Section 9.2.2. On the other hand the second stage of FD collimation can be omitted altogether. Present system length will remain adequate.

In Table 9-5 we see that within the collimation system gaps run between a comfortable 0.8 mm and 2.0 mm. The ratio of the wake kick to the kick of the same collimator with an optimum gap, denoted by ϕ_i, has a maximum value of 2.4.

All absorbers have a length of 20 r.l. If necessary this could be reduced to 10 r.l. In this sense we feel there is some margin in the aperture limits derived.

Collimator Focusing and Near-Wall Kick Strengths. Table 9-6 presents the results from calculating the focusing strengths and near-wall kick strengths of all collimators. The focusing totals agree with the estimations of Section 9.2.4 and, because of a waist location at the IP that varies longitudinally within the bunch, will add 0.5% to the luminosity loss.

The near-wall kick strengths for on-axis beams for the 1st vertical FD stage (FD_V_1) add up to 9.8 σ'. The estimate of Eq. 9.79 equal to 0.28 n_y has the value 10 σ'. Because the second-stage IP collimators must be opened up to 150 σ to insure passive protection from mis-steered beams, the 10 σ plays little role here. And because the second-stage collimators are set this wide, the wake effects on the second (or third) stage FD collimators is tiny total of 0.4 σ.

The second-stage collimators must be opened up because of the large near-wall wakes for mis-steered beams. The total for the resistive-wall and geometric wakes in the second-stage vertical FD phase is 119 σ'. To this we must consider that the beam could have already been at 30 σ in the IP phase without having been too distorted, hence the aperture must be 150 σ. As we have argued in Section 9.2.4, a kick of 135 σ' would imply a beam sigma product of 4.8 mm^2, which would also insure the survival of the collimator from an impact of a full train. In the horizontal plane the 1st stage total is 9.2 σ'. The aperture has been opened to 15 σ to account for the fact that the beam could already have been close to 5 σ. There is no danger of the beam ending up near these apertures and being steered into collimators in the final-focus system.

Location	Sect.	Collimator	$\sqrt{\beta}$ ($\sqrt{m}$)	n_i	g_i (mm)	g_i/g_{0i}	ϕ_i	$t_i'/t_i = (\overline{n}_i/n_i)^2\phi_i$
FD_V_1	1	Ti Spoiler	200	36	2.0	11.1	2.3	0.22
		Cu Absorber	40		0.8	0.8	1.0	0.10
	2	W Absorber	200		1.4	3.7	1.4	0.09
FD_V_2	2	W Absorber	140	40	0.9	2.3	1.2	(0.06)
FD_V_3	1	W Absorber	550	42	6.5	17.0	2.8	0.13
	2	W Absorber	550		6.5	17.0	2.8	0.13
Totals:		**6 Collimators**						**0.66**
FD_H_1	1	Ti Spoiler	140	5	1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.7	1.0	0.09
	2	Ti Spoiler	140		1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.7	1.0	0.09
FD_H_2	1	W Absorber	80	6	1.0	2.8	1.2	(0.04)
	2	W Absorber	80		1.0	2.8	1.2	(0.04)
FD_H_3	1	W Absorber	300	7	4.7	12.1	2.4	0.06
	2	W Absorber	300		4.7	12.1	2.4	0.06
Totals:		**8 Collimators**						**0.63**
IP_V_1	1	Ti Spoiler	200	36	2.0	11.1	2.3	0.22
		Cu Absorber	40		0.8	0.8	1.0	0.10
	2	W Absorber	200		1.4	3.7	1.4	0.09
IP_V_2	2	W Absorber	140	150	3.3	3.3	1.3	0.01
Totals:		**4 Collimators**						**0.42**
IP_H_1	1	Ti Spoiler	140	5	1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.5	1.0	0.09
	2	Ti Spoiler	140		1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.5	1.0	0.09
IP_H_2	1	W Absorber	75	15	2.7	2.6	1.2	0.01
	2	W Absorber	75		2.7	2.6	1.2	0.01
Totals:		**6 Collimators**						**0.52**

Table 9-5. *A list of all the collimators in the collimation and final-focus system showing the number of σ collimated at each and the contribution to the total collimation budget. The β functions from the lattice shown in Figure 9-33 have been used to calculate beam sizes at each collimator. The second stage of the FD collimation is optional and has not been included in totals. The large second-stage IP apertures are dictated by machine protection requirements. Calculations are based on $\alpha_G = \alpha_R = \pi^2/8$.*

Location	Sect	Collim.	Quad "$k\beta$" Strength for On-Axis Bm.		Near Wall Kick for On-Axis Bm.		Near Wall Kick for Near Wall Beam			
			RW	G	RW (σ')	G (σ')	RW nr	RW (σ')	G nr	G (σ')
FD_V_1	1	Ti Spoil	0.01	0.02	1.9	3.4	48	34	24	17.0
		Cu Abs.	0.01	0.01	1.4	1.0	124	37	16	5.0
	2	W Abs.	0.01	0.01	0.8	1.3	40	20	12	6.0
FD_V_2	2	W Abs.	0.00	0.00	0.7	0.9	64	18	16	4.0
FD_V_3	1	W Abs.	0.01	0.01	1.3	2.3	53	102	6	11.0
	2	W Abs.	0.01	0.01	1.3	2.3	53	102	6	11
Totals:		**6 Coll.**	**0.05**	**0.06**	**7.4**	**11.2**				
FD_H_1	1	Ti Spoil	0.01	0.01	0.2	0.3	110	0.7	25	1.5
		Cu Abs.	0.01	0.00	0.1	0.1	270	1.7	11	0.7
	2	Ti Spoil	0.01	0.01	0.2	0.3	180	0.7	25	1.5
		Cu Abs.	0.01	0.00	0.1	0.1	270	1.7	11	0.7
FD_H_2	1	W Abs.	0.00	0.00	0.1	0.1	240	0.9	13	0.5
	2	W Abs.	0.00	0.00	0.1	0.1	240	0.9	13	0.5
FD_H_3	1	W Abs.	0.00	0.00	0.1	0.2	6	0.8	6	0.9
	2	W Abs.	0.00	0.00	0.1	0.2	6	0.8	6	0.9
Totals:		**8 Coll.**	**0.04**	**0.02**	**0.9**	**1.4**				
IP_V_1	1	Ti Spoil	0.01	0.02	1.9	3.4	48	34.0	24	17
		Cu Abs.	0.01	0.01	1.4	1.0	120	37.0	16	5
	2	W Abs.	0.01	0.01	0.8	1.3	40	20.0	12	6
IP_V_2	2	W Abs.	0.000	0.00	0.5	0.6	240	69.0	10	3
Totals:		**4 Coll.**	**0.03**	**0.04**	**4.6**	**6.3**				
IP_H_1	1	Ti Spoil	0.01	0.01	0.2	0.3	18	1.1	25	1.5
		Cu Abs.	0.005	0.00	0.1	0.1	27	1.7	11	0.7
	2	Ti Spoil	0.01	0.01	0.2	0.3	18	1.1	25	1.5
		Cu Abs.	0.005	0.00	0.1	0.1	27	1.7	11	0.7
IP_H_2	1	W Abs.	0.00	0.00	0.1	0.1	76	2.7	10	0.4
	2	W Abs.	0.00	0.00	0.1	0.1	76	2.7	10	0.4
Totals:		**6 Coll.**	**0.03**	**0.02**	**0.8**	**1.0**				

Table 9-6. *A list of all the collimators in the collimation and final-focus system showing the number of σ collimated at each and the contribution to the total collimation budget. The β functions from the lattice shown in Figure 9-33 have been used to calculate beam sizes at each collimator. Geometric wake numbers should be considered to be tentative for parallel-plate tapers.*

Tolerances

The lattice has large β functions and some very large R_{12} and R_{34} values which will lead to important tolerances on quadrupole strength and position stability and on sextupole position stability. Since we are designing near the limit allowed by wakefield intensities, we can also expect important tolerances on collimator jaw stability.

There are four important time domains for tolerance specification:

- capture tolerances,
- long term stability,
- short term stability, and
- jitter tolerances.

Capture tolerances are the absolute alignment or strength tolerances required at start-up so that the system can be aligned and tuned using beam-based techniques.

Long term stability tolerances are those limits that when exceeded would require re-execution of these beam-based alignment strategies.

Several tuning knobs will be activated perhaps every hour at the IP to adjust sensitive aberrations affecting the IP beam spot size. Change in waist and skew aberrations can also be observed in the skew correction system (SCS). Short-term stability tolerances are the tolerances that must be held until tuning knobs are reset or aberrations are rechecked and corrected. We estimate that the SCS scans occur ten times per hour, if required, since this process can be non-invasive.

The beam centroid is stabilized with feedback systems based on orbit measurements and a rule of thumb estimate is that they are effective for times longer than or equal to 15 beam pulses, or about 1/10th of a second. Jitter tolerances refer to requirements on the stability of the system for times shorter than this time.

Except for initial collimator jaw alignment, this section discusses only the stability tolerances associated with the last two time scales.

Tolerance Budgets. The alignment of the beam collision at the IP is maintained with a fast feedback system. As noted above, the jitter tolerances refer to beam-line changes on time scales which can not be corrected by this feedback system.

In addition to limiting luminosity loss, a budget is necessary because with a large beam jitter it becomes difficult to carry out beam-based alignment, wakefield effects are enhanced, and synchrotron radiation in the final doublet is increased. (See Section 11.7). Table 9-7 is a proposal for a jitter budget.

There are many aberrations which must be tuned in the final-focus system using the spot size at the IP as the diagnostic tool (see Table 11-9). Once the aberrations are tuned, changes in the beam line can cause an increase in aberration strength, and that aberration must be re-tuned, or compensated based on an SCS measurement. Several of these aberrations can result from changes in the collimation beam line. These are: horizontal and vertical waist, horizontal and vertical dispersion, and one skew aberration. The collimation system has been allotted 0.5% luminosity change for each of these seven aberrations.

Quadrupole Strength Stability. Each FD-phase section has four quadrupoles at particularly large values of the vertical β function, about 30 km. A small change in the strength of any of these quads would cause a shift of the

Machine Section	Steering Budget (H/V) σ	Accumulated Jitter (H/V) σ	Luminosity Loss (H/V) %
To end of linac	0.1 / 0.25	0.10 / 0.25	0.5 / 0.8
Collimation system	0.1 / 0.2	0.14 / 0.32	1.0 / 1.3
w/ amplification		0.18 / 0.40	1.6 / 2.0
Big bend	0.05 / 0.1	0.19 / 0.41	1.8 / 2.1
Final focus	0.1 / 0.2	0.21 / 0.46	2.2 / 2.6
Final doublet	0.1 / 0.2	0.23 / 0.50	2.7 / 3.1

Table 9-7. *A proposal for a jitter budget.*

waist at the IP. On the assumption that changes in the beam size can be measured with a relative resolution of 10%, it follows that each independent aberration can be adjusted so as to leave a residual increase in beam size of only 2%. SLC experience indicates this may be less than 0.5%. The waist position is one of these aberrations, thus a waist-stability budget of 2% must be apportioned to the various modules of the collider. Let us assume that 1/2 of this is assigned to the collimator system, and of that is divided further into 1/2 for quadrupole-strength stability and 1/2 for sextupole-position stability. Hence the quadrupole-stability budget is 1/2%. Assuming that the majority of this can be assigned to the sixteen quadrupoles at large vertical β, and that they are on independent power supplies, it can be estimated that

$$\frac{dk}{k} \leq \frac{1}{60}\ \frac{1}{k\beta} \tag{9.82}$$

For our lattice $k\beta = 1,200$ hence $dk/k < 1.3\ 10^{-5}$.

Based on sensitivity calculations for each element, Table 9-8 assigns strength stability tolerances to all collimation system quadrupoles so that the collimation system meets the allotted 0.5% horizontal and vertical waist budget. We have conservatively included both phases in the list of Table 9-8, hence there are in fact 32 quadrupoles that have been assigned a tolerance of $dk/k < 1\ 10^{-5}$. This could be reduced to 16 with a tolerance of $1.8\ 10^{-5}$.

Sextupole Position. The results of this section are summarized in Table 9-9.

- Horizontal Sextupole Position—The calculation of the preceding section can be used to estimate the horizontal position stability of the sextupole since any sextupole motion adds a quadrupole field to the beam line of magnitude $k_S dx$. For a 0.5% contribution to spot size increase the displacement would be limited by the relationship (valid for the FD-phase sextupoles only)

$$k_s dx \leq \frac{1}{10}\ \frac{1}{\beta_y} \tag{9.83}$$

  depending on whether one is looking at the vertical waist or horizontal waist.

  For the sextupoles of the stage 1 of the FD phase at large $\beta_x = 1.7\ 10^4$ where $k_S = 1.16\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 5\ \mu$m. For the sextupoles of the stage 2 of the FD phase at large $\beta_x = 5.3\ 10^3$ where $k_S = 2.6\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 7.5\ \mu$m. These four pairs sextupoles contribute to the horizontal waist aberration. Since the quadrupole stability tolerances are mandated by vertical waist considerations, the entire 1/2% horizontal waist aberration budget may be allocated to these four sextupole pairs. If we require the same tolerance for all pairs, the result is $dx < 3\ \mu$m.

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
Q703	17.00	1.00	1.900	0.100	11.00	0.5
Q704	15.00	1.00	0.810	0.050	113.00	1.0
Q704	15.00	1.00	0.820	0.050	12.00	0.5
Q705	18.00	1.00	16.000	1.000	2.90	0.2
Q705	18.00	1.00	58.000	1.000	2.90	0.2
Q706	11.00	1.00	0.710	0.050	25.00	0.5
Q706	11.00	1.00	0.720	0.050	32.00	0.5
Q707	9.50	1.00	2.800	0.100	5.40	0.5
Q707	9.80	1.00	2.900	0.100	5.40	0.5
Q708	10.00	1.00	4.800	0.300	19.00	1.0
Q708	10.00	1.00	4.700	0.300	19.00	1.0
Q709	6.90	1.00	4.100	0.300	23.00	1.0
Q709	6.90	1.00	3.900	0.300	24.00	1.0
Q710	20.00	1.00	2.200	0.100	76.00	1.0
Q710	20.00	1.00	2.200	0.100	66.00	1.0
QCL1	94.00	1.00	13.000	1.000	230.00	1.0
QCL2	12.00	1.00	16.000	1.000	13.00	0.5
QCL3	0.15	0.01	10.000	1.000	140.00	1.0
QCL4	0.54	0.03	14.000	1.000	34.00	0.5
QCX1	0.18	0.03	16.000	1.000	130.00	1.0
QCX2	0.13	0.01	41.000	1.000	210.00	1.0
QCX2	0.12	0.01	45.000	1.000	220.00	1.0
QCX3	0.11	0.01	310.000	1.000	360.00	1.0
QCX3	0.11	0.01	370.000	1.000	310.00	1.0
QCX2	0.12	0.01	56.000	1.000	42.00	0.5
QCX2	0.14	0.01	49.000	1.000	43.00	0.5
QCX1	0.18	0.01	19.000	1.000	37.00	0.5
QCX1	0.18	0.01	16.000	1.000	130.00	1.0
QCX2	0.13	0.01	41.000	1.000	210.00	1.0
QCX2	0.12	0.01	45.000	1.000	220.00	1.0
QCX3	0.11	0.01	310.000	1.000	360.00	1.0
QCX3	0.11	0.01	370.000	1.000	310.00	1.0
QCX2	0.12	0.01	56.000	1.000	42.00	0.5
QCM1	0.16	0.01	57.000	1.000	50.00	0.5
QCM2	0.23	0.03	24.000	1.000	45.00	0.5
QCM3	5.60	0.10	10.000	1.000	75.00	0.5
QCM4	98.00	1.00	5.600	0.300	37.00	0.5
QCM4	108.00	1.00	6.300	0.300	29.00	0.5
QCM3	7.30	0.10	0.850	0.050	2.00	0.2
QCM2	0.22	0.03	0.120	0.025	0.42	0.05
QCM1	0.15	0.01	0.210	0.030	0.43	0.05

Table 9-8. *The vibration and strength stability sensitivities and tolerances of all the quadrupoles in the collimation system. The sensitivities are calculated as if the one quadrupole was assigned the entire budget of the section. There are about 100 quadrupoles, so in general the tolerance is 1/10th of the sensitivity. Continued ...*

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
QCX2	0.12	0.01	0.170	0.030	0.36	0.05
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCX2	0.14	0.01	0.180	0.030	0.38	0.05
QCX1	0.18	0.01	0.097	0.025	0.36	0.05
QCX1	0.18	0.01	0.097	0.025	0.36	0.05
QCX2	0.13	0.01	0.180	0.030	0.38	0.05
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCM1	0.15	0.01	0.210	0.030	0.45	0.05
QCM2	0.23	0.03	0.130	0.025	0.44	0.05
QCM3	5.60	0.10	1.100	0.100	1.80	0.20
QCM4	98.00	1.00	5.800	0.300	35.00	0.50
QCN4	98.00	1.00	4.900	0.300	51.00	0.50
QCN3	11.00	1.00	8.800	0.300	43.00	0.50
QCN2	0.72	0.03	11.000	1.000	44.00	0.50
QCN1	0.47	0.03	24.000	1.000	52.00	0.50
QCX2	0.38	0.03	24.000	1.000	46.00	0.50
QCX3	0.33	0.03	100.000	1.000	1300.00	1.00
QCX3	0.34	0.03	110.000	1.000	710.00	1.00
QCX2	0.38	0.03	35.000	1.000	35.00	0.50
QCX2	0.42	0.03	30.000	1.000	36.00	0.50
QCX1	0.57	0.03	11.000	1.000	29.00	0.50
QCX1	0.57	0.03	8.700	0.300	36.00	0.50
QCX2	0.42	0.03	22.000	1.000	46.00	0.50
QCX2	0.38	0.03	24.000	1.000	46.00	0.50
QCX3	0.33	0.03	100.000	1.000	1300.00	1.00
QCX3	0.34	0.03	110.000	1.000	710.00	1.00
QCX2	0.38	0.03	35.000	1.000	35.00	0.50
QCP1	0.48	0.03	34.000	1.000	40.00	0.50
QCP2	0.75	0.03	14.000	1.000	36.00	0.50
QCP3	9.40	0.10	8.800	0.300	61.00	0.50
QCP4	73.00	1.00	5.800	0.300	47.00	0.50
QCP4	78.00	1.00	6.200	0.300	38.00	0.50
QCP3	11.00	1.00	1.900	0.100	3.10	0.20
QCP2	0.72	0.03	0.220	0.030	0.78	0.10
QCP1	0.47	0.03	0.370	0.030	0.76	0.10

Table 9-8. *Cont.*

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
QCX2	0.38	0.03	0.320	0.010	0.64	0.10
QCX3	0.33	0.03	0.240	0.030	2.10	0.20
QCX3	0.34	0.03	0.240	0.030	2.10	0.20
QCX2	0.38	0.03	0.320	0.030	0.64	0.10
QCX2	0.43	0.03	0.320	0.030	0.68	0.10
QCX1	0.57	0.03	0.170	0.030	0.64	0.10
QCX1	0.57	0.03	0.170	0.030	0.63	0.10
QCX2	0.41	0.03	0.330	0.030	0.67	0.10
QCX2	0.38	0.03	0.320	0.010	0.65	0.10
QCX3	0.33	0.03	0.240	0.030	2.10	0.20
QCX3	0.34	0.03	0.240	0.030	2.10	0.20
QCX2	0.38	0.03	0.320	0.030	0.65	0.10
QCX2	0.43	0.03	0.330	0.030	0.68	0.10
QCX1	0.57	0.03	0.170	0.030	0.65	0.10
QMD1	3.00	0.10	0.490	0.050	1.70	0.20
QMD1	2.20	0.10	0.520	0.050	1.40	0.20
QMD2	0.98	0.10	0.570	0.050	0.90	0.10
QMD2	0.99	0.10	0.560	0.050	0.90	0.10
QMD3	3.60	0.10	0.740	0.050	6.80	0.20
QMD3	3.80	0.10	0.760	0.050	7.50	0.20
QMD4	25.00	1.00	1.100	0.100	5.90	0.20
QMD4	26.00	1.00	1.100	0.100	5.70	0.20

Table 9-8. *Cont.*

For the sextupoles of stage 1 of the FD phase at large, $\beta_y = 4\ 10^4$ where $k_S = 1.5\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 1.6\,\mu$m. For the sextupoles of stage 2 of the FD phase at large $\beta_y = 1.3\ 10^4$ where $k_S = 2.8\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 2.7\,\mu$m. These sensitivities are based on the vertical waist aberration, for which we have budgets 0.5% for both pairs. If the tolerance of both pairs is taken to be identical, the result is $dx \leq 1.4\,\mu$m.

The exact meaning of these sextupole-position tolerances must be stated with some care. A change in the position of any quadrupole between large-β_y paired sextupoles that would cause a displacement of the beam at the second sextupole by 1.4 μm would give rise to the vertical waist aberration. Thus one sees that it is not the absolute position of the sextupole that is of concern, but its position with respect to the sextupole with which it is paired, and the position of this sextupole pair with respect to the beam line as defined by the quadrupoles. For details on stabilization strategies see Section 9.2.6 below.

- Vertical Sextupole Position—A change in the vertical sextupole position gives rise to a skew quad aberration of strength $k_S dy$. For a 0.5% spot size increase we have, for FD-phase sextupoles, the condition

$$k_s dy \leq \frac{1}{10}\sqrt{\frac{\varepsilon_y}{\varepsilon_x}}\,\frac{1}{\sqrt{\beta_x \beta_y}} \tag{9.84}$$

At the $\beta_{y,\max}$ in stage 1 we have $\sqrt{(\beta_x\beta_y)} = 1.7\,10^4$. The resulting 0.5% sensitivity, for $k_s = 1.5\,\mathrm{m}^{-2}$, is $dy \leq 0.5\,\mu$m. At the $\beta_{x,\max}$ in stage 1 we have $\sqrt{(\beta_x\beta_y)} = 0.95\ 10^4$ and $k_s = 1.1m^{-2}$, with a sensitivity $dy \leq 1.2\,\mu$m.

Beam-line Location	Sextupole Name	Horizontal Sensitivity (μm)	Horizontal Tolerance (μm)	Vertical Sensitivity (μm)	Vertical Tolerance (μm)
V_FD_1	SCY1	1.6	1.4	0.5	0.3
H_FD_1 (2)	SCX1&2	5.0	3.0	1.2	0.4
V_FD_2	SCY4	2.7	1.4	0.8	0.4
H_FD_2 (2)	SCX7&8	7.5	3.0	1.6	0.5

Table 9-9. *Stability sensitivities and tolerances for collimation system sextupole positions.*

If the 1/2 percent budget is divided with 1/6% going to each of the $\beta_{x,\max}$ sextupoles and 1/3% going to the $\beta_{y,\max}$, we obtain tolerances of $dy \leq 0.7$ μm and $dy \leq 0.4$ μm, respectively.

At the $\beta_{y,\max}$ in stage 2 we have $\sqrt{(\beta_x \beta_y)} = 5.3\ 10^3$. The resulting sensitivity, for $k_s = 2.8\,\mathrm{m}^{-2}$, is $dy \leq 0.8$ μm. At the $\beta_{x,\max}$ in stage 2 we have $\sqrt{(\beta_x \beta_y)} = 3.1\ 10^3$ and $k_s = 2.6\,\mathrm{m}^{-2}$, with a sensitivity $dy \leq 1.6$ μm.

If the 0.5% budget of the collimation system for the principal skew aberration is divided amongst these six sextupole pairs, one possible assignment are the tolerances listed in Table 9-9. We note that the vertical sextupole position tolerance has the same meaning as the horizontal tolerances discussed in the previous section. See Section 9.2.6 for stabilization strategies.

Energy Jitter Considerations. Variations in bunch intensity give rise to variations in energy through beam-loading effects in the linac, hence the energy can vary bunch-to-bunch within a bunch train, and also bunch-train to bunch-train. This variation can be no larger than the band width of the final-focus system, about $\pm 0.4\%$, but we consider here whether such variations give rise to wakes in the collimation system. This jitter is about 1/10th of 4% and hence is 1/10th of the collimation aperture.

By designing the collimation system with a symmetric dispersion function and in such a way that the transverse wake is the same magnitude at the horizontal spoiler and/or absorber at each end of the $-I$s, the energy jitter will produce the same $\Delta x'$ at the second collimator as the first. However because of the $-I$ transform between these locations, the kicks cancel. Thus a properly designed collimation system can be insensitive to energy jitter.

Quadrupole Position. Several quadrupoles between sextupoles have $k_Q R_{12}$ multipliers to the sextupole of 2.2. Thus the tolerances in Table 9-9 must be divided by 2.2 to get the stability tolerance for such quadrupoles.

Jitter of a quadrupole in the FD phase will cause beam jitter at the IP. Using 1/2% of the jitter budget for the collimation system, and assuming that the 8 quadrupoles located at $\beta_{\max}$ points are moving independently, the tolerance on the quadrupole jitter is given by

$$\Delta y < \frac{1}{5}\,\frac{1}{\sqrt{8}}\,\frac{\sigma_{y'}}{k_Q} = \frac{1}{5}\,\frac{1}{\sqrt{8}}\,\frac{\sigma_y}{k_Q \beta_y} \tag{9.85}$$

Since $\sigma_y = 40$ μm, and $k\beta_y = 1,200$, the tolerance on jitter is $\Delta y < 2$ nm. A similar consideration of the quadrupoles at the β_x maximum leads to a horizontal jitter tolerance of $\Delta x < 2$ nm. The $k\beta$ is smaller, but there are twice as many quadrupoles and there is no help from disruption. See Table 9-8 for a list of vibration tolerance for all quadrupoles in the collimation system. These vibration tolerances are chosen so that the contribution from all quadrupoles in the collimation system adds up to at most a 0.5% spot size increase.

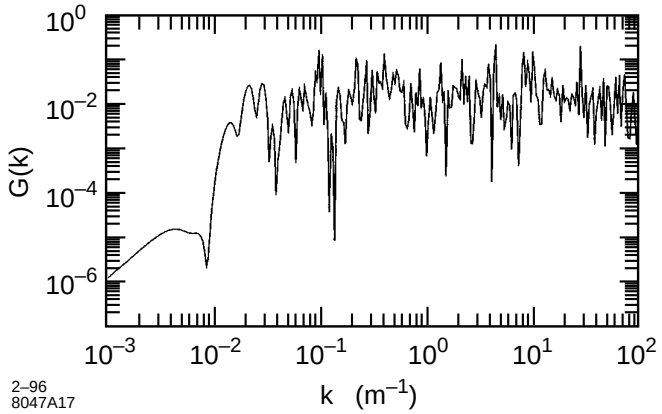


Figure 9-35. *The vertical lattice response function for the 1-TeV-c.m. collimation system.*

Note that these vibration tolerances do not apply to seismic ground motion. They can be taken to be the tolerance of the quadrupole vibration with respect to ground beneath it, or to high frequency cultural noise. In the FFTB the vibration between the quadrupole and the ground beneath it is the order of 1 nm, hence the stands and vibration from cooling water in the FFTB meet the requirements of the vibration tolerances of Table 9-8.

The seismic motion, although larger than the tolerances calculated above, is very correlated at low frequencies (smaller than 5 Hz) and the spectrum is small for frequencies above 5 Hz in tunnels below the ground surface. See Appendix C for details on ground motion. Because of the correlation, one must calculate a lattice response function to see the net effect on the IP beam positions. Figure 9-35 shows this lattice response function for the collimator section for the vertical plane. The response function is also discussed in Appendix C. and its application to the final-focus lattice is discussed in Section 11.5.5.

The lattice response function gives the square of the displacement at the IP for a unit amplitude wave of wave number k in the collimation section. At $k > 0.1$ (corresponding to $\lambda < 60$ m, or with a velocity of 600 m/s at frequencies $f > 10$ Hz) the response function average is about 0.04. This means if we adopt the IP position requirement as 0.5 nm, the tolerance in the collimation system is 2.5 nm. This number is larger than quiet ground conditions in this frequency region, but care must be taken not to introduce noise of this magnitude. At lower k the tolerance drops off abruptly. At $k = 0.01$ (corresponding to $\lambda < 600$ m, or with a velocity of 600 m/s a frequency of $f > 1$ Hz) the response function $G(k)$ is 10^{-5}, corresponding to a tolerance of 160 nm. The integrated ground motion to this frequency is only a few nm, so is well within this limit.

One can fold the quiet spectrum ground motion together with the lattice response function to calculate the relative beam motion at the IP due to the entire spectrum of seismic noise. The result is a beam displacement at the IP that is the order of 0.1 nm for seismic motion in the collimation system.

Dipole Strength Stability. The tightest tolerance on dipoles occurs for dipoles between the vertical spoiler and vertical absorber of stage 1 of the FD-phase collimation. Since the dispersion function at these sextupoles is $\eta = .026$ m, we may deduce that $\Sigma\theta_i R_{12,i} = 2\eta = 0.052$ m, where the sum is over the bends between the sextupoles. The tolerance condition may be written as $\Sigma\Delta\theta_i R_{12,i} \leq 1.4\,\mu$m. Since all θ_i are equal, these two relations can be combined to conclude that $\Delta\theta/\theta \leq 2.7\ 10^{-5}$. Using the relation for the two sextupole pairs at the horizontal spoilers and absorbers we have $\Sigma\theta_i R_{12,i} = 0.08$ m and $\Sigma\Delta\theta_i R_{12,i} \leq 3\,\mu$m, we conclude $\Delta\theta/\theta \leq 3.8\ 10^{-5}$. Since these are different aberrations, the smallest of the two estimates prevails.

Beam Line Location	Dipole Name	$\Sigma\theta_i R_{12,i}$ (cm)	0.02 σ (mm)	Jitter Sens. (10^{-5})	Jitter Toler. (10^{-5})	Stab. Toler. (10^{-5})
Angle Adjust	B2M (2)					N/A
Match In	BCA1 (2)	4	6	15	5	N/A
IP_1	BCA1 (4)	8	6	8	5	N/A
IP→FD	BCM1&2	4/4	6/6	15/15	5	N/A
FD_1	BCA2 (4)	8	6	8	2.7	2.7
FD→IP	BCN1&2	4/2	6/4	15/20	5	N/A
IP_2	BCB1 (4)	4	4	10	5	N/A
IP→FD	BCP1&2	2/2	4/4	20/20	5	N/A
FD_2	BCB2 (4)	4	4	10	4.8	4.8
Match Out	BCB2 (2)	2	4	20	5	N/A

Table 9-10. *Tolerances for the dipoles in the collimation system.*

For comparison, the tolerance for the dipoles of the stage 2 FD phase are $\Delta\theta/\theta \leq 4.8\ 10^{-5}$, the vertical pair ($\Sigma\theta_i R_{12,i} = 0.029\,\text{m}, \Sigma\Delta\theta_i R_{12,i} \leq 1.4\,\text{mm}$), and $\Delta\theta/\theta \leq 6.7\ 10^{-5}$ for the horizontal pairs ($\Sigma\theta_i R_{12,i} = 0.045\,\text{m}, \Sigma\Delta\theta_i R_{12,i} \leq 3\,\text{mm}$). The smallest of these will prevail.

The tolerances calculated in the previous two paragraphs are stability tolerances. Beam spray and change of centroid due to beam loss at collimators, would make it very difficult to use beam-based techniques in FD_1. Since these tolerances relate to conditions at the IP, and do not affect collimation system function, they may be monitored and compensated in the skew correction system. Also, see Section 9.2.6.

There are no similar stability implications for the remaining dipoles, and tolerances can be considered as jitter tolerances. The dipoles of the match in and match out would give jitter in the FD phase. The transition-region dipoles between IP and FD can cause jitter in both planes. The dipoles within a stage, can cause jitter at the phase of that stage. The change in position due to a change in angle can be found from the dispersion function change created by the bends. These are listed in column 3 of Table 9-10.

To calculate the tolerances for the remaining dipoles we use the condition that the dipoles should not give rise to detectable beam jitter. Quantitatively we will require that the contribution to jitter by all collimation system dipoles should be less than 0.02 σ (this should be chosen to be negligible compared to expected jitter values). This number, at the location where the dispersion was measured, is given in column 4 of Table 9-10. The sensitivities are the ratios of the dispersion function to the 0.02 σ numbers. These are given in column 5 of Table 9-10.

The sensitivities are equally divided between the two phases. Assuming that the bends in each location listed in Table 9-10 are on the same power supply, and that the errors in the bends at different locations are independent, we arrive at the tolerances listed in column 5.

Collimator Jaw Stability. In this section we investigate the effect of the wake from the two horizontal collimators between the vertical sextupoles on the offset of the beam at the downstream sextupole. Pulse-to-pulse displacement at the downstream sextupole from the wakefield kick must be smaller than the sextupole position tolerance of $dx = 0.3\,\mu$m.

The horizontal collimators are almost π apart in x-phase, hence the beam jitter at the horizontal collimators will give rise to a wake kick of opposite sign. But since the R_{12}s to the vertical sextupoles are the same sign, the net displacement at the vertical sextupole will be zero. In this regard it is important that the two collimators have the same wakefield kick strength.

If one of the jaws moves there will be a net kick of course. Since the wake kick from a 1-σ jitter is less than 0.7 σ' for the sum of all spoilers or absorbers at a given phase, the wakefield from a single collimator for a 1-σ motion will give a kick of about 0.7/4 $\sigma' \approx 0.2\sigma'$. Since 1σ at the horizontal collimators is about 180 μm, the kick will be less than 0.2 (17 nr)=3.4 nr. The displacement at the downstream "vertical" sextupole will be about 0.17 μm.

This consideration also yields the stability tolerance on the collimator jaw of about 320 μm. This is very loose, and corresponds to the absolute alignment tolerance for the jaws which will be calculated next.

If these considerations had been carried out for the vertical collimators the tolerance would be about 1σ, which for the vertical collimation is about 40 μm.

Collimator Jaw Alignment. If the beam passes off-center through a collimator the centroid receives a kick, and also the beam emittance grows slightly. These effects were calculated in Section 9.2.4. For an optimized taper it was shown that the luminosity loss due to emittance growth was 2.3 times smaller than the luminosity loss due to the centroid kicks. When kicks arise because of beam jitter, they all have the same sign, and the luminosity loss goes as the number of collimators squared. When the collimator jaws are misaligned, the kicks will have a random sign, and thus the luminosity loss will grow linearly with the number of collimators. So we can calculate the luminosity loss due to each collimator and then add the effects.

Suppose we assume a 1-σ collimator jaw misalignment. Then the t-ratio for that collimator gives the kick magnitude (in numbers of σ'). The luminosity loss, from emittance growth, due to that collimator will be $0.055(t/t')^2$ (see Eq. 9.40). So the luminosity loss for all collimators at a given phase will be $0.055\Sigma(t/t')_i^2$. For the vertical final-doublet phase (FD_V) this sum is 0.1. Hence a random displacement of the vertical collimators by 1σ would result in a luminosity loss of 1/2%. (Note that the emittance growth which occurs in the IP phase does not effect the luminosity. The growth occurring in this phase effects the angular distribution at the IP.) The horizontal sum is 0.05 so the luminosity loss from a random 1σ horizontal collimator displacement would be 1/8%.

Thus the tolerances on the collimator jaw alignment can be taken to be about 1σ. This would result in a total luminosity loss of 3/4%. See Section 9.2.6 for a discussion of how to achieve this alignment. Also see Section 11.5.5 in Chapter 11.

Particle Tracking and Mapping Studies

Tracking Between Collimator Apertures. In order to insure the passive protection of the collimation system absorbers and the final-focus system collimators, it is essential that they lie in the shadow of the initial collimation system spoilers. Our concern is that the presence of uncompensated sextupoles and chromaticity within the system could lead to unexpected beam offsets at the absorbers and collimators. To check this we have traced particles through the system.

The clearest method we have found to verify this functionality and diagnose any failures is to initialize particles uniformly in an x-x' or y-y' (or both simultaneously) phase space for a fixed energy offset δ. In practice one need consider only negative values of δ, since large positive values of δ can not occur. Figure 9-36 shows the result of such an exercise for the horizontal phase space for an energy off-set of $\delta = -0.02$. The profile is shown at four important horizontal absorbers in the collimation system, and for the first collimator in the final-focus system. Two horizontal

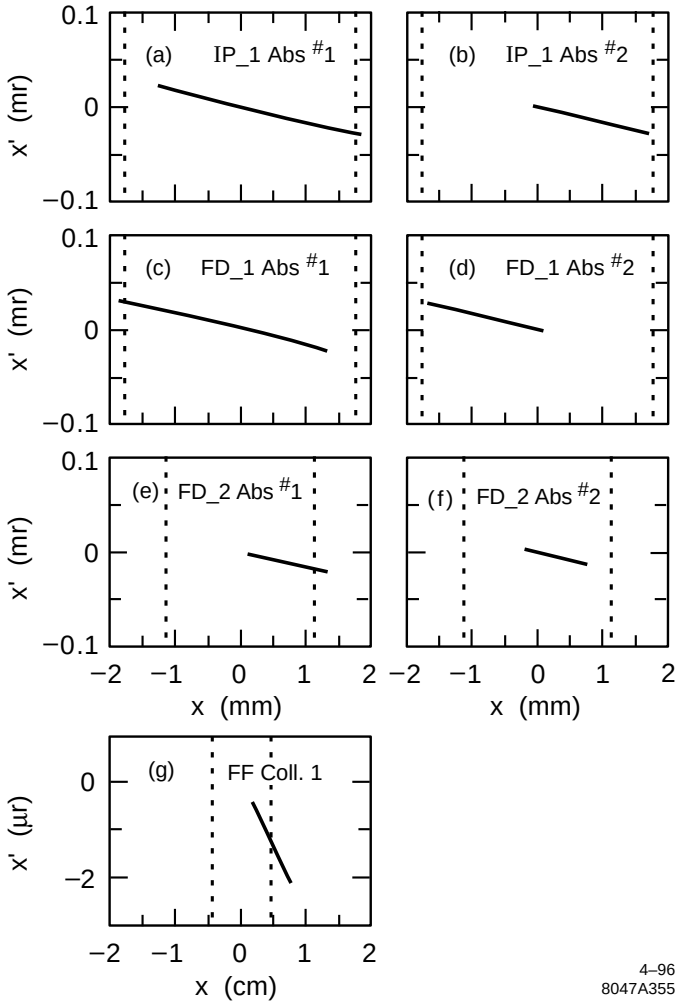


Figure 9-36. *This figure shows the downstream image of particles released at the entrance to the collimation system at $\delta = -0.02$ and on a 6σ square in $x - x'$ phase space. The four pictures are the image at all absorbers in the collimation system and at the collimators in the final-focus system.*

absorbers not shown have very large apertures and two are in the stage-2 FD-phase section which we have argued is unnecessary.

We first note that the distributions of Figure 9-36 look like straight lines since the x' range arising from even small α functions is much larger than the internal spread of x' within the bunch. Secondly we note that except for the profile at the final-focus collimator, the distributions fall within the collimated aperture, or in the few exceptions, the apertures can be enlarged to accommodate. The apertures of the absorbers in IP_1 and FD_1 can be enlarged to accommodate the slight beam offsets. They must be set back somewhat to avoid beam that could pass near the spoilers and be deflected into the absorbers. The apertures of FD_2 are not crucial, and can be enlarged, since this phase of collimation could be omitted altogether.

The large offset of the profile at the collimator in the final-focus system was a surprise to us because the profiles tracked in an earlier lattice, that were shown above in Figure 9-6, were not a problem. The analysis of this situation is described in the following paragraph.

Mapping between Collimator System and IP. The results of tracking described in the previous section show that there is a problem with the beam displacement at the first collimator in the final-focus section. The changes between this lattice and an earlier lattice that did not have this problem are: i) a modified big bend without chromatic correction, ii) the addition of a 400-m skew correction section, iii) a new angle adjustment module at the entry to the final-focus system, and iv) an enlarged beta match to the final focus which includes a phase adjustment capability.

To pinpoint the source of the position offset, we created maps [Yan] of each section of the transport line from the collimation system to the final-focus system. Analysis of these maps showed that linear chromaticity and δ^2 dispersion must be corrected. When these terms are corrected, the image falls within the aperture. Figure 9-37 shows the maps for each section and the complete beam line before the chromatic corrections are introduced, and Figure 9-38 shows the situation afterward. The pictures are drawn in a normalized phase space with the IP phase horizontal and the FD phase vertical.

As pointed out, the stage-2 FD-phase collimation is redundant, and can be removed. Then the IP-phase stage-2 collimation will be located between the last FD-phase collimation and the final-focus system and could be used for chromatic correction.

Next-generation lattice. Following up on the mapping studies described in the previous paragraph, we have developed a next-generation lattice that incorporates the following features:

- Removes the FD stage-2 collimation section,
- Inserts sextupoles into the big bend to chromatically fix the big bend,
- Uses the big-bend sextupoles to also compensate the chromaticity of the skew-correction system and the final-focus angle-adjustment module,
- Uses the IP stage 2 sextupoles to compensate the chromaticity of the IP switch,
- Fixes the phase of modules with second-order dispersion so that the dispersion is always in the IP phase,
- Compensate the effects of interleaved sextupoles in the collimation-system modules with octupoles.

This lattice has been tracked and mapped, and has achieved the behavior described in Figure 9-38.

Power Deposition from Spoilers. Figure 9-39 shows the result of tracking particles initialized at the edge of the last spoiler in stage 1 with the angular and energy distributions of Section 9.2.2 [Cai]. We have carried out similar studies for all of the spoilers, and both edges of each. The results of this study show that about 1% of the particles survive to the first section of stage 2, and none hit the beam line after that [Drozhdin 1996].

A study of an earlier but similar beam line is shown in Figure 9-40 [Keller]. This figure shows a section of the beam line corresponding to the first half of one phase of one of the $-I$s in the collimation lattice indicating the placement of absorbers and spoilers. Note the horizontal and vertical scale. The beam pipe aperture is a small slit along the bottom of the frame. The horizontal and vertical spoiler are indicated by h/e and v, respectively, and the corresponding absorbers by H/E and V. Following the first h/e spoiler is a drift length $L = 37$ m, then a quadrupole pair with total

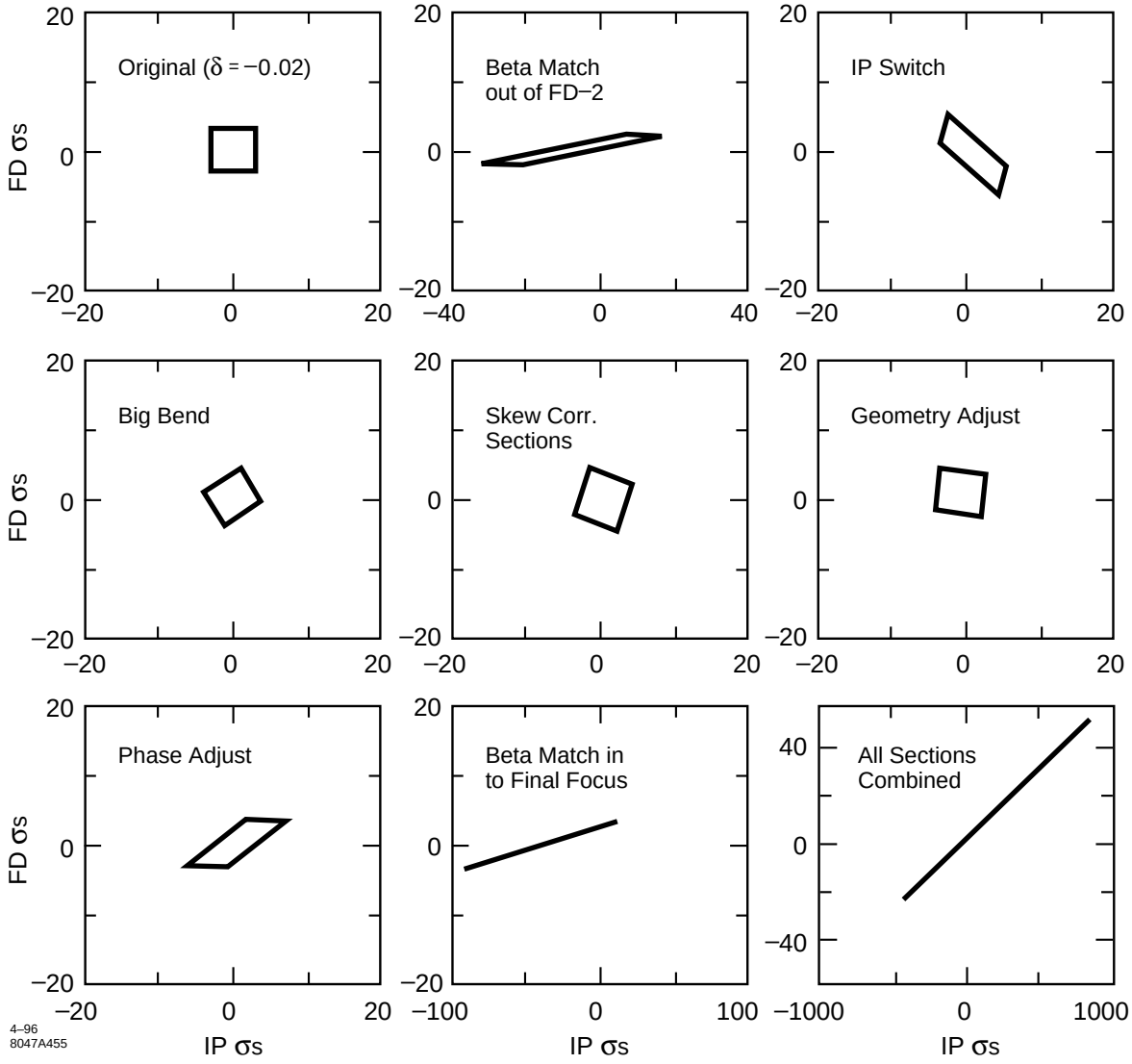


Figure 9-37. *This figure shows a sequence of maps for sections of the beam delivery system. Each picture corresponds to the image of a $3\sigma \times 3\sigma'$ piece of phase space with $\delta = -0.02$. The map is the composition of a linear map from IP to section entrance, then through the section, then a linear map back to the IP. The last picture is a map through all sections.*

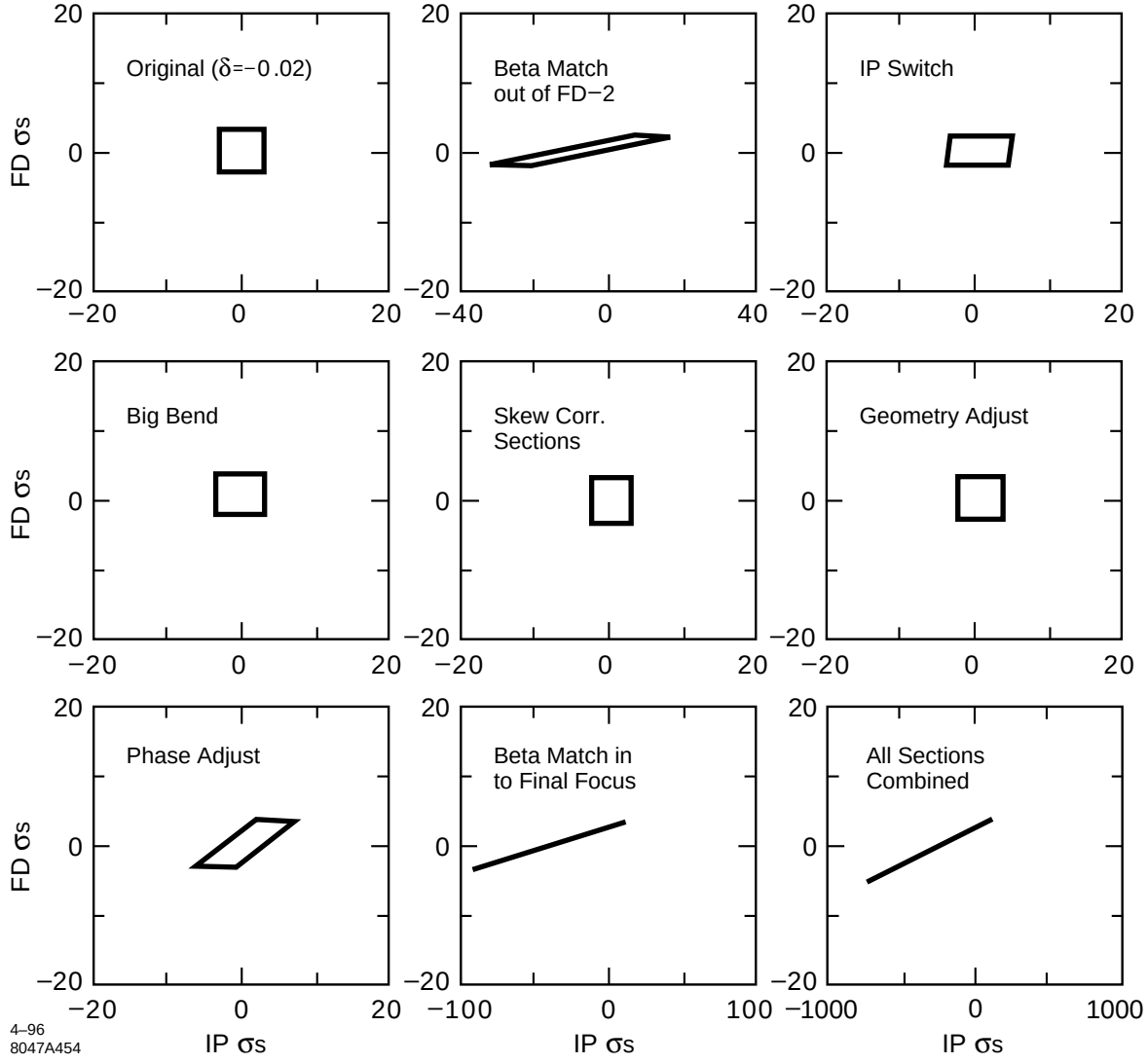


Figure 9-38. *This set of maps is the same as Figure 9-37 with the additional feature that the linear chromaticity and 2nd-order dispersion have been locally corrected. Note the dramatic change in the final map through all sections. (The scale has changed.) The FD phase is now less than 5σ and will pass through the collimator.*

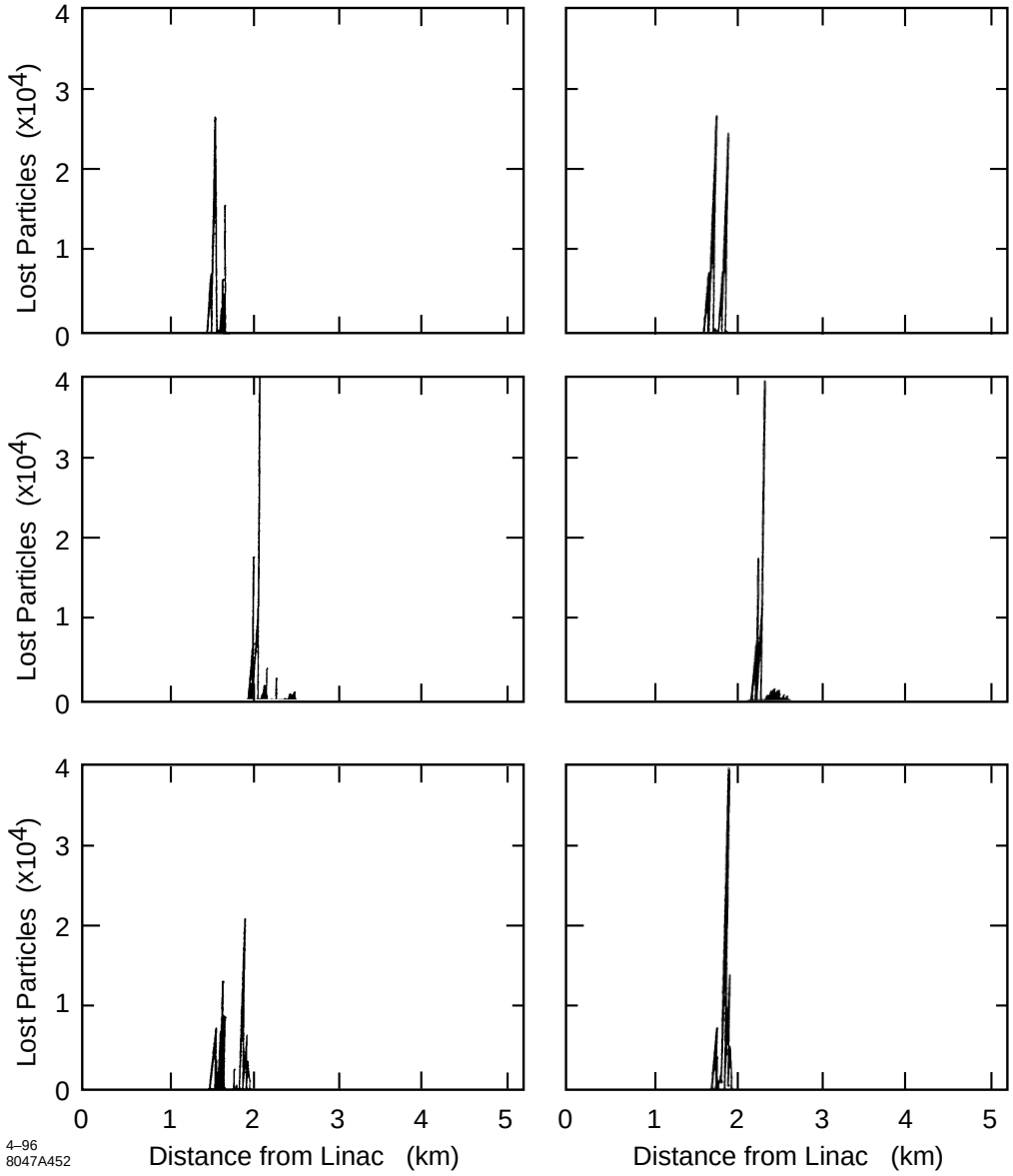


Figure 9-39. *These six pictures show the energy deposition from particles impacting the spoilers in the collimation system. The initial angular and energy distribution correspond to those shown in Figure 9-13–9-14. The six pictures are for the left or top edge of the six spoilers.*

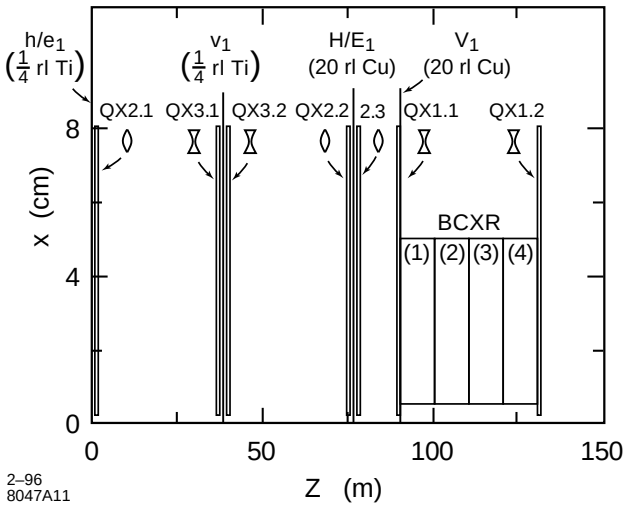


Figure 9-40. *A close-up view of a 150-m section of lattice, starting at the first horizontal spoiler. This physical arrangement was used for an EGS study of energy deposition in the 1-TeV-c.m. collimation system. The results of this study are presented in Table 9-11.*

Element	% Energy Absorbed
QX2.1	0.04
QX3.1	4.0
V1 (0.25 rl Ti)	0.1
QX3.2	1.0
QX2.2	30.1
HE1 (20 rl Cu)	36.0
QX2.3	0.4
QX1.1	0.2
V1 (20 rl Cu)	11.6
Thru V1	16.4

Table 9-11. *Results of an EGS run showing energy deposition downstream from the first spoiler.*

$k_Q = 0.53$, then a second drift of length 37 m before the absorber H/E. Using the formula of Section 9.2.2 yields $\sqrt{(R_{12}R_{34})} = 55$ m. Following the first vertical spoiler (v) is a drift length of 37 m, then a quadrupole pair with $k_Q = .037$, followed by a drift of length 11.6 m, for which $\sqrt{(R_{12}R_{34})} = 43$ m. A photon dump will be required in the region of the four bends, indicated in Figure 9-42 by the letter BCXR, and an absorber will be located before QX1.2 as well as QX1.1.

The incident beam was 50 mm from the edge of the spoiler slit h/e_1. The quadrupoles are simulated by 8-cm-radius copper cylinders with a 0.25-cm-radius bore. The four 10-m dipoles were simulated by 5-cm radius copper cylinders with a 0.5-cm radius. The percent energy absorbed is indicated in Table 9-11.

For a 500-GeV beam energy, at 120 Hz, with 0.9 10^{12} particles per bunch train, the average power in the beam is 8.4 MW. Assuming 1% is incident on a spoiler, downstream absorbers will need to accept a continuous 84 kW of power. The main absorber placed at the end of the second straight section following the spoiler will receive about

one-half of this energy, hence it will need to disperse 40 kW of energy. About 20% of the energy exiting the spoiler is in the form of photons. A large number of electrons will pass through the bend and impact absorbers in the second half of the collimation section, or in the $3\pi/2$ section between collimation sections.

Power Deposition from "Virtual" Spoilers. There have been suggestions that it would suffice to have just one stage of collimation, omitting the second IP-stage section. This seems on the face of it improbable because, according to our assumptions, there are potentially 10^{10} particles impacting the last spoiler. Of these about 2/3 are expected to impact absorbers placed immediately downstream from the spoiler. Indeed tracking shows about 79% hit absorbers in this section. Another 19% are lost in the bend that terminates the dispersion and the beta-match out of the collimation system. That leaves 4 10^8 particles traveling toward the IP with large IP-phase amplitudes. If a particle is off-energy one would expect it to be lost in the bends. But according to the energy distribution of Section 9.2.2, 35% of the particles have an energy loss less than 4%. So many particles have energies near the beam energy, and the majority of the ones that get lost because of low energy are lost at the end of the collimation system. Only 0.5% are lost between the end of the collimation system and the beginning of the final-focus system. 1% are lost in the final-focus system and the remaining 2.5% pass through the IP. Since 1% represents 10^8 particles, this is two orders of magnitude larger than the number which can be collimated in the final-focus system. The only caveat to this statement would be a detector design that could veto 100 muons arriving parallel to the beam line.

Figure 9-41 shows the loss distribution along the beam line from "virtual" spoilers placed at locations in stage 2 where there would be spoilers if this was the only collimation stage [Cai].

9.2.6 Operational Issues

Orbit Stabilization

Changes in waist and skew aberrations that arise from orbit changes within the collimation system can be monitored non-invasively in the skew correction system (SCS) downstream, and corrections introduced. This strategy will relieve the tolerances listed in Table 9-9 in the sense that the stability tolerance can be considered to be for the time between skew and waist checks in the SCS rather than the time between skew and waist scans at the IP. Also the results are intrinsically better since the two beams are not intertwined here as at the IP.

However, it will still be important to occasionally monitor the orbit within the collimation system. This requires BPMs located in these sections. BPM performance will be adversely affected by the presence of intense backgrounds originating primarily from spoilers. Hence the BPMs must be carefully protected from the spray originating at the spoilers, or from any secondary particles that might be generated by this spray.

One can take advantage of the fact that the number of bunches in the beam-train and the emittance of a single bunch can be changed rapidly (a criteria that is important for the machine protection systems). If the collimators are designed so that they can be withdrawn rapidly, then the beam could be reduced to a single small-emittance bunch and the beam trajectory can be observed without the presence of spray and offsets due to beam loss. We note that it takes an acceleration of 2 10^{-3} g, where g is the acceleration of gravity, to move an object 5 mm in 1 second, with an original and final velocity equal to zero. The concern is that the collimator alignment accuracy not suffer in this process. Collimator alignment tolerances are typically 100 μm horizontally and 40 μm vertically, so this may not be a problem.

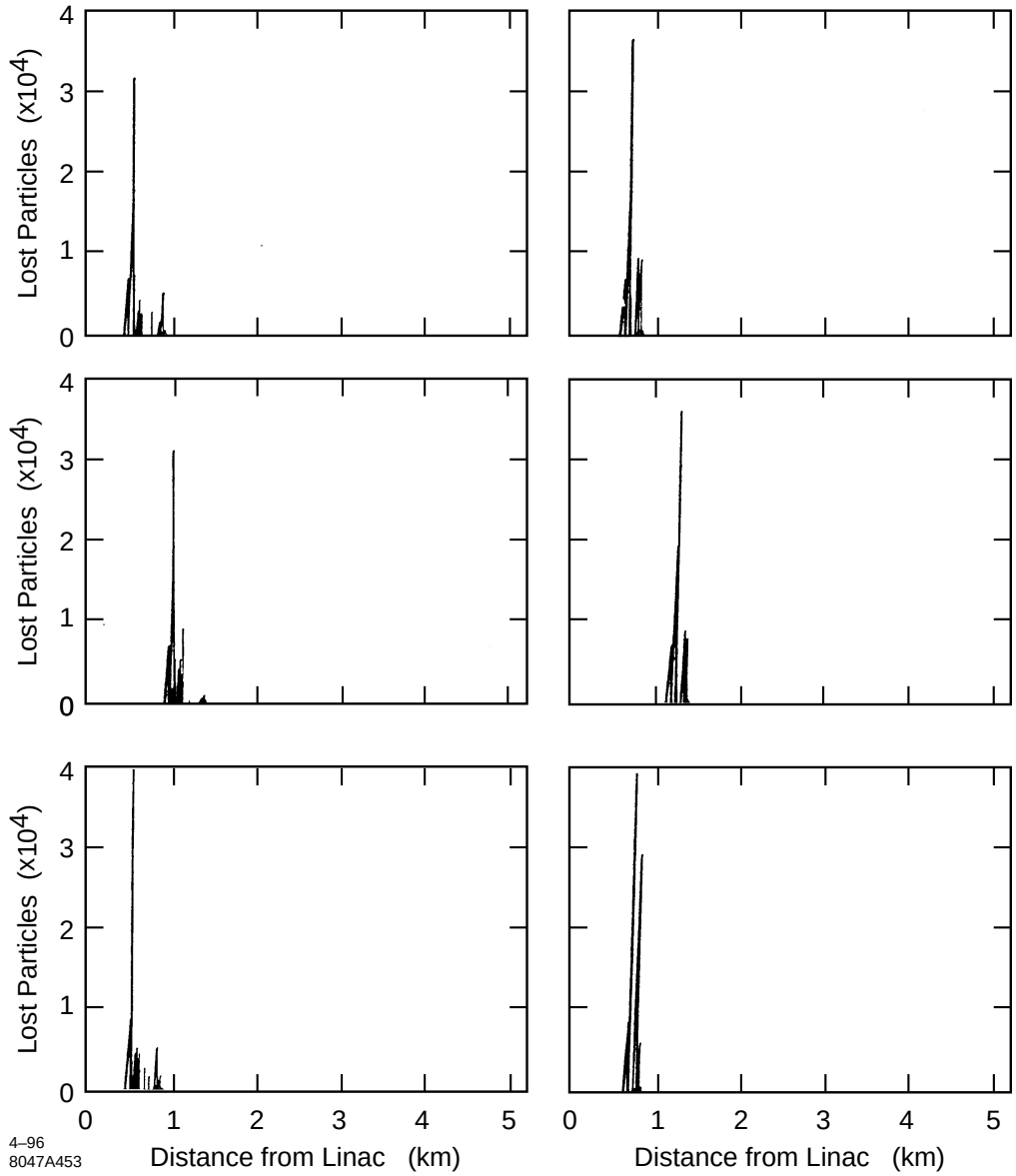


Figure 9-41. *These six pictures show the energy deposition from particles impacting "virtual" spoilers in stage 2 of the collimation system. The intent is to explore the consequences of having only 1 stage of collimation.*

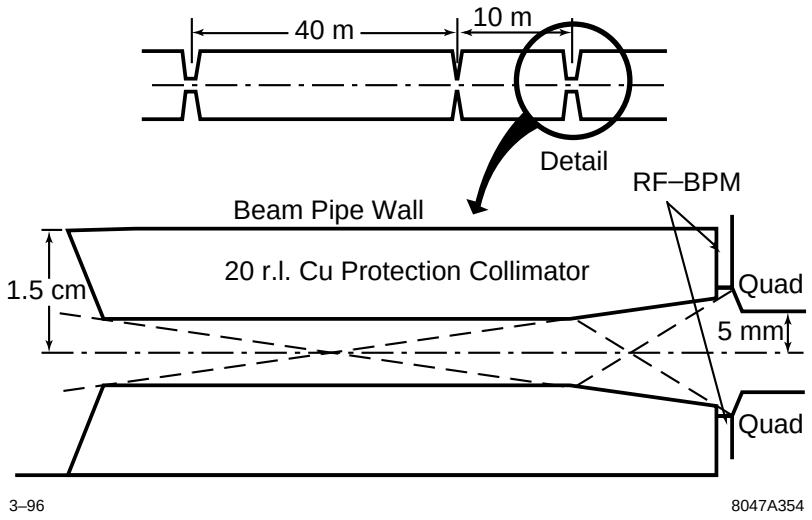


Figure 9-42. *A conceptual sketch of a possible protection collimator for BPMs of the collimation system.*

BPM Protection

Even with a single bunch it is important to protect and shield the BPMs. Because of their small length, and high resolution, and with an inner diameter that is twice the size of the quadrupole bore, the rf BPM [Shintake 1995] tested in the December 95 FFTB run should be easy to protect. One should choose the recess of the BPM and the shape of the protection collimator in such a way that particles could not reach or influence the BPM with a two-surface process (that is the beam would have to scatter or create a secondary electron at two distinct surface before reaching the BPM). The rf BPMs tested in the FFTB had an inner radius of 10 mm, and a resolution of 40 nm. Figure 9-42 shows a conceptual sketch of a protection collimator for a recessed rf BPM with an inner radius of 8 mm. This sketch is only intended to be suggestive, and does not represent a design. Research with rf BPMs has been very promising. It will be important to do studies of spray protection geometries to determine their effectiveness.

System Alignment

Magnet Alignment. It is presumed that all magnetic elements will be located on movers similar to the type used in the FFTB. These have step sizes of 1/3 μm with a total range of 1 mm. If the magnets are originally aligned to say 250 μm it is a straightforward task to align the magnets to tens of micrometers using beam-based techniques that rely on changing the current in each magnet. Such techniques were employed successfully in the FFTB.

Collimator Jaw Alignment. Collimator jaws can be aligned by changing the separation of the jaws while keeping the mid point between jaws fixed, and observing any centroid disturbance of the beam. As clarified in Section 9.2.5, the tolerance on collimator position is about 1 σ. For most collimators a 1-σ beam displacement produces about a 0.1-σ kick. In the first stages of the collimation system, with the large R_{12}s and R_{34}s to downstream stages, it is easy to detect these 0.1-σ kicks. Vertically 0.1 σ can be 4 μm, horizontally it is 30 μm. This will be detectable with the contemplated BPMs.

In the last stage of the collimation system this is more difficult. The typical beta functions in the big bend and skew correction system are 40 m, hence 0.1 σ is be 1.4 μm horizontally and 0.14 μm vertically. The former could be detected

with an FFTB type strip line BPM. The latter could be detected with the rf BPM that has been tested. In principle there is no problem detecting these orbit changes. One has the additional advantage that one can observe the trajectory through the entire remaining beam delivery system, so there is a lot of redundancy.

The alignment of the final-focus system collimator is more sensitive yet. Here the maximum R_{34} to a downstream BPM will be much smaller, say 200 m. A 0.1-σ kick at the CCY collimators has a magnitude of 0.05 nr, and the downstream displacement will be 10 nm. It is expected that this tolerance can be achieved with an rf BPM since the first efforts resulted in a 40-nm resolution. This BPM would have to be placed at an optimum location in the final telescope.

We note that if the beam has a slight banana shape as a result of upstream wakes, this collimation alignment technique will be weighted to the head of the bunch, since the head is causing the wake on the tail. Any result coming from the banana part of the bunch would tend to move the collimator in the direction of the bunch. This will have the effect of causing a kick to the tail in the direction opposite to the banana tail. In a phase space diagram this kick is perpendicular to the tail in phase space.

Tuning and Stabilization

Since the IP-phase quadrupoles are in phase with the IP they have little effect on the IP beam size, hence the tolerances of the two stages of IP-phase collimation are much looser than the tolerances on the FD-phase collimation, and so we shall concentrate on the FD phases. Furthermore, stage 2 of the FD-phase collimation is optional, since the beam has been already well collimated by previous stages, and since the β functions are smaller, the tolerances are looser than stage 1. Thus we can focus our attention on stage 1 of the FD-phase collimation. We will assume that the BPMs have the required resolution and have been protected from spray.

There are five quadrupoles in each of two sections of this stage. Each section is separated from the other by π phase advance. The first section has one horizontal spoiler and one vertical spoiler; the second section has one horizontal spoiler. If we assume that 1% of the beam is indeed intercepted at the spoilers, then since their gaps are about 2 mm, a change in the beam centroid of 20 μm can occur. Since the tolerance on the centroid position at the sextupole located next to the vertical spoiler is about 0.4 μm, we must determine the vertical centroid to a factor of 50 better than the change due to beam scraping. Thus stabilization to maintain the tolerances of Section 9.2.4 is not possible when the beams are being collimated to this extent.

As mentioned above, either the aberrations caused by a change in the collimation system orbit can be measured out and compensated in the skew correction section, or the collimators can be withdrawn after reducing the train to a single bunch. In the latter case methods similar to those described in Section 11.5.6.

Another strategy is to insure that the elements (strengths and positions) of the beam line are stable without any intervention. This is always a good course of action if possible. This strategy is greatly enhanced by the ability to non-invasively monitor the waist and skew aberration in the skew correction system (SCS) which follows the big bend. In this case the required stability times is a fraction of an hour rather than an hour. The stability tolerances on the bends are all a few times 10^{-5} which is achievable and should not cause a problem. The smallest vertical beam position tolerance at the sextupole is 0.3 μm, which could result from the motion of a quadrupole by abut 0.1 μm. Recent FFTB measurements show a beam-line element motion of about 0.1 μm per hour, which would suggest a drift of less than 10 nm in a six-minute stability time [Assman]. The FFTB measurements require further analysis and a model development to extrapolate to longer beam-line lengths. Since the FFTB line is located in a concrete block structure in the research yard at SLAC, not 20 m underground in an environment that should be thermally more stable, it appears we can achieve the required stability without intervention. (Recall that the tolerances were based on 1/2% luminosity loss for the skew and waist aberrations from the entire collimation system.)

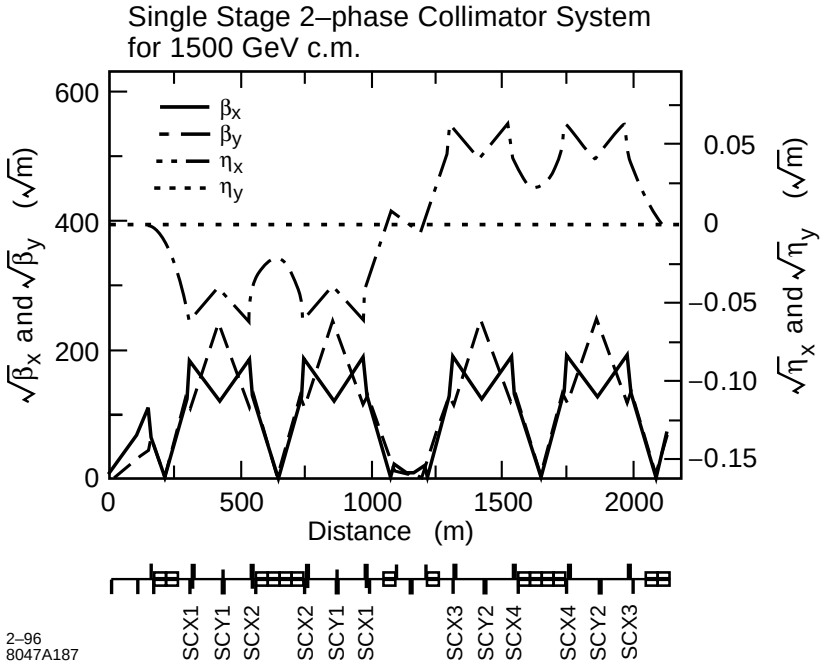


Figure 9-43. *The lattice functions for a 1.5-TeV-c.m. collimation system.*

9.2.7 Energy Scaling Laws

As we indicated in Section 9.2.4 the $\overline{n}$ are independent of energy, and since the equation for the gap is independent of energy, so is g_0. It follows that the spot sizes at the collimators are independent of energy but that β_0 must increase linearly with energy. Since energy deposition at the surface of a collimator is independent of energy, the requirements of beam-size to insure spoiler survival remain the same.

According to this scaling the beam divergence at the spoilers will scale as $\sqrt{(\varepsilon/\beta)} \sim 1/\gamma$. But the rms angle in the multiple scattering formula also varies as $1/E$, so the scattering angle to beam divergence ratios remain unchanged. Thus many of the considerations of particle counts we have presented in Section 9.2.3 above remain unchanged.

A Collimation System for 1.5-TeV c.m.

A lattice for a 1.5-TeV-c.m. system is shown in Figure 9-43. This lattice was designed so that the horizontal and vertical planes could be collimated at the same location, where the vertical β-function is a maximum. At this energy, collimating each phase one time requires a beam-line length of 2 km. As described in the text, the optimum system will have 2 IP phases and 1 FD phase of collimation. The second IP phases can have a smaller β-function. If more beam-line length is required we can extend the collimation system in to the linac tunnel.

Figure 9-44 shows the band pass and emittance growth for this lattice. Both are satisfactory.

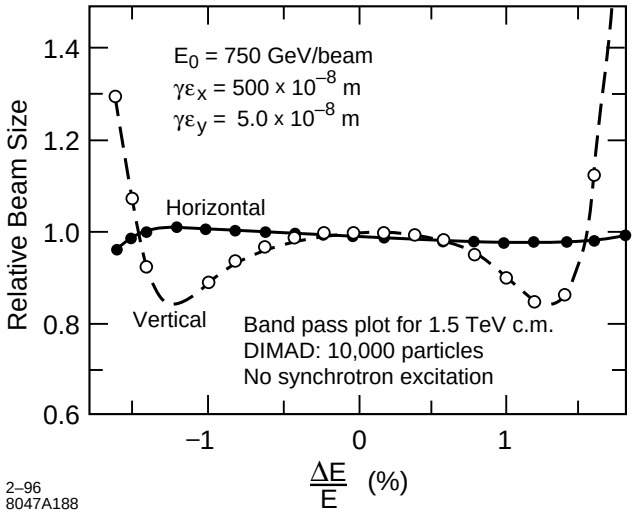


Figure 9-44. *The band pass plot for the 1.5-TeV-c.m. collimation system.*

9.2.8 Nonlinear Collimation Systems

Nonlinear collimation systems were proposed by Merminga *et al.*, [Merminga 1992] and described in several papers. The collimation system proposed for the JLC was also nonlinear and described in the JLC report [Oide]. It had a new interesting feature that the horizontal and vertical planes were collimated at the same time. Before beginning the design and analysis of the linear collimation system described here we assessed the results of these papers. We were impressed that the lengths of the nonlinear systems were not obviously shorter, and either the sextupoles were impossibly strong, or the beta functions were as large as linear system beta functions. This led us the conclusion that it would be best to begin with the more conservative linear system, pursue an optimization of that design, calculate its tolerances and assess its operational behavior, and then return to the nonlinear system for a comparison.

Now that the linear design has reached some maturity, it is appropriate to turn to the nonlinear systems, or to a combination of linear and nonlinear designs, and carry out the above intentions. Unfortunately that has not been completed.

9.2.9 Summary and Conclusions

We have described the specifications, discussed the relevant properties of materials, presented spoiler and edge-scattering distributions, clarified the relevant wake-field of tapered collimators and defined the optimum choice of collimator shape, quantified the impact of near-wall wakes for on-axis and mis-steered beam, designed lattices that implement the required functionality, calculated the position and strength tolerances of the magnetic elements, clarified the impact of ground motion, tracked the lattices to confirm their functionality, tracked spoiler distributions in the collimation system to determine power deposition, traced edge-scattered distribution in the final-focus system to determine the probability of particle impacts on the final doublet, determined the extent of tail repopulation due to gas scattering, discussed the operations problems, and addressed machine protection issues. To our knowledge all issues have been addressed with the satisfactory conclusion that it is possible to build and operate a collimation system for the proposed beam parameters from beam energies of 175 GeV to 750 GeV, that will collimate the beam at apertures

that are required by the final-focus system and final-doublet apertures. We are confident that we have an existence proof. However the systems are long and they are sensitive.

Because of the system length and its delicate nature, we feel that further research is warranted in four directions:

1. As the machine protection system (MPS) is better determined and clarified, it may evolve that the passive protection requirement we have assumed is not warranted by the magnitude of additional risk avoided. The second level of the MPS already requires careful monitoring of all system changes and limiting the range of change that can occur in fast systems. At the present moment the only fast change, which to our knowledge can occur upstream of the collimation system, is the sudden dropping of klystron families, resulting in a sudden large energy change. This is a special kind of change that must already be addressed in the linac beam diagnostic sections (chicanes). If the requirement of passive collimator protection were dropped, the collimation system would become much shorter and less delicate.

   It would be also worthwhile to better determine the worst-case tail population expected from the linac. Is it larger or smaller than the number we have assumed? This does not alter the lattice design, but does impact the power that must be handled by the absorbers.

2. Experiments should be carried out to verify the wakes of tapered collimators.

3. Nonlinear systems and combinations of linear with nonlinear systems should be studied. It is time to search for the "optimum" collimation system.

4. Improvements in the present system should be pursued, including a redesign to meet recent beam parameter changes.

5. Finally, non-conventional collimation ideas should be pursued. Though they have been discussed, it would be a worthwhile study to clarify the impracticality of all wild ideas, with the hope that one will actually survive.

9.3 Pre-Linac Collimation

The pre-linac collimation system is intended to remove the low energy, 10-GeV, contribution to the beam halo produced from the electron/positron source, damping ring, pre- accelerator and bunch compressor. The pre-linac collimation system should be able to collimate continuously 1% of the beam, 10^{10} particles, at the IP and FD phase in both planes and must withstand one full mis-steered bunch-train. Using simple energy scaling laws the post-linac collimation-system design can be scaled to produce the pre-linac design. There are some differences based on requirements for the dispersion function, and with synchrotron radiation in dipoles.

9.3.1 Pre-Linac Collimation Function:

Halo particles before the main linac are present upon extraction from the damping ring and generated in the bunch compressor. If not removed this low energy halo will be injected into the main linac and accelerated up to full energy. By placing a collimation system at the entrance to the main linac the amount of this low-energy halo accelerated to full energy will be minimized easing the load on the final collimation system, and reducing potential background in the IR region. The strategy used to achieve the needed passive protection is "spoiler followed by absorber." The spoilers will be constructed of titanium. The spoilers function is to increase the angular divergence of an accidentally mis-steered beam so that it can be intercepted by the downstream copper absorbers.

9.3.2 Pre-Linac Collimation System Requirements:

To match the post-linac specifications, the beta functions at the collimators must collimate at a depth of 6 σ_x, and 35 σ_y. The collimator jaw gap is determined by two conditions. The first is the spoiler survivability which requires that $\sqrt{(\sigma_x \sigma_y)} \leq 100\ \mu$m so that the spoiler does not exceed its tensile strength and break due to thermal expansion. The second condition on the jaw gap is determined from minimizing the wakefields induced by the collimator jaws. The equation for the optimal collimator jaw gap is independent of energy so the optimum pre-linac collimation jaw gap will be the same as the post-linac gap, approximately 1 mm.

In order for the spoilers to survive it is required that the thermal stress created by the temperature rise remain below the tensile limit of the spoiler for one mis-steered bunch train. A titanium alloy has been found to offer the best performance and it can be coated with pure titanium or titanium nitride to reduce resistive-wall wakefield effects. Following the arguments for the temperature rise and tensile strength requirements for the post-linac collimation system one can scale the results for the pre-linac collimation system. The energy deposition at the surface of a material is independent of energy above 1 GeV, hence the instantaneous heat load at 10 GeV is the same as 500 GeV. The heat load in the absorber however will simply decrease from 84 kW to 1.7 kW due to the fact that the beam has less energy to deposit. Hence the spoiler material and thickness remain the same as the 500-GeV design and the absorber design can be relaxed.

The beam divergence varies as $1/\gamma$ and the rms multiple scattering angles also vary as $1/\gamma$, thus the ratio of scattering angles to beam divergence remain constant. The energy scaling laws show that much of the previous work done on the post-linac collimation system can simply be carried over for the pre-linac collimation system. A detailed discussion of heat loads, multiple scattering, absorber properties, spoiler transmission, wakefield tolerances, etc., can be found in the post-linac Section 9.2 Energy scaling is specifically discussed in Section 9.2.7.

9.3.3 Pre-Linac Collimation Optical Design:

The optical design follows the post linac design utilizing an interleaving of the horizontal and vertical scraping to minimize the collimation section length. The beam is collimated at the IP and FD phase in both planes one time. Using the constraint $\sqrt{(\sigma_x \sigma_y)} \leq 100$ mm for spoiler survival the beta functions at the design emittances must satisfy the condition $\sqrt{(\beta_x \beta_y)} \leq 80$ m at 10 GeV. Figure 9-45 shows a 10-GeV lattice which satisfies these conditions. The spoiler and absorber gaps are about 1 mm. There are two subsections each having a set of horizontal and vertical spoilers and absorbers each separated by $-I$ transformations. The two subsections are separated by a $5\pi/2$ transform.

9.3.4 Pre-Linac Collimation System Bandwidth

The pre-linac collimation system bandwidth is shown in Figure 9-46. Though it is larger than the post-linac bandwidth, the energy spread at the beginning of the linac is quite large because of the second bunch compressor. The bandwidth should be three times this large. This problem must be addressed, either by attempting to design a collimation system with 4.5% bandwidth or locating the collimation system down-stream in the linac after the energy spread is reduced. The latter option, though seemingly attractive, has the problem that the first 600 m (to 30 GeV) of the linac is used to induce a coherent energy spread to facilitate BNS damping in the linac. Hence at 30 GeV the energy spread is still 1.5%, and it would be necessary to place the collimation system at 90 GeV. This is nine times the energy of the existing design, so the system would have to be redesigned for that energy. Since pre-linac collimation is optional we have not investigated this subject further.

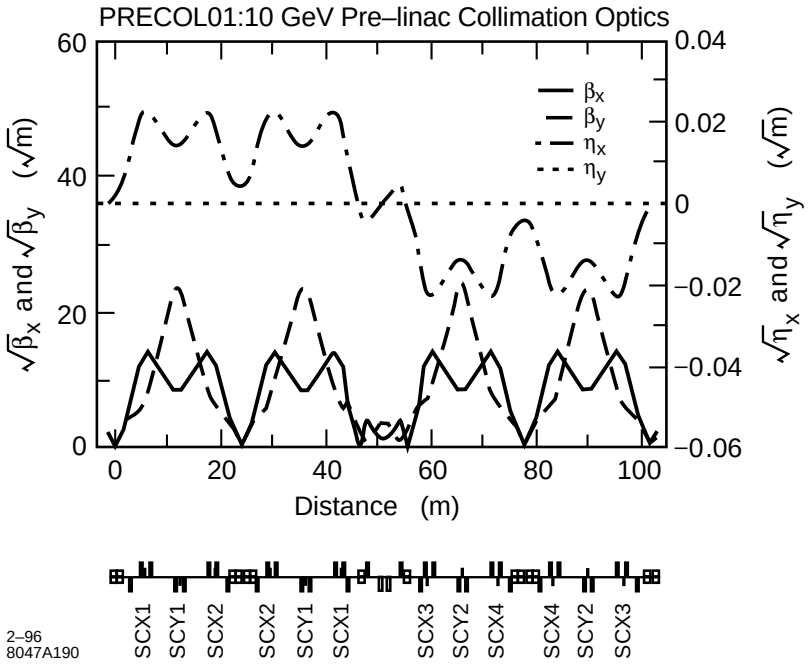


Figure 9-45. *The pre-linac collimation system lattice.*

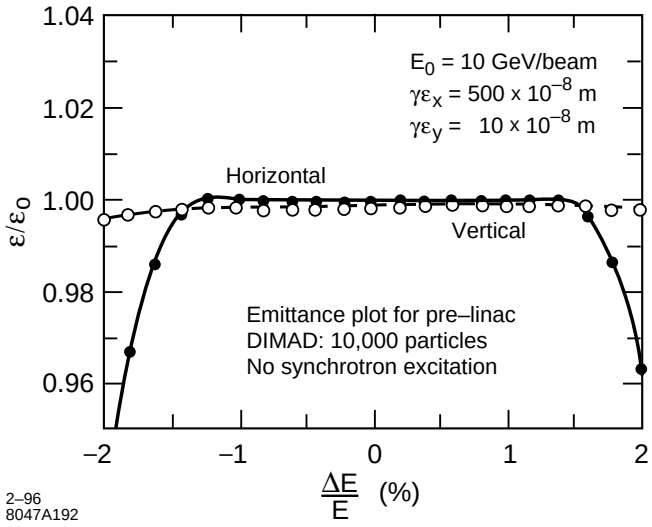


Figure 9-46. *The pre-linac emittance growth.*

9.3.5 Summary

We haved presented a design for an optional 10-GeV pre-linac collimation system. This system is a scaled-down version of the 500-GeV post-linac collimation system. The beta functions were scaled so that the spoiler survival criteria of $\sqrt{(}\sigma_x\sigma_y) \leq 100\,\mu$m, a 1% beam halo collimation, and a minimal wakefield condition are satisfied. The system uses the same quadrupole magnet design as the post-linac system and its length is 100 m. The design presented has a bandwidth of 1.5% which is too small to be located immediately after the bunch compressor. If pre-linac collimation is deemed necessary this system will need to be improved upon.

9.4 Bunch Length Collimation

9.4.1 Introduction

In order to achieve high luminosity, the NLC requires very short bunches. The three primary considerations that determine the bunch length are: the bunch length should not be much larger than β_y^*, the emittance growth due to the transverse wakefields increases with the bunch length, and the energy spread due to the longitudinal wakefield decreases with increasing bunch length. For the NLC parameters, the optimal bunch length is between 100 and 200 μm.

When generating these bunches in the bunch compressors, one typically populates long bunch length tails. These tails will be deflected to large transverse amplitudes by the transverse wakefields. In addition, the long longitudinal tails will be converted into energy spread as they go through the main linac. In the post-linac collimation system and the final-focus system, the energy tails will generate transverse halo due to chromatic effects. Both sources of transverse tails will create unacceptable backgrounds in the detector if not collimated. Therefore, the longitudinal tails have to be collimated before entering the final focus. It is better, and presumably easier, to collimate them before the main linac.

For these reasons we decided to design a bunch length collimation system for the NLC. The system should be located before or at the beginning of the NLC main linac, and remove particles beyond $3\sigma_z$. The usual method to cut longitudinal tails is to convert them to transverse tails and remove them with a collimator. The ideal case is when the bunch length has a monotonic energy distribution. Then the energy spread can be transferred into horizontal displacement with dispersion.

9.4.2 Collimator location

Figure 9-47 is the layout of the damping ring to the final focus of the NLC, showing the damping ring, two steps of bunch compression and the X-band main linac.

The longitudinal phase $\Phi = n\pi$ is the necessary phase to place the collimator. Hence we may choose the first collimation at the exit of the damping ring. Here, the energy spread is small (0.1%), and for $3\sigma_z$ bunch length collimation, it will need to cut beyond $+/-0.3\%\Delta E/E$. Assuming the collimator half gap equals 1.5 mm and $\sigma_e/E = 0.1\%$, then a dispersion $\eta = 0.5$ m is needed. Since there is already a spin rotation system at this location, the 0.5-m dispersion is available.

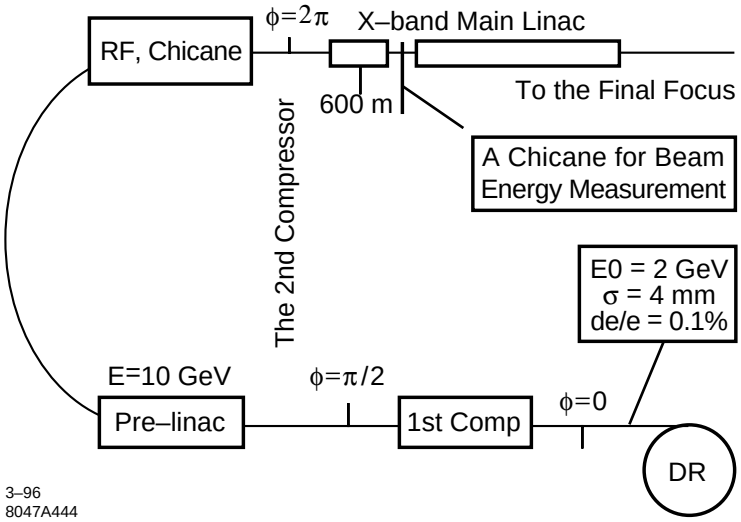


Figure 9-47. *The layout of the NLC.*

Collimation location	Beam loss %
1st collimation only	4%
2nd collimation only	2%
1st and 2nd together	5%

Table 9-12. *Particles loss percentage at the collimation systems.*

A second bunch-length collimation could be put after the second bunch compressor. Due to the space limitation, we choose to put this collimator after the first 600 m of acceleration in the X-band main linac, where there is a chicane for beam energy measurement.

9.4.3 Particles Loss at Bunch Length Collimators

We put the two bunch length collimation systems into the present two-step compression system and the main-linac lattice design. Table 9-12 presents the expected particle loss at the collimators with a cut at $2\sigma_e$ [Raubenheimer 1994, Zimmermann].

The following figures show the bunch shape at the end of the main linac as calculated by a modification of the LITRACK program [Bane].

9.4.4 Collimation in the Linac

It is important to have the beam energy be longitudinally monotonic, so that the bunch length collimation can be made by energy tail cuts. To meet the above requirement, we can put the beam behind the rf crest, so that only 10% of the

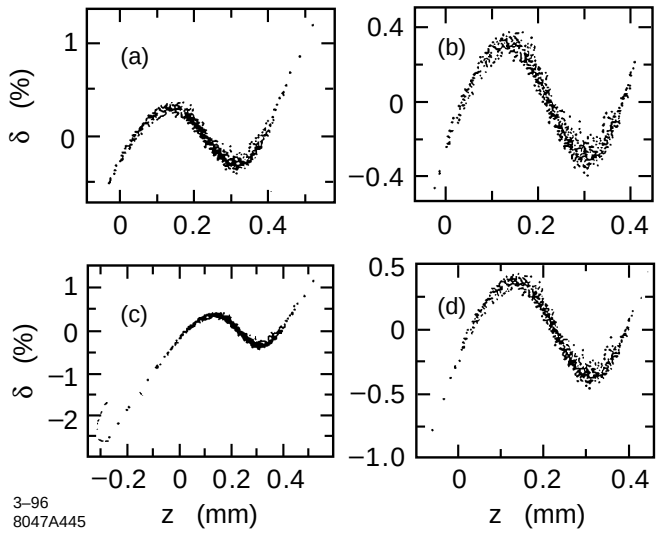


Figure 9-48. *(a) The bunch shape without collimation. (b) The bunch shape with the first collimation only. (c) The bunch shape with the second collimation only. (d) The bunch shape with both bunch length collimation systems.*

	NLC-I	NLC-II	NLC-III
Rf phase	16°	16°	16°
Energy gain	33 MV/m	55 MV/m	70 MV/m
In 600 m	20 GeV	33 GeV	42 GeV
# of particles	7 10^9	11 10^9	14 10^9
Energy spread	1.0%	1.2%	1.28%
# of particles	1.54 10^8	3.30 10^8	7.84 10^8
(%) Cut	(2.2%)	(3.0%)	(5.6%)

Table 9-13. *Some parameters for the three NLC design energies.*

particles are beyond $4\sigma_z$ and have energy deviation greater than $4\sigma_e$. For a collimator half gap equal 750 μm, and $\sigma_e/E = 1\%$, a dispersion of $\eta = 2.5$ cm implements a $+/-3\sigma_e$ energy tail cut.

At the beginning of the main linac, the beam bunch length is 100 μm. The energy distribution may be calculated with the formula:

$$E(z) = E_0(z) + V_{\rm rf}\cos(\Phi_0 + kz) - \int W_L(z'-z)\rho(z')dz' \quad . \tag{9.86}$$

Here $E_0(z)$, is the initial energy distribution at the beginning of the main linac, with $< E_0(z) >= 10$ GeV. $V_{\rm rf}$ is the main linac rf peak voltage. Φ_0 is the center particles' accelerating phase, and k is the wave number of the rf wave form. $W_L(z'-z)$ is the main X-band linac longitudinal wakefield, which has been calculated by K. Bane. $\rho(z')$ is longitudinal distribution, which we will assume to be Gaussian. The minus sign indicates beam energy loss.

For the three NLC designs, when $\Phi = 16°$ behind the crest of the rf phase is chosen, the beam energy spreads are about 1%. If we set the collimator to cut $+/-2\sigma_e$ (*i.e.*, $+/-2\%$), less than 10% particles will be cut. Table 9-13 lists some parameters and percentages of particles that are lost at the collimators. Figure 9-49 shows the energy distribution along the bunch.

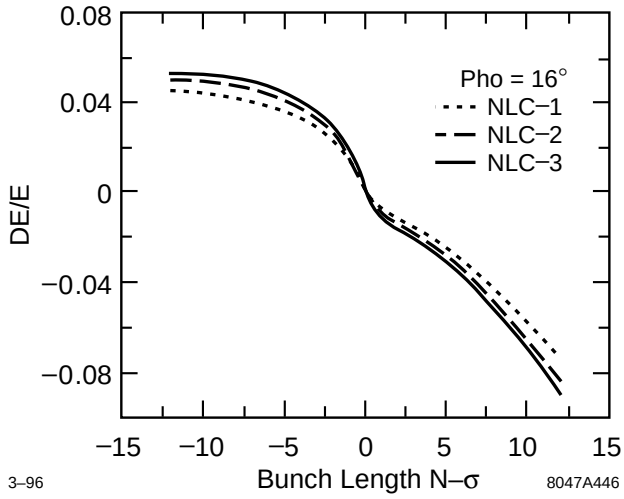


Figure 9-49. *Energy distributions for the three NLC design phases.*

9.4.5 Conclusion and Discussion

It is possible to have two bunch-length collimation systems in the NLC design lattice. The bunch has a nearly monotonic energy distribution when we choose the rf phase to be 16° behind the rf crest for all three NLC design phases. The energy spreads are about 1%. If we decide to collimate $+/-2\sigma_e$ energy tails, there will be less than 10% of particles cut.

Before determining the utility of the collimators, a number of additional questions remain to be answered. In particular, we need to study the longitudinal and transverse wakefields induced by the collimators and related tolerances.

Our results show that the collimator at the exit of the damping ring is relatively straightforward and effective. Because of a non-linear $\delta - z$ relationship, and therefore a nonlinear $x - z$ dependence at the collimator, the second collimation is more difficult and less beneficial.

References

[Assman] R. Assman and C. Montag have studied quadrupole motion in the FFTB using the DESY wire system. Data analysis is in process.

[Bane] K. Bane wrote the program LITRACK (1994). Undocumented.

[Brown 1979] K. Brown and R. Servranckx, "Chromatic Corrections for Large Storage Rings", SLAC–PUB–2270, in *IEEE Trans. Nucl. Sci.* **26**, 3598 (1979).

[Cai] Y. Cai performed the tracking studies for this chapter using a C++ mapping and tracking code he has written called LEGO.

[Chao 1992] A. Chao, "Coherent instabilities of relativistic bunched beam", SLAC–PUB–2946 (1982).

[DeStaebler-Walz] Spoilers followed by absorbers have been in use for many years at the Stanford linac.

[DeStaebler 1994] H. De Staebler first pointed this out to us. Private communication.

[Drozhdin 1996] A. Drozhdin and N. Mokhov, "NLC-1000 Beam Collimation System Simulation", Jan. 18, 1996. The authors have prepared distributions similar to those presented in the text generated by Y. Cai. We thank them for supplying us with a preprint of this study they did of the NLC collimation system.

[FFTB BPMs] The FFTB strip line BPMs have a resolution of 1 mm. The stability of the zero point has not yet been determined. According to S. Smith the one-hour stability could conceivably be a factor of ten better than the resolution. A proposal has been submitted to study the zero point drift of these BPMs.

[Helm] All of the collimation lattices have been prepared by R. Helm. Lattice angle control was studied by Y. Nosochkov.

[Irwin 1991] J. Irwin, N. Merminga, R. Helm, and R. Ruth, "Optimizing a Nonlinear Collimation System for Future Linear Colliders", SLAC–PUB–5507, in *Proc. 1991 Part. Acc. Conf.* (1991).

[Keller] L. Keller provided this tracking study.

[Merminga 1992] N. Merminga, R. Helm, J. Irwin, and R. Ruth, *Particle Acclerators* (1992).

[Nelson 1993] R. Nelson has provided the EGS simulations described in this chapter. See The EGS Code System, SLAC–PUB–265 (1985).

[Oide] JLC collimation system, Section 4.8.6, "JLC-I", KEK Report 92–16 (1992).

[Piwinski] A. Piwinski, "Wake Fields and Ohmic Losses in Flat Vacuum Chambers", DESY–HERA–92–04 (1992).

[Rossi 1952] B. Rossi, *High-Energy Particles*, (Prentice-Hall 1952).

[Raubenheimer 1995] T. Raubenheimer has carried out these first estimates of expected tail populations. Further detailed studies are recommended. Private communication (1995).

[Raubenheimer 1994] T.O. Raubenheimer, "Bunch Compressor Parameters", NLC–Note–2 (1994).

[SLC] Recent experiments with collimator wakefield kicks are indicating that the wake kick is about three times larger than predicted by theoretical formulas. It appears to have the gap behavior appropriate to a resistive-wall wake, but a geometric source is not ruled out.

[Shintake 1995] Rf BPMs with a resolution of 40 nm were demonstrated in the Dec. 95 FFTB run.

[Smith 1995] Private communication. Steve Smith designed and built the BPMs for the FFTB.

[Stupakov 1995] "Geometric Wake of a Smooth Taper", SLAC–PUB–7086 (Dec. 1995).

[Stupakov 1996] This limit was computed by G. Stupakov. Private communication.

[von Holtey] G. von Holtey suggested the use of curved collimators for this reason. G. von Holtey, "LEP Main Ring Collimators", CERN/LEP–BI/87–03.

[Walz 1973] D. Walz, D. Busick, T. Constant, K. Crook, D. Fryberger, G. Gilbert, J. Jasberg, L. Keller, J. Murray, E. Seppi, and R. Vetterlein, "Tests and Description of Beam Containment Devices and Instrumentation–A New Dimension in Safety Problems", SLAC–PUB–1223 (1973).

[Walz 1991] D. Walz has provided the information on properties of materials used in this chapter.

[Walz 1992] "NLC Final-focus Collimation and Dumping", Working Group Talk, Final Focus and Interaction Region Workshop, (1992).

[Warnock] According to Warnock the resistive-wall wake is not given by this integral in perturbation theory (for small disturbances of the beam wall). Private communication. However, we know of no better approximation to use for these deep tapers.

[Yan] Y. Yan has created and analyzed the maps used to analyze the chromatic behavior of the beam delivery beam line.

[Yokoya 1988] "Impedance of Slowly Tapered Structures", CERN SL/90-88 (AP).

[Yokoya 1991] K. Yokoya and V. Telnov prepared estimates of this at the LC91 workshop in Protvino, USSR. The argument here follows a transparency prepared by K. Yokoya.

[Yokoya 1995] K. Yokoya has pointed out the importance of the on-axis wakes for parallel plate collimators. We thank K. Bane for communicating these results.

[Zimmermann] F. Zimmermann, "Longitudinal Single-bunch Dynamics and Synchrotron Radiation Effects in the Bunch Compressor", NLC–Note 3 (1994).

Contributors

- K. Bane
- Y. Cai
- A. Drozhdin
- R. Helm
- L. Keller
- J. Irwin
- R. Messner
- R. Nelson
- T. Raubenheimer
- G. Stupakov
- F. Tian
- D. Walz
- M. Woodley
- Y. Yan
- F. Zimmermann

10

IP switch and big bend

Contents

10.1 Introduction

The IP switch follows the main linac and collimation section and allows slow switching between multiple IPs. The big bend provides muon protection and IP separation. It also generates the IP crossing angle which facilitates extraction of the spent beams. Figure 10-1 shows a schematic layout (drawn to scale) of the IP switch, the two big bend sections, and the skew correction and diagnostic sections (Chapter 11). These sections follow the collimation section (Chapter 9) and a 100-m-long emittance diagnostic section. An NLC design with two IPs will require two IP switches and four big bends.

10.2 The IP Switch

The purpose of the IP switch is to provide capability for switching the beam between two alternate final focus beam lines. The IP switch should provide enough separation so that most of the major transport elements are not shared by the two beam lines. Rapid switching of the beam between alternate transport lines is not necessary. It is probably possible to make the switch in a period of less than one hour. Emittance growth from aberrations and synchrotron excitation should be negligible (*i.e.,* a few percent).

10.2.1 Optics Design

The IP switch bends the beam a total of 1.5 mr. Figure 10-2 shows the IP switch optics. The QS quadrupole is horizontally movable in order to switch between IPs. It is displaced by ± 3.25 cm (± 2.6 cm) for the 500-GeV/beam (750-GeV/beam) configuration.

The upgrade to 750 GeV/beam is accomplished by adding ten 3-m-long dipoles inboard of the existing dipoles (Figure 10-3a and 10-3b). Some quadrupole strength changes are also necessary. Figure 10-3 shows the beam line offsets for the 500-GeV/beam (Figure 10-3a) and the 750-GeV/beam (Figure 10-3b) cases. The plots start from the end of the collimation section and continue through the big bend matching section. The BS dipoles should be thin C-magnet types to fit in the 11.7-cm (7.8-cm) center-to-center separation at the face of the first dipole just downstream of the QS quadrupole. The first five BS dipoles (first ten for the 750 GeV/beam configuration) provide the separation, while the QS quadrupole and the next five (ten) BS dipoles make the system achromatic. Beam line elements through QS are common to both beam lines. Three configuration modes are possible. The first is at 500 GeV/beam with 10 dipole magnets installed (Figure 10-3a), the second is at 500 GeV/beam with 20 dipole magnets installed (Figure 10-3b), and the third is 750 GeV/beam with 20 dipole magnets installed (also Figure 10-3b). Tables 10-1 and 10-2 list the magnets for both energies. The match into the IP switch and also into the big bend has been accomplished with two sets of four quadrupoles (QM1,…8) of modest design. The center-to-center separation at the face of QM5 is 15.7 cm, requiring a special thin quadrupole design at least for QM5 and QM6.

10.2.2 Chromatic Emittance Dilution

Tracking studies using the tracking code TURTLE [Carey 1982] have been made for beams with Gaussian energy distributions having rms of 0.2, 0.4, and 0.6%. The results for the entire beam line (from end of collimation section

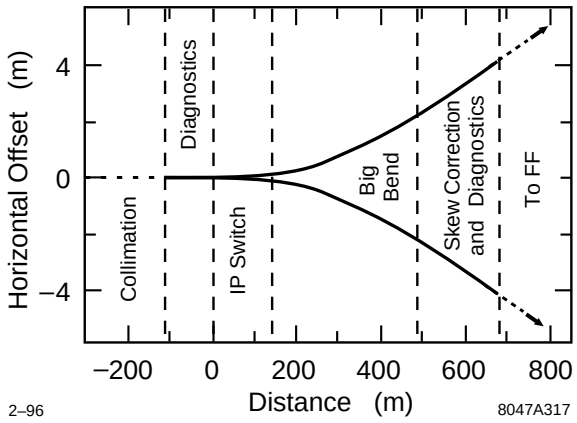


Figure 10-1. *IP switch/big bend layout (to scale).*

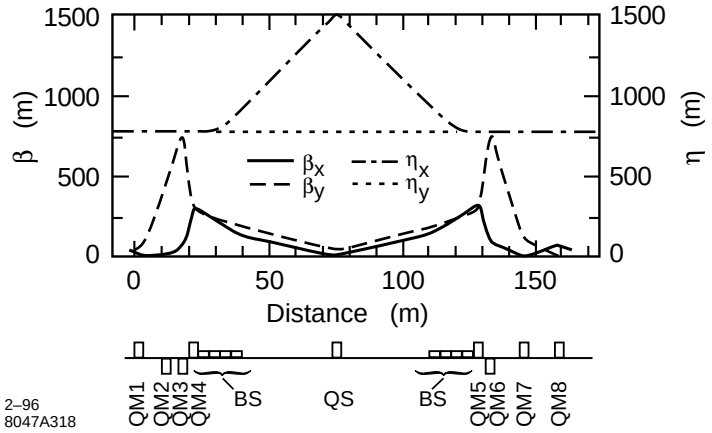


Figure 10-2. *IP Switch Optics (500 GeV/beam).*

Name	Number	Length (m)	Radius (mm)	Pole-Tip Field (kg)
QM1	1	2.5	6	+5.85 (+8.78)
QM2	1	2.5	6	−0.33 (−0.38)
QM3	1	2.5	6	−6.57 (−10.0)
QM4	1	2.5	6	+6.62 (+9.97)
QS	1	2.0	6	+2.35 (+4.43)
QM5	1	2.5	6	+6.08 (+9.15)
QM6	1	2.5	6	−6.60 (−9.97)
QM7	1	2.5	6	+3.01 (+4.62)
QM8	1	2.5	6	+6.66 (+10.0)

Table 10-1. *IP-switch quadrupole magnets for 500 GeV/beam (750 GeV/beam). Fields for the 20-dipole, 500-GeV/beam mode (not shown) are simply scaled from the 750-GeVcase.*

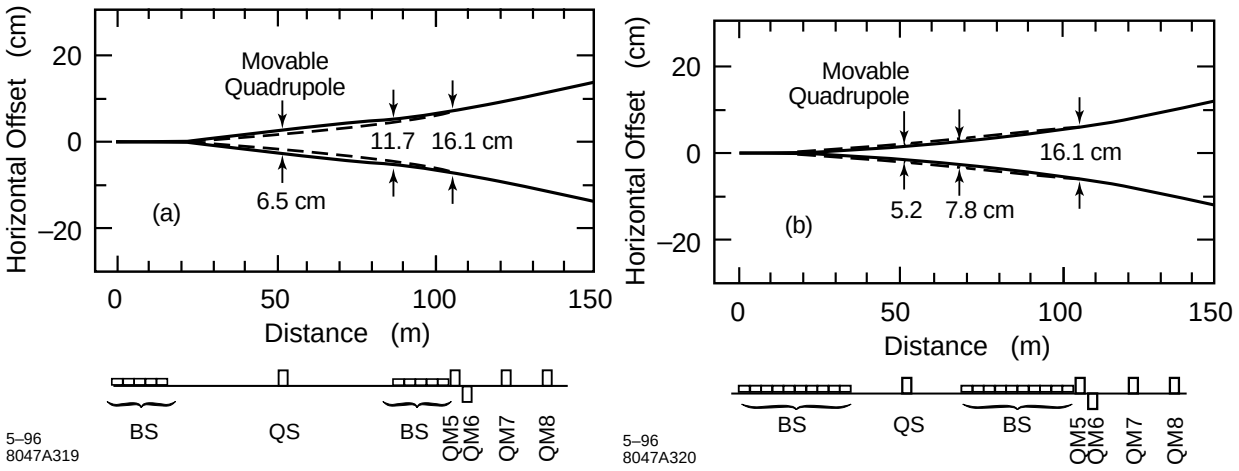


Figure 10-3. *(a) IP switch beam line (500 GeV/beam). Dashed line is nominal beam line at 750 GeV. (b) IP switch beam line at 750 GeV/beam after adding 10 more 3-m dipoles inboard of existing dipoles. Dashed line is nominal beam line at 500 GeV/beam.*

Beam Energy (GeV)	Name	Number	Length (m)	Half Gap (mm)	Field (kGauss)
500	BS	10	3.0	6	0.834
500	BS	20	3.0	6	0.417
750	BS	20	3.0	6	0.625

Table 10-2. *IP switch dipole magnets for different energy modes.*

to beginning of final focus) are tabulated in Table 10-9. The chromatic contribution to emittance increase for the IP switch alone at 0.3% rms energy spread is <1% in each plane.

10.2.3 Synchrotron Radiation

The horizontal emittance dilution, energy spread increase, and energy loss due to synchrotron radiation (SR) through the IP switch are summarized in Table 10-3. The 7.3% emittance increase quoted in Table 10-3 is with respect to the main damping ring extracted emittance. For a more realistic end-of-linac emittance of $\gamma\epsilon_{x0} = 5 \times 10^{-6}$ m the increase is 4.4%.

10.2.4 Tolerances

The single-element tolerances for the IP switch magnets are listed in Tables 10-4 (dipoles) and 10-5 (quadrupoles). Each tolerance represents a 2% luminosity loss for that single element's effect on one beam. The effects of these errors generally increase the IP beam size except in the case of dipole field regulation and quadrupole transverse vibration which continuously steer the beams out of collision. In this case, since the exact betatron phase to the IP is

Beam Energy (GeV)	Number of Dipoles	$\Delta\epsilon_{x\mathrm{SR}}/\epsilon_{x0}$ (%)	$\sigma_{\delta\mathrm{SR}}$ (10^{-4})	Energy loss (MeV)
500	10	3.3	7	66
500	20	0.7	3	33
750	20	7.3	10	167

Table 10-3. *Horizontal emittance dilution, energy spread increase and energy loss due to synchrotron radiation through the IP switch ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m).*

NAME	Quantity	roll (mr)	$\Delta B/B_0$ (%)	b_1/b_0 (%)	b_2/b_0 (%)
BS	10	10	0.033	3.4	300

Table 10-4. *IP switch dipole magnet single element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 4 mm. The sextupole field harmonics for the dipoles are extremely loose ($\sim 300\%$ at $r = 4$ mm).*

NAME	Quantity	roll (mr)	Δx offset (μm)	Δy offset (μm)	Δx_{rms} vibrate (μm)	Δy_{rms} vibrate (μm)	$\Delta B/B_0$ (%)	b_2/b_1 (%)
QM1	1	4.1	360	21	1.1	0.063	4.5	1650
QM2	1	27	6800	130	20	0.400	11	4000
QM3	1	0.47	150	5.2	0.45	0.016	0.33	55
QM4	1	0.40	83	7.9	0.25	0.024	0.56	52
QS	1	1.2	1680	76	5.0	0.230	5.1	150
QM5	1	0.43	90	8.7	0.27	0.026	0.61	58
QM6	1	0.46	150	5.2	0.44	0.016	0.33	55
QM7	1	12	1500	30	4.5	0.090	4.9	3750
QM8	1	7.2	190	61	0.58	0.180	4.5	1000

Table 10-5. *IP switch quadrupole magnet single element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Sextupole field harmonics (b_2/b_1) are evaluated at a radius of 4 mm and are very loose ($> 50\%$).*

not calculated, phase averaging is applied. The tolerances given in the tables have not yet been distributed out into a weighted tolerance budget; the numbers are for reference. In fact, given multiple errors over multiple elements, these tolerances are much too loose. However, since tuning considerations have not been folded in, most static, non-steering errors may also be corrected over some reasonable range.

10.2.5 Diagnostics and Correctors

Beam position monitors (BPM) will be required, probably one horizontal and one vertical BPM per $\pi/2$ of betatron phase. A minimal number of dipole orbit correctors should be used; the optimal locations for these devices have yet

to be established. The diagnostic section at the end of the big bend will be used to determine the quality of the match into the big bend.

The horizontal dispersion of 32 mm (26 mm at 750 GeV/beam) at QS provides an excellent location for the measurement of beam energy and energy spread, using a BPM and a profile measurement device such as a wire scanner. At the entrance to QS, the dispersive horizontal spot size for a beam with 0.3% rms energy spread is 96 μm (78 μm), while the betatron spot size ($\gamma\epsilon_x = 3 \times 10^{-6}$ m) at this location is only 5.2 μm (4.2 μm), allowing a good energy spread measurement. In addition, placement of a 1-μm-resolution BPM at this location will provide a relative energy measurement resolution of $\sim 3 \times 10^{-5}$.

10.2.6 Beam Correction Issues

Beam-based techniques will be used to verify transverse alignment during commissioning of the beam line, but will probably be infrequently necessary thereafter. Orbit correction algorithms remain to be studied, but a simple point-to-point scheme will probably be sufficient. Feedback stabilization of beam position at the entrance to the big bend will probably be desirable, depending on the stability of the incoming beam.

10.2.7 Other Issues

Some issues remain to be considered, including:

- The need for machine protection collimators and their locations.
- Vacuum and pumping requirements.

10.3 The Big Bend

The big bend is a low-angle arc after the main linac which provides detector muon protection [Keller 1993], an IP crossing angle to facilitate extraction of the spent beams, and allows switching between multiple IPs. The total bend angle (including 1.5-mr IP switch angle) is 10 mr (20-mr IP crossing angle) which provides ∼40-m spatial separation between the two IPs (∼700-m transport to an ∼1600-m-long final focus). At 500–750 GeV/beam, the horizontal emittance growth due to SR sets lower limits on the system design length. The following Section describes an optimized optical design of this big bend section for 500 GeV/beam and 750-GeV/beam electrons (or positrons).

10.3.1 Optical Design

For electrons, the emittance growth due to SR is calculated using [Helm 1973, Raubenheimer 1993]:

$$\Delta\gamma\epsilon_x \approx \left(4 \times 10^{-8}\,\mathrm{m^2\,GeV^{-6}}\right) \cdot E^6 \sum_i \frac{L_i \langle H \rangle_i}{|\rho_i|^3} \quad , \tag{10.1}$$

where the summation is over bending magnets, L is the magnet length, ρ is the bend radius of the magnet, E is the beam energy, and $< H >$ is the mean of the usual "curly-H" function.

$$\langle H \rangle = \frac{1}{L} \int_0^L \frac{\eta^2 + (\eta\prime\beta + \eta\alpha)^2}{\beta} dz \tag{10.2}$$

This integral has been solved [Helm 1973] for a magnet with bending and focusing. The mathematical result is lengthy and is not reproduced here. To find the optimal parameters for a string of FODO cells, this result is used with Equation (10.1) in a convenient computer program to calculate the SR emittance growth with maximum quadrupole pole-tip fields of 10 kg at 750 GeV/beam with a 6-mm radius. Preliminary resistive wall calculations [Bane 1995] indicate that this radius might be decreased to 3 mm, shortening the system length by ∼100 m. However, it is thought that this possibility may not provide an adequate safety factor.

The number of FODO cells and the phase advance per cell are varied to find the minimum total length for a ∼2% horizontal SR emittance growth ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m). Both separated function and combined function magnet systems were explored. The parameters reached represent a compromise between theoretical optimal values and realistic constraints on magnet lengths and reasonable phase advance per cell.

The phase advances per cell chosen ($\Psi_x = 108°, \Psi_y = 90°$) do not represent the precise optimum ($\Psi_x = 135°, \Psi_y < 72°$). The values are biased towards a more reasonable design considering beam position monitor sampling, chromaticity, a potential sextupole resonance and magnet alignment tolerances. The effect on the total length of this slight bias is small (<10%). Splitting the horizontal and vertical tunes in the separated function lattice improves the bending magnet density because it allows the D-quadrupoles to be shorter than the F-quadrupoles for a constant pole-tip field. However, in a combined function design, this split does not improve the SR emittance dilution. The tune split may also desensitize the beam to a potential ion instability. The choice between combined function (CF) or separated function (SF) lattice reduces to a few points listed below.

Combined Function: The CF lattice is more space-efficient and only one type of magnet needs to be built. The main disadvantage is that the focusing strength is not tunable without changing the bending strength. The magnets are long (∼6 m) since they bend and focus, but the net length is still shorter than the SF design (∼30%).

Separated Function: In its favor, the SF lattice may be more easily tuned since the focusing strength will be independent of the bend strength. For example, the phase advance per cell may be changed to provide a trombone tuner between the collimation phase and the IP. Also beam-based alignment techniques can be applied. However, there are three types of magnets to build in this scheme and the overall length needs to be longer than the CF design (∼30%).

Due to its tunability and beam-based alignment potential, we have chosen the SF design. The dispersion and beta functions of the 500-GeV/beam SF design are shown in Figure 10-4. The final design uses 15 FODO cells with four dipoles per cell. Table 10-6 lists the FODO parameters for the big bend SF design.

A 76-cm quadrupole-to-dipole space has been maintained for BPM and ion pump placement and a 50-cm dipole-to-dipole space is held so that the dipole and quadrupole magnet lengths are no more than 3 m and 2.5 m, respectively. A "missing-magnet" dispersion matcher/suppressor is used at the entrance and exit of the FODO string. The matcher/suppressor magnets are identical in size and strength to those in the FODO section. Only their longitudinal position has been adjusted to obtain the periodic dispersion function in the FODO section. In this way, all 64 dipoles may be powered in series with one power supply, the 17 defocusing quadrupoles powered with a second supply, and the 16 focusing quadrupoles on a third. This allows tuning of the phase advance per cell independently in each plane. Table 10-7 lists the dipole magnet parameters and Table 10-8 lists the quadrupole magnet parameters for the big bend as shown in Figure 10-4. The 3-m-long dipole magnet design may also be used in the IP switch as long as it is a C-magnet design with dimensions which meet the requirements described in Section 10.2.1. The 2.5-m-long QF design may be used for the QM1-8 matching quadrupoles in the IP switch. The center-to-center beam line separation between

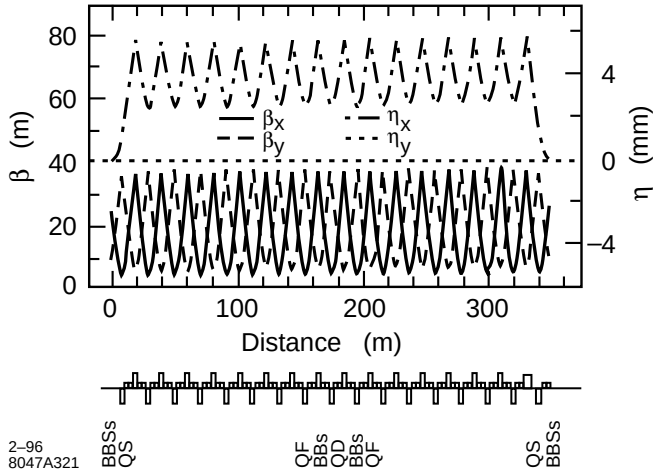


Figure 10-4. *Big bend optics for separated function lattice.*

Total Length (FODO cells)	(m)	312
Total bend angle	(mr)	8.5
Number of FODO cells [a]		15
Maximum $\beta_{x,y}$	(m)	36
Maximum η_x	(mm)	5.6
x-phase advance/cell	(°)	108
y-phase advance/cell	(°)	90
Spin phase advance/cell	(°)	37 (55)
Dipole magnet length	(m)	3.0
Bend radius (per dipole)	(km)	22.6
F-quad length	(m)	2.50
D-quad length	(m)	2.26

[a] There are four dipole magnets per cell.

Table 10-6. *Optimized big bend parameters at 500 GeV/beam (750 GeV/beam) for 6.7-kGauss (10-kGauss) quadrupole pole-tip fields at 6 mm radius and tolerable SR emittance growth.*

the two big bends at the face of the first QD magnet is 28.8 cm (for either 10 or 20 IP switch dipoles) which sets an upper limit on the outer horizontal dimension of the big bend quadrupole magnets of $<$28 cm full width.

10.3.2 Chromatic Emittance Dilution

Tracking studies using TURTLE [Carey 1982] have been made for the entire beam line described in this chapter (nearly 800 meters of beam line from the end of the collimation section to the beginning of the final focus). A Gaussian energy distribution with rms of 0.2%, 0.4%, and 0.6% energy spread was used. For all cases, the emittances used were $\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m and $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m. The chromatic emittance dilution at 500 GeV/beam for each case is listed in Table 10-9 (SR effects not included). Sextupole compensation is not necessary.

Name	Number	Length (m)	Half Gap (mm)	Field (kGauss)
BB	64	3.0	6	0.738 (1.108)

Table 10-7. *Big bend dipoles for 500 GeV/beam (750 GeV/beam).*

Name	Number	Length (m)	Radius (mm)	Pole Tip Field (kGauss)
QD	17	2.258	6	-6.70 (-10.05)
QF	16	2.500	6	+6.70 (+10.05)

Table 10-8. *Big bend quadrupoles for 500 GeV/beam (750 GeV/beam).*

Rms energy spread (%)	$\Delta\epsilon_x/\epsilon_{x0}$ (%)	$\Delta\epsilon_y/\epsilon_{y0}$ (%)
0.2	1.2	0.2
0.4	3.7	1.6
0.6	7.7	4.2

Table 10-9. *Chromatic emittance increase at 500 GeV/beam for various rms Gaussian energy spreads for beam line including initial diagnostic section, IP switch, all matching sections, big bend, and the skew correction and diagnostic section (synchrotron radiation effects not included).*

10.3.3 Synchrotron Radiation Effects

Optics

Table 10-10 lists the SR parameters for the 500-GeV/beam big-bend design of Figure 10-4. The 750 GeV/beam parameters are also given. The fractional emittance growth referred to is the main damping ring extracted emittance of $\gamma\epsilon_x = 3 \times 10^{-6}$ m. The energy loss across the length of the big bend is 0.07% at 500 GeV (0.22% at 750 GeV). Given the large chromatic bandpass of the big bend (less than 2% emittance increase at 1% rms Gaussian energy spread), it is not necessary to taper the fields through the system.

Beam Energy	(GeV)	500	750
Critical Energy (u_c)	(MeV)	12	42
SR-generated rms Energy Spread	(%)	0.015	0.040
Energy Loss	(GeV)	0.331	1.675
Horizontal SR Emittance Growth	(%)	0.3	3.3
Number of photons/electron		90	456

Table 10-10. *Synchrotron radiation parameters for big bend at 500 GeV and 750 GeV/beam for $\gamma\epsilon_x = 3 \times 10^{-6}$ m.*

NAME	Quantity	roll (mr)	$\Delta B/B_0$ (%)	b_1/b_0 (%)	b_2/b_0 (%)
BB	64	45	0.014	35.0	7500

Table 10-11. *Big bend dipole magnet single-element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3\times10^{-6}$ m, $\gamma\epsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.3\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 4 mm. The sextupole component tolerances for the dipoles at $r = 4$ mm are extremely loose.*

Detector Backgrounds

The energy distribution generated by SR has a long tail which falls off rapidly for energies, u, well above the critical energy, u_c ($u/u_c \equiv \xi >> 1$) [Sands 1970].

$$n(\xi) \propto \xi^{-1/2} e^{-\xi} \tag{10.3}$$

To estimate the number of electrons in the tail which achieve an oscillation amplitude comparable to one rms horizontal beam size (in the interest of staying clear of the final doublet face), the necessary energy deviation, u_1, is written in terms of the rms SR energy spread, σ_δ, the relative horizontal SR emittance increase, $\Delta\epsilon_x/\epsilon_{x0}$, and the beam energy, E_0.

$$\xi_1 \equiv \frac{u_1}{u_c} = \frac{\sigma_\delta}{u_c/E_0} \cdot \frac{1}{\sqrt{\Delta\epsilon_x/\epsilon_{x0}}} \sim 40 \ (\text{at } E_0 = 750\,\text{GeV}) \tag{10.4}$$

For the worst case (750 GeV/beam), a particle which is 40 critical energies lower than nominal will oscillate at one sigma. The number of electrons per bunch at or beyond this energy, N_{e^-}, is calculated in Equation (10.5) where N_γ is the total number of photons/e^-. Even for 10^{10} electrons per bunch at 750 GeV/beam, this energy tail is insignificant and will not generate a background.

$$N_{e^-} = 10^{10} \cdot \frac{9\sqrt{2}}{15} N_\gamma \int_{\sqrt{2\xi_1}}^{\infty} e^{-x^2/2} dx \sim 10^{-4} \tag{10.5}$$

10.3.4 Tuning, Tolerances, and Corrections

The single-element tolerances for the big bend magnets are listed in Tables 10-11 (dipoles) and 10-12 (quadrupoles). Each tolerance represents a 2% luminosity loss for that single element's effect on one beam. The effects of these errors generally increase the IP beam size except in the case of dipole field regulation and quadrupole transverse vibration which continuously steer the beams out of collision. In this case, since the exact betatron phase to the IP is not calculated, phase averaging is applied. The tolerances given in the tables have not yet been distributed out into a weighted tolerance budget; the numbers are for reference. In fact, given multiple errors over multiple elements, these tolerances are much too loose. However, since tuning considerations have not been folded in, most static, non-steering errors may also be corrected over some reasonable range.

At this time, detailed tolerance and tuning studies have not been performed. However, the big bend design must include dispersion tuning elements for both planes and betatron phases to correct any residual dispersion due to magnet misalignments and gradient errors. Coupling and matching corrections as well as diagnostics exist just after the big bend (Chapter 11). Vertical dispersion correction can be provided by adding four small skew quadrupoles (of zero nominal field)—one per cell in the last four cells. This scheme takes advantage of the 90° vertical phase advance per cell by pairing skew quadrupoles at $-I$ transfer matrix (2 cell) separation so that, for equal and opposite skew quadrupole settings, no betatron cross-plane coupling is generated. The second pair of skew quadrupoles then handles

NAME	Quantity	roll Δx (mr)	Δy offset (μm)	$\Delta x_{\rm rms}$ offset (μm)	$\Delta y_{\rm rms}$ vibrate (μm)	$\Delta B/B_0$ vibrate (μm)	b_2/b_1 (%)	(%)
QD	17	4.4	680	26	2.1	0.077	6.9	2400
QF	16	4.3	230	57	0.70	0.170	3.5	500

Table 10-12. *Big bend quadrupole magnet single-element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Sextupole field component tolerances (b_2/b_1) are evaluated at a radius of 4 mm and are very loose ($\geq 500\%$).*

Quad Type	Quantity	Length (m)	pole radius (mm)	max. field (kGauss)	rms reg. tolerance (%)	$\Delta\eta_{x,y}^{\rm max}$ at $\beta_{x,y}^{\rm max}$ (mm)	max. ϵ_y/ϵ_{y0}
skew	4	0.5	6	± 5	0.1	4.6	13
normal	4	0.5	6	± 8	0.5	14	5.1

Table 10-13. *Big bend dispersion tuning magnet specifications at 500 GeV/beam for 0.3% rms energy spread, $\gamma\epsilon_{x0} = 3\times10^{-6}$ m and $\gamma\epsilon_{y0} = 3\times10^{-8}$ m. There is one skew quadrupole in each of the last four FODO cells 10 cm downstream of the QD and one normal quadrupole in each of cells 9, 10, 14 and 15 at 10 cm upstream of the QF.*

the other betatron phase. The range of correction for one pair of 50-cm-long, ± 5-kGauss pole-tip field, 6-mm pole-tip radius skew quadrupoles located 10 cm downstream of each QD at $\beta_y = 32$ m, $\eta_x = 2.8$ mm is $\epsilon_y/\epsilon_{y0} = 13.4$ at 0.3% rms energy spread for $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m at 500 GeV/beam (the vertical dispersion induced at the center of a QD magnet is as much as 4.6 mm). No significant coupling or horizontal beta function perturbation is generated over this range. However, for very large corrections some second order dispersion may be induced which will limit the correction range or require similar skew sextupole tuners. This level has not yet been studied.

The horizontal dispersion may be controlled similarly by adding two pairs of small normal quadrupoles (of zero nominal field). Due to the 108° horizontal phase advance per cell these quad pairs must be spaced by 5 cells to provide a $-I$ separation. If the 9th and 10th as well as the 14th and 15th cell include a 50-cm-long, ± 8-kg, 6-mm-radius quadrupole which is 10-cm upstream of each QF (at $\beta_x = 31$ m, $\eta_x = 5.2$ mm) the emittance correction range per pair will be $\epsilon_x/\epsilon_{x0} = 5.1$ (the additional horizontal dispersion induced at the center of a QF magnet is as much as 14 mm). However, since the vertical transfer matrix between paired normal quadrupoles (5 cells) is not equal to $-I$, there will be a small perturbation to the vertical beta function which amounts to a 10% beta beat amplitude at full horizontal dispersion correction (± 8 kGauss quadrupole fields). This small effect is correctable with the matching quadrupoles just upstream of the pre-final focus emittance diagnostic section (Chapter 11). The dispersion correction specifications are summarized in Table 10-13.

Given the dispersion correction available, the big bend quadrupoles will probably not require movers. However, beam-based alignment techniques will greatly benefit in speed and convergence if movers (at least in the vertical plane) are available. The movers should control the vertical position to $\sim$5-μm resolution over a range of approximately $\pm 500\,\mu$m. Roll control is not required given the fairly loose roll tolerances as well as the skew (Chapter 11) and vertical dispersion corrections available. Beam-based alignment of the big bend quadrupoles has not yet been studied in detail. However, an independent partial current shunting switch across each big bend quadrupole will probably be a significant advantage for any beam-based alignment algorithm.

Horizontal and vertical dipole correctors at each QF and QD, respectively, will be required to initially steer the beam line and to use in fast feedback applications. Correctors with ± 1.0-kg fields and 25-cm length will be adequate to displace the beam nearly $\pm 500\,\mu$m at the next similar quadrupole at 500 GeV/beam. The horizontal correctors will

then need to regulate at $\sim 1 \times 10^{-3}$ over the 100-ms (10 pulse) range while similar vertical correctors will need $\sim 1 \times 10^{-4}$ regulation ($\sim$0.3% luminosity loss due to all correctors in both big bends for both planes).

The tolerance on the beta match into the big bend is quite loose. It can be shown that the SR emittance increase approximately scales with the amplitude of the incoming beta mismatch.

$$\Delta\varepsilon_{SR} \approx B_{\rm mag}\Delta\epsilon_{\rm SR-nom.} \tag{10.6}$$

Here $B_{\rm mag}(\geq 1)$ is the beta mismatch amplitude in the horizontal plane and $\Delta\varepsilon_{\rm SR-nom.}(\ll \varepsilon_{x0})$ is the nominal SR emittance increase for a matched incoming beam. A very large mismatch of $B_{\rm mag} = 2$ $(\beta_x \approx 4\beta_{x0}, \alpha_x = \alpha_{x0} = 0)$ will amplify a nominal 0.3% SR emittance increase to 0.6%. The vertical match has no such constraint.

10.3.5 Spin Transport and Depolarization

The spin phase advance per cell (spin tune) has also been tabulated in Table 10-6. A spin tune-betatron tune resonance is to be avoided, or small vertical oscillations will precess the electron spin into the vertical plane [Limberg 1993]. However, even without a resonance there may be significant vertical alignment error induced spin rotation due to the large gradient magnets and the extremely high energy. A 100-μm vertical beam offset at 500 GeV/beam in a single QD magnet will rotate a longitudinally oriented spin vector 1° into the vertical plane. If the errors are static this may be compensated by properly orienting the incoming spin vector using the 2-GeV solenoid rotator system (Chapter 5).

The depolarization for a bend through θ, at an energy γ, and an incoming Gaussian rms energy spread of σ_δ is

$$\overline{P}/P_0 = \frac{1}{\sqrt{2\pi}\sigma_\delta}\int_{-\infty}^{\infty} e^{-\delta^2/2{\sigma_\delta}^2}\cos(a\gamma\delta\theta)d\delta = e^{-(a\gamma\theta\sigma_\delta)^2/2} \quad . \tag{10.7}$$

where $a = (g-2)/2$ is the anamolous magnetic moment. For a 10-mr bend at 500 GeV/beam with a 0.3%-rms energy spread the relative depolarization is 0.06% (0.13% at 750 GeV).

10.3.6 Vacuum System

The pressure requirements for the big bend section are set by tolerable detector background levels [Irwin 1993]. At present, it is desirable to achieve an average pressure of $\sim 5 \times 10^{-8}$ Torr in the big bend. If the chamber is cylindrical and made of aluminum with specific outgassing rate $q = 5 \times 10^{-10}$ T-l/s-m (similar to mature SLC arcs), a specific conductance for a 6-mm-radius of $c = 0.18$ m-l/ and ion pumps of speed $S > 5$ l/s placed three per FODO cell ($L \sim$7 m for a total of 45 ion pumps), the system is conductance-limited with average pressure [Ziemann 1992].

$$\overline{P} \cong qL^2/3c \approx 5 \times 10^{-8}\text{Torr} \quad . \tag{10.8}$$

An order-of-magnitude-lower pressure is probably achievable by using a baked stainless steel chamber with a much lower specific outgassing rate.

References

[Keller 1993] L.P. Keller, "Muon Background in a 1.0-TeV Linear Collider", SLAC–PUB–6385 (October 1993).

[Helm 1973] R.H. Helm, M.J. Lee, P.L. Morton, M. Sands, "Evaluation of Synchrotron Radiation Integrals", SLAC–PUB–1193 (March 1973).

[Raubenheimer 1993] T.O. Raubenheimer, P. Emma, S. Kheifets, "Chicane and Wiggler Based Bunch Compressors for Future Linear Colliders", SLAC–PUB–6119 (May 1993).

[Bane 1995] K. Bane, private communication (1995).

[Sands 1970] M. Sands, "The Physics of Electron Storage Rings", SLAC–121 (November 1970).

[Limberg 1993] T. Limberg, P. Emma, R. Rossmanith, "The North Arc of the SLC as a Spin Rotator", *Proc. of the 1993 Part. Acc. Conf*, Washington, DC (1993).

[Irwin 1993] J. Irwin, R. Helm, W.R. Nelson, D. Walz, "Conventional Collimation and Linac Protection", SLAC–PUB–6198 (May 1993).

[Brown 1977] K.L. Brown *et al.*, "TRANSPORT", SLAC–91 (May 1977).

[Carey 1982] D.C. Carey *et al.*, "Decay TURTLE", SLAC–246 (March 1982).

[Ziemann 1992] V. Ziemann, "Vacuum Tracking", SLAC–PUB–5962 (October 1992).

[Emma 1994] P. Emma, T. Limberg, R. Rossmanith, "Depolarization in the SLC Collider Arcs", *Proc. of the 1994 European Part. Accel. Conf.*, London, England (July 1994).

Contributors

- Paul Emma
- Dick Helm
- Tor Raubenheimer
- Mark Woodley

11

Final Focus

Contents

11.1 Introduction

The final-focus system comprises the region between the big bend and the main beam dump. Its function is to demagnify the transverse sizes of electron or positron beams by a factor 80 horizontally and 300 vertically, down to a value of about 250 nm $\times$ 4 nm at the interaction point (IP), where the beams collide. After the collision, an extraction line guides the beam remnants onto the beam dump.

The final-focus designs for 500 GeV and for 1 TeV which are described in this chapter, have a comfortable momentum bandwidth and can operate in the full ZDR parameter plane. A dedicated geometry-adjustment section at the entrance to the final focus facilitates the adiabatic upgrade from 350-GeV to 1.5-TeV-c.m. energy, with a constant IP position and only minor transverse displacements.

Important design issues of the final-focus system are linear and nonlinear optics; momentum bandwidth; effects of synchrotron radiation; tuning schemes; sensitivity to varying beam conditions, such as incoming orbit, emittance, and energy; tolerances on alignment, vibrations and field changes; response to ground motion; wakefield effects; stability of the final doublet; design and tolerances of the crab cavity; compensation of the solenoidal detector field; beam removal from the IP; maintenance, tuning and stabilization systems.

When compared to previous colliders, the NLC beam parameters and tolerances enter a distinctly new regime. The spot size at the IP is 125 times smaller than that at the SLC, which is already as small as 500 nm, and is 15 times smaller than the 70-nm spot size achieved at the FFTB. In addition to delineating the stringent tolerances which are a consequence of the small spot size, we will attempt to clarify how we propose to achieve these tolerances. The story is complex since there are several significant timescales, several distinct aberrations to consider at each timescale, many elements and element parameters which can influence each aberration, and many sources of change for the parameters of each element. Each timescale is associated with a tuning system and a maintenance system which stabilizes the final focus between tunings. We use the word tuning when the diagnostic involves determination of IP beam and collision conditions. Hence missteering at the IP is one of the aberrations which is tuned.

In the case of alignment, because everything is moving, it is often difficult to clearly convey what elements need to stay aligned, by how much, and with respect to what other elements or coordinate system. An important result of ground-motion studies, which are described in Appendix C, is that the ground (bedrock) can be used as a reference, in the sense that if all elements were moving as the bedrock below them, then the beams would remain in collision. Section 11.5.5 discusses these results as they apply to the final-focus system.

Aberrations are described in Section 11.5.3. A table in that Section (Table 11-9) lists all aberrations that will be tuned, how they will be tuned, frequency of tuning, what budget of luminosity is allotted to the final-focus system (since other upstream elements can contribute to these aberrations), and what systems maintain (stabilize) the system between tunings. Tuning and maintenance systems are described in more detail in Sections 11.3.1, 11.3.2, and 11.5.4. 11.5.4 and 11.5.8.

In Section 11.5.5, relevant timescales, each of which corresponds to a tuning frequency, are listed and named, and all tolerances described are specified relative to these timescales. Tolerances can be achieved if:

- The diagnostic systems used for tuning have the required resolution.
- Tuning and adjustment knobs have the required sensitivity.
- Maintenance systems stabilize the final focus between tunings.

The performance of the maintenance systems in turn depends on achieving the required resolution in their diagnostic systems. Diagnostic systems are described in Section 11.5.8.

In the SLC, only the IP steering and incoming beam launch conditions have been automated. Though a dither feedback has been considered for the other IP aberrations, it would be speculative at this point to believe such a system would work at the NLC, and we do not make this assumption. Automated tuning (steering) and automated maintenance systems are described in Section 11.5.6.

Tables 11-13 through 11-18 list the sensitivities of the parameters of each element. Using these sensitivities a tolerance is assigned to each element (Table 11-19) so that the system budget of each aberration is met. Elements with similar tolerances have been grouped together. Table 11-21 gives the luminosity loss that results from these tolerance assignments, to be compared with the final-focus system budget of Table 11-9. Corresponding to each tolerance in Table 11-19, there is a brief statement on how this tolerance can be met (Table 11-20). Finally, in Section 11.5.8, we present a table that lists every diagnostic element and the resolution required.

It is worth noting that both the optical design and proposed tuning algorithms draw heavily from SLC and FFTB experience. In the limited time available, every small aspect of the design could not be studied in minute detail. However, we feel that all potentially critical or problematic topics have been addressed in depth. Outstanding work includes, for instance, a more detailed study of certain tuning and alignment procedures, and a more exhaustive description of commissioning, operation and machine protection. These items are not expected to be difficult.

One of the reviewers of the August 95 ZDR workshop (K. Oide) has created an alternative final-focus design for NLC parameters at 500-GeV-c.m. energy, which is based on the odd-dispersion scheme [Oide 1992]. The bandwidth of this system, without any additional `Brinkmann'-sextupoles, is comparable to that of the present NLC design. An advantage of the odd-dispersion final focus is that it only uses about half the number of quadrupoles. A potential disadvantage is the nonexistence of an IP pre-image point, which may or may not be an operational aid. The tunability, upgradability, performance at higher energy, and background situation of the odd-dispersion design need to be evaluated, prior to a final decision on this alternative. In the present report, only the modular final-focus design by R. Helm is discussed, the performance of which appears to be entirely satisfactory.

11.2 Parameters and Specifications

11.2.1 Goals and System Boundaries

The purpose of the NLC final-focus system is to transport electron and positron beams of energy 180 GeV to 750 GeV from the end of the big bend to the IP, where the demagnified beams are collided, and to remove the beam remnants cleanly to facilitate crucial post-IP diagnostics. To accomplish its task, the final-focus system has to be stable and reliable, and it needs redundant diagnostics to detect and compensate all changes of beam parameters, magnet positions, or field strengths, which would otherwise reduce the luminosity.

For a c.m. energy of 500 GeV, the design spot size at the interaction point (IP) is about 4.2–6.5 nm vertically and 250–300 nm horizontally. The normalized emittances at the entrance of the final focus are assumed to be $\gamma\epsilon_x \approx 4\times10^{-6}$ m and $\gamma\epsilon_y \approx$ 7–10.5 $\times10^{-8}$ m. At a c.m. energy of 1 TeV, the design spot size is 3.4–5.2 nm vertically and 200–250 horizontally, for normalized emittances of $\gamma\epsilon_x \approx 4\times10^{-6}$ m, and $\gamma\epsilon_y \approx$ 9–13.5$\times10^{-8}$ m. The beam energy distribution in the final focus depends on bunch compressor and linac configurations. For most studies presented in this chapter, a Gaussian momentum distribution with a relative rms momentum spread of $\delta \approx 0.3\%$ is assumed. This distribution makes it easy to study the effect of a typical energy spread, but the real energy distribution will be non-Gaussian (compare Chapters 5 and 7).

The entrance of the final-focus system is formed by a skew correction section (SCS) and a diagnostics section (DS). These are followed by a geometry-adjustment section (GAS), beta- and phase-matching section (BMS), horizontal and vertical chromatic correction sections (CCX and CCY), which are separated by a beta-exchange module (BX), the final transformer (FT), the interaction region (IR), and the beam removal system (BRS). This chapter describes all these sections except for the IR which is discussed separately in Chapter 12.

The final-focus system is flexible enough to be operated in the entire c.m. energy range from 350 GeV to 1.5 TeV. Most of the data and figures in the following sections refer to the design for either 500-GeV or 1-TeV-c.m. energy.

11.2.2 Parameter List

Table 11-1 exemplifies beam parameters at the interaction point for c.m. energies of 500 GeV and 1 TeV. In this table, the values listed for luminosity and spot size do not include any dilutions, which may arise from high-order aberrations, synchrotron radiation residual uncorrected low-order aberrations, timing offsets or crab crossing errors. When dilutions are taken into account, the expected luminosity is reduced by about 20–30%, see Table 11-4 and the discussions in Section 11.5.

Most beam parameters are variable, as the NLC should function at any point inside a multi-dimensional operating plane [Burke 1995]. From the final-focus point of view, the largest IP divergences represent the worst case, since in this case the effect of nonlinear aberrations, the aperture requirements, and the Oide effect are most severe. The beam parameters listed in Table 11-1 refer to this case.

The minimum horizontal beta function at the IP is limited by nonlinear aberrations, by the maximum number of beamstrahlung photons that can be tolerated Chapter 12. and by the Oide effect (Sections 11.5.3 and 11.6.4). The value presently chosen was imposed by the beamstrahlung, whereas Oide effect and nonlinearities would allow for a 20% smaller beta function. It is, therefore, possible to compensate the increase of the horizontal spot size due to residual low-order aberrations by reducing β_x^*.

11.2.3 Energy Flexibility

The final focus should operate at least in the energy range from 350 GeV to 1.5 TeV. In the present design, this flexibility is accomplished with slightly different geometries for three overlapping energy ranges (Figures 11-9 and 11-10). The implication is that two minor reinstallations of magnets and supports are necessary during the energy upgrade, at around 500 GeV and 1.1 TeV. The IP position is held constant by means of a special bending section—the geometry-adjustment section—which is located at the entrance to the final focus.

At 500-GeV-c.m. energy, the final focus is operated with quadrupoles scaled down from the 1-TeV design. A complication arises in the final doublet, since the last quadrupole is a permanent magnet. Here, the energy scaling is accomplished by an adjacent superconducting quadrupole whose field changes sign during the energy raise from 500 GeV to 1 TeV. In order to keep the cost of a further upgrade low, the remaining final-focus magnets could be designed such that their strength can be increased to the 1.5-TeV values.

Comments	c.m. energy	500 GeV	1 TeV
Luminosity w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.52	1.16
Luminosity w. pinch w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.75	1.63
# particles per bunch	N_b (10^{10})	0.65	0.95
# bunches	n_b	90	90
# bunch trains per s	f	180	120
Enhancement factor w/o dil.	H_D	1.46	1.41
Hor. spot size w/o dil.	σ_x (nm)	253	200
Vert. spot size w/o dil.	σ_y (nm)	4.18	3.35
Hor. IP beta function	β_x^* (mm)	8	10
Vert. IP beta function	β_y^* (mm)	0.125	0.125
Norm. hor. emittance	$\gamma\epsilon_x$ (10^{-8} mr)	400	400
Norm. vert. emittance	$\gamma\epsilon_y$ (10^{-8} mr)	7	9
Hor. IP divergence	$\sigma_{x'}$ (μrad)	31.6	20.0
Vert. IP divergence	$\sigma_{y'}$ (μrad)	33.5	26.8
Bunch length	σ_z (μm)	100	125
Crossing angle	θ_c	20 mr	20 mr
Energy spread	δ_{rms}	$\geq 3 \times 10^{-3}$	$\geq 3 \times 10^{-3}$
Free length from IP	l^* (m)	2	2

Table 11-1. *Basic worst-case interaction-point beam parameters without dilutions (see also Table 11-4).*

11.2.4 Overview

On the next pages, the skew correction section (SCS) is discussed, along with two options for the adjacent diagnostics section (DS). Section 11.4 is devoted to the beta- and phase-matching section (BMS), which will be used to adapt the IP beta functions and the waist position to varying incoming beam conditions, and to adjust the betatron phase advance between the collimator section and the IP. The geometry-adjustment section (GAS) located upstream of the BMS, the horizontal chromatic correction section (CCX), the beta-exchange module (BX), the vertical chromatic correction section (CCY) and the final transformer (FT), as well as a tolerance analysis for this region are described in Section 11.5, which also contains lists of magnets and diagnostics for the entire region between the SCS and the IP. Section 11.6 discusses the final doublet; Section 11.7 discusses effects of the solenoidal detector field and the crab cavity. The beam removal system (BRS) and beam dump are described in Section 11.8. Section 11.9 concludes the chapter with a short perspective on the present design and on outstanding questions.

11.3 Skew Correction and Diagnostic Section

This beam line section follows the big bend and is used to measure and correct any anomalous cross-plane coupling and to precisely match the beam into the main body of the final focus. In order to minimize the projected emittance, especially for a flat beam, it is necessary to remove all cross-plane correlations. SLC experience has shown this to be a very difficult problem when provisions are not included in the design to reliably measure and/or correct all four betatron correlation phases. Rather than placing skew quadrupoles and wire scanners as post-design space allows, a dedicated coupling correction and diagnostic section is highly desirable, especially for large emittance aspect ratios as in the NLC ($\varepsilon_{x0}/\varepsilon_{y0} \sim 100$).

11.3.1 The Skew Correction Section

The ideal Skew Correction Section (SCS) contains four skew quadrupoles separated by appropriate betatron phase advance in each plane such that the skew correctors are orthonormal (orthogonal and equally scaled). A simple realization of such a section is possible if the skew quadrupoles each correct one of the four beam correlations $< xy >$, $< x'y >$, $< xy' >$, $< x'y' >$ and if, in addition, the values of the product of horizontal and vertical beta functions are equal at each of the skew quadrupoles. The relative emittance dilution for a thin skew quadrupole of focal length f is

$$\frac{\varepsilon_y}{\varepsilon_{y0}} = \sqrt{1 + \frac{\varepsilon_{x0}}{\varepsilon_{y0}} \cdot \frac{\beta_x \beta_y}{f^2}} \quad , \qquad \frac{\varepsilon_x}{\varepsilon_{x0}} = \sqrt{1 + \frac{\varepsilon_{y0}}{\varepsilon_{x0}} \cdot \frac{\beta_x \beta_y}{f^2}} \quad . \tag{11.1}$$

For a flat beam $(\varepsilon_{x0}/\varepsilon_{y0} \gg 1)$, and the vertical emittance is much more sensitive. Orthogonality of the skew quadrupoles is achieved by separating the first and second and also the third and fourth skew quadrupoles by FODO cells with betatron phase advances of $\Delta\psi_x = \Delta\psi_y = \pi/2$, and separating the second and third skew quadrupoles by $\Delta\psi_x = \pi, \Delta\psi_y = \pi/2$. Then if the first skew quadrupole controls the xy phase (by definition here) the second controls the $x'y'$ phase, the third is the $x'y$ phase and the fourth is at the xy' phase. This scheme allows total correction of any arbitrary linearly coupled beam with correction range limited only by available skew quadrupole strength. Initially (at zero strength) the skew quadrupoles are orthonormal. As correction strength is applied there is some deviation from this ideal situation due to the slight effect on the in-plane optics. For large corrections $(\varepsilon_y/\varepsilon_{y0} > 2)$. applied by repeated minimization of the projected emittance, some iteration may be necessary (Section 11.3.3). Figure 11-1 shows beta functions and quadrupole locations for the SCS. The full length of the correction section, $L_{\rm SCS}$, is limited by achievable skew quadrupole pole-tip field, $|B_0|_{\rm max}$, its length, l, and its pole-tip radius, r, as well as the maximum correctable vertical emittance dilution, $(\varepsilon_y/\varepsilon_{y0})_{\rm max}$, the nominal emittance aspect ratio, $(\varepsilon_{x0}/\varepsilon_{y0})_{\rm nom}$, and the beam energy, expressed here as magnetic rigidity, $(\beta\rho)$,

$$L_{\rm SCS} \geq \frac{15}{2} \cdot \frac{r(B\rho)}{l|B_0|_{\rm max}} \sqrt{\frac{\left(\varepsilon_y/\varepsilon_{y0}\right)^2_{\rm max} - 1}{\left(\varepsilon_{x0}/\varepsilon_{y0}\right)_{\rm nom}}} \quad . \tag{11.2}$$

Note, Eq. 11.2 is a thin-lens approximation for both the skew and the FODO cell quadrupoles. It does not hold for skew or FODO cell quadrupoles with focal lengths comparable to their magnetic lengths. For a 500-GeV/beam system intended to correct up to a factor of three in emittance dilution, and for an emittance aspect ratio of 100 with 30-cm-long, 6-mm-radius, $\pm$4-kGauss-skew quadrupoles, the system is $\sim$170 m in length. While it would be possible to reduce this length by $\sim$20%, the matching into the diagnostic section (Section 11.3.2) is simplified with the slightly longer design.

11.3.2 The Diagnostic Sections

In order to tune the skew quadrupoles and other matching elements an emittance measurement must be made downstream of the correction. In what follows, only emittance measurements made with multiple wire scanners [Ross 1991] are considered. The emittance measurement may be either a single plane (2D) measurement where the coupling is only inferred by the dilution, or it may be a fully coupled (4D) measurement where all four coupling parameters as well as the two intrinsic transverse emittances are measured. The following presents a design for each scheme with ideal wire-to-wire phase advances and also constant matched spot sizes and aspect ratios. In the present NLC design, a 2D section will be used immediately downstream of the post-linac collimation section, while a 4D section will be used in conjunction with a skew-correction section at the input to the final focus.

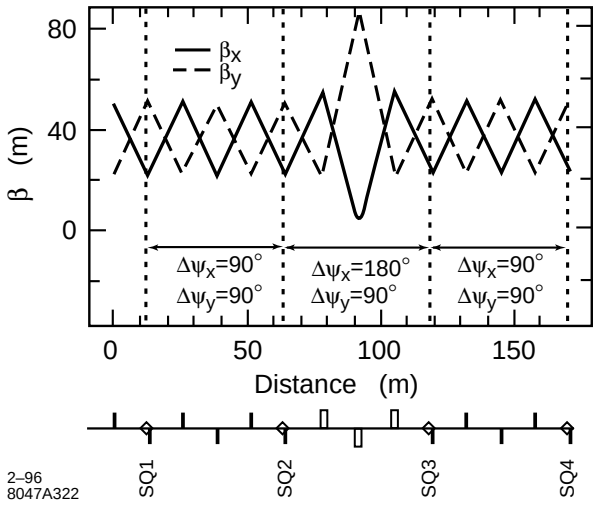


Figure 11-1. *Skew correction section (SCS) optics.*

The 2D Emittance Measurement Scheme

A space-saving, economical 2D emittance measurement system is probably adequate for regions in NLC where low levels of coupling are expected. In this case the optimal wire-to-wire phase advance per plane is π/N, where N is the number of wire scanners.[1] This conclusion is clear by viewing the normalized matched phase space as a circle with beam size measurements made at π/N phase intervals. With three single plane parameters to measure (ε, β, α) a four-scanner measurement provides some redundancy as well as better phase coverage for poorly matched beams. By separating each wire with a FODO cell of $\mu_{x,y} = 45°$, the phase coverage is optimum and, in addition, the matched beam will have a constant size per plane. Furthermore the aspect ratio, a, produced at (or near) a vertically focusing quadrupole is very reasonable and does not require precise wire-scanner roll alignment tolerances.

$$a \equiv \frac{\sigma_x}{\sigma_y} \cong \sqrt{\frac{\varepsilon_{x0}}{\varepsilon_{y0}} \cdot \frac{1-\sin\frac{\mu}{2}}{1+\sin\frac{\mu}{2}}} \approx \frac{2}{3}\sqrt{\varepsilon_{x0}/\varepsilon_{y0}} \tag{11.3}$$

In the case of the NLC $a \approx 6.7$ which requires an easily achieved roll tolerance[2] of $|\psi| \leq \sin^{-1}\left(\sqrt{2}/10\sqrt{3}a\right) \sim 1°$. Each wire scanner need only measure the x and y beam sizes (the correlation is not necessary in this 2D scheme). Figure 11-2 shows the lattice which includes the SCS followed by a matched 2D/four-scanner diagnostic section. The coupling is fully corrected by repeatedly minimizing the measured vertical emittance with each skew quadrupole in turn (some iteration may be necessary). Such a system should probably be preceded by four beta-matching quadrupoles to facilitate compensation of the slight in-plane focusing effect of energized skew quadrupoles. A clear advantage to this scheme is that the uncoupled, matched beam will be clearly evident as four precisely equal x and four precisely equal y beam sizes on the scanners. The lower limit on the length of the diagnostic section, $L_{2\mathrm{D}}$, is set by the minimum measurable beam size, σ_y, the normalized emittance, ε_{yN}, and the beam energy, γ.

$$L_{2D} \geq \frac{3}{2}\frac{\sigma_y^2\gamma}{\varepsilon_{yN}} \tag{11.4}$$

[1] The arguments presented here are the same whether these are actual carbon-filament-wire scanners or multibunch-capable laser-wire scanners.
[2] The following relation holds only for $a >\sim 3$ and achieves a systematic emittance measurement error of <2%.

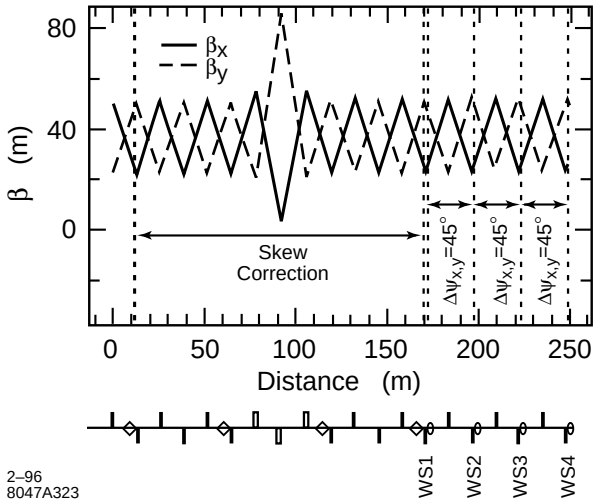


Figure 11-2. *SCS plus 2D emittance diagnostics section optics.*

For a 500-GeV beam with ∼1.5-μm vertical beam size and a normalized emittance of 5×10^{-8} m, the minimum length is ∼70 m.

The 4D Emittance Measurement Scheme

If it is expected that coupling may be a significant problem or the application of a faster, calculated correction is desirable, the 4D measurement scheme may be preferable. In this case a precise measurement of the four linear coupling coefficients is made with a set of σ_x, σ_y, and σ_{xy} measurements made for each of six wire scanners. Each wire scanner must have three independent angle filaments—in the simplest case, a horizontal, a vertical and a 45° filament typically used in the SLC [Ross 1991]. The wire-to-wire phase advances are chosen in a similar way to the skew quadrupole placement in the SCS (Section 11.3.1). With the first wire measuring the xy correlation, a second wire at $\Delta\psi_x = \Delta\psi_y = \pi/2$ measures $x'y'$ and a third wire advanced again by $\Delta\psi_x = \pi, \Delta\psi_y = \pi/2$ measures $x'y$ and finally a fourth at another $\Delta\psi_x = \Delta\psi_y = \pi/2$ sees xy'. The problem with this four-wire scheme is that the in-plane measurements are not determined. A fifth and sixth wire must be added to cover the missing single plane phases. Figure 11-3 shows the six-wire solution including the SCS, which is ideal for both the single plane and the coupling measurements. There are 10 parameters to measure ($\varepsilon_{x,y}$, $\beta_{x,y}$, $\alpha_{x,y}$, $< xy >$, $< x'y' >$, $< x'y >$ and $< xy' >$) and up to 18 profiles are scanned leaving 8° of freedom. It is also possible to use a subset of the 18 profiles in order to speed the measurement process. Note that this system can also be used as a 2D system by making single plane profile scans with wires WS3, WS4, WS5 and WS6 of Figure 11-3. The minimum length is approximately twice that of the 2D system.

$$L_{4\mathrm{D}} \geq \frac{\pi\sigma_y^2\gamma}{\varepsilon_{yN}} \tag{11.5}$$

11.3.3 Tuning Simulations

Tuning simulations were run using the Final Focus Flight Simulator [Woodley 1994] to test the convergence of the skew correction for a 500-GeV beam with an intrinsic emittance aspect ratio ($\varepsilon_{x0}/\varepsilon_{y0}$) of 100. Figure 11-4 shows the

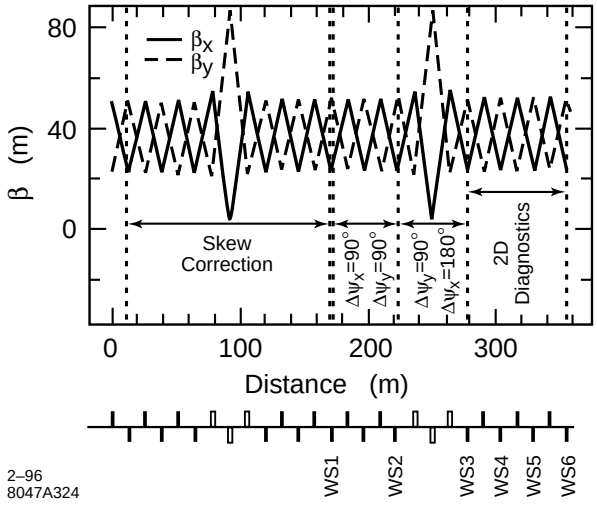


Figure 11-3. *SCS plus 4D emittance diagnostics section optics.*

relative vertical emittance achieved with each skew quadrupole scan. In this case, with the input coupling diluting the vertical emittance by more than a factor of two, a correction to less than 1% dilution is achieved after the first pass through all four skew quadrupoles. This beam line setup operation could probably be executed in less than an hour of real machine time. Of course, using the 4D system here could conceivably produce a reasonably well-calculated correction within a period of minutes. The final choice of appropriate systems depends on the expected phase space stability and magnitude of the errors.

11.4 Beta-Matching Section

The purpose of the NLC Beam Delivery Beta-Matching section is to provide a set of magnets upstream of the horizontal chromatic correction section which can be used to respond to variations in the incoming beam and/or the desired beam parameters at the IP, and also can be used to set the desired phase advance between the collimators in the Post-Linac Collimation region (Section 9.4) and the Final Doublet. Experiences at the SLC Final Focus and the Final Focus Test Beam have demonstrated that it is useful to be able to reduce the angular divergences at the IP (and hence the beam sizes inside all the other magnets in the Final Focus), while maintaining the chromatic correction and final telescope optics which are used in collision. Reducing the IP divergence is also useful during the early stages of a run cycle, when the emittances in the beam delivery region may be too large to permit collisions at the design betatron functions. Finally, the beta-matching quadrupoles can be used to correct various mismatches between the design incoming beam and the beam which is actually delivered from the linac.

In the design of the NLC Beta Match, it has been assumed that coupling correction and measurement of the incoming beam phase space has been accomplished in the preceding Decoupling and Diagnostics section (Section 11.3), leaving only the problem of the uncoupled beam matching. The design constraints on the beta-matching region can be summarized as follows:

- The region should contain at least six quadrupoles, in order that the six independent first-order transfer matrix values to the IP can be independently adjusted, and that for a given input beam and set of Twiss parameters there is a solution which provides the desired output conditions and phase advances.

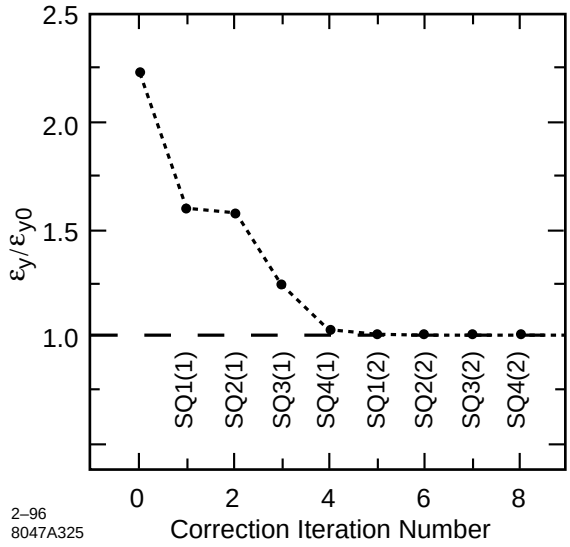


Figure 11-4. *Skew tuning simulation results.*

- The quadrupole specifications are consistent with warm-iron devices which can be easily manufactured and do not introduce emittance growth due to wakefields. Tolerances on field quality and mechanical stability should be no tighter than those upon quadrupoles in the chromatic correction sections. In general, the magnet specifications should be consistent with the technology and techniques of the rest of the beam delivery section.
- The contribution of the beta-matching region to overall chromaticity should be small enough to be easily cancelled out by reasonable increases in chromatic correction sextupole strengths.
- The range of accessible solutions for beam conditions at the IP should be as large as possible. Independent control of magnification and waist position control in x and y is essential.
- A dispersion-free pre-image point of the IP should be provided in each final focus, at which the magnification and waist position can be verified. This is necessary because the resolution of the incoming beam phase space is rarely adequate to match to collision conditions without such pre-IP verification. In addition, high-resolution (rf) BPMs that are placed close to the pre-image point will allow to correct the beam-beam deflection scans for pulse-to-pulse trajectory jitter in the IP betatron phase. Similar jitter-correction techniques have proven very successful at the SLC [Raimondi 1995] and are also being tested at the FFTB.
- The system should be as short, simple, and robust as possible.

11.4.1 Optical Design

At this time, there are two linear collider final-focus systems in existence, each of which has a unique beta-matching system. The SLC Final Focus uses a cluster of quadrupoles to focus on an image of the IP in the center of the first chromatic correction bend (Figure 11-5). A wire scanner in the center of the bend is used for match verification and tuning. In this case, the image is an existing symmetry of the design, and no additional quads are necessary to match from the image onto the CCS proper. The Final Focus Test Beam has no such naturally-occurring image point, and the five quadrupoles of its beta-matching section are used to focus the beam onto a pair of wire scanners in between

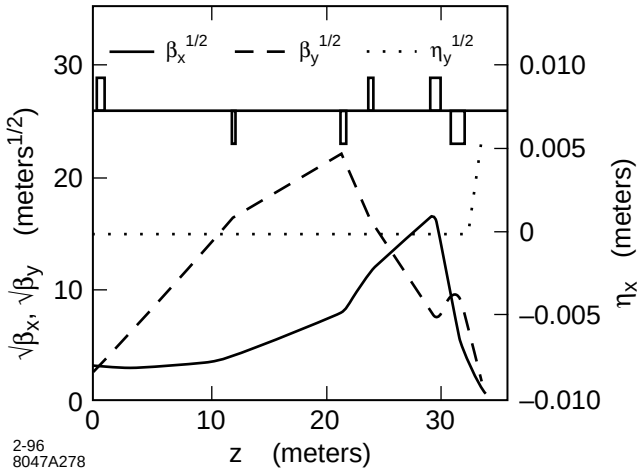


Figure 11-5. *Schematic representation of the SLC Final Focus Beta Matching region.*

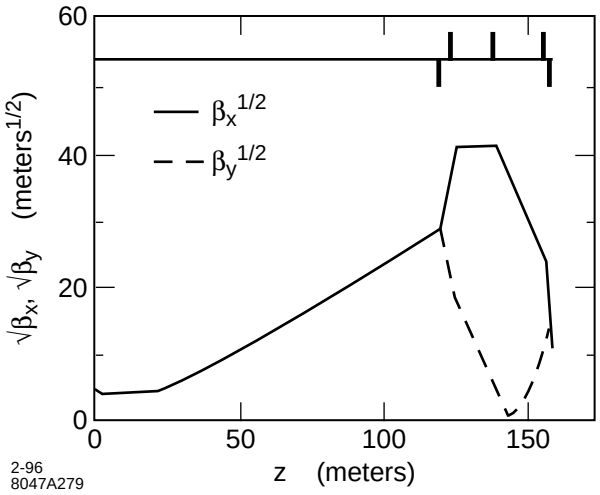


Figure 11-6. *Schematic representation of the Final Focus Test Beam Beta Matching region.*

the two chromatic correction sections (Figure 11-6). The SLC design is a more conventional mix of quadrupoles and modest drift regions, while the FFTB design precedes the first matching quadrupole with a 120-m drift space.

The current design of the NLC Final Focus includes several images of the IP, including a dispersion-free image in the beta exchange region (between the horizontal and vertical chromatic correction sections). The NLC Beta Matching section has been designed to resemble the FFTB system, including the use of the beta exchange image for tuning the match. This is for several reasons. First, the optics of the beta exchanger guarantees the presence of horizontal and vertical IP images which can be exploited for match verification, while use of an upstream image constrains the design of the dispersion matching section. In addition, the energy spread of the NLC is sufficient that chromatic correction of the beam at the IP image will be necessary during tuning. This can only be done downstream of the first chromatic correction section, by retuning the sextupoles in that section. As it is, only one plane can be corrected at a time at the

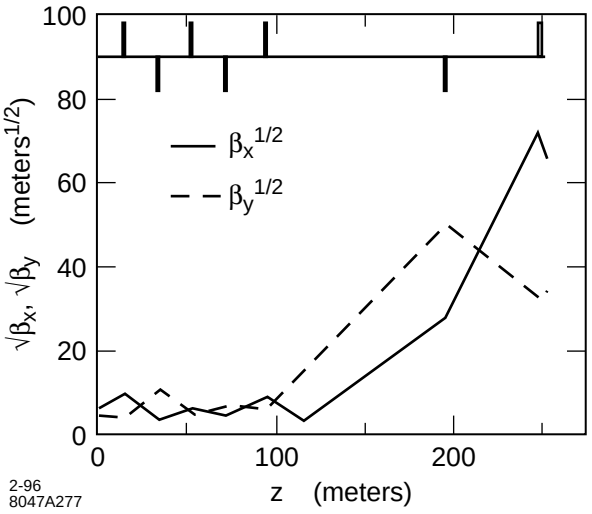


Figure 11-7. *Betatron functions of the beta-matching section.*

beta exchanger tune-up point; and the sextupole field required to correct the vertical chromaticity is opposite in sign to that used in normal running, necessitating a reversible power supply for the CCX sextupoles.

The optical functions of the NLC Beta Match are shown in Figure 11-7. The current design includes a total of seven quadrupoles, which satisfies the numerical requirement set forth above.

The total length of the system is 254 m. The maximal value of β_x within the section is 5,150 m, while the maximal value of β_y is 2,530 m. These values are far smaller than the typical values within the downstream region of the Final Focus, and the tolerances are expected to be correspondingly looser. Note that the beta-match region contains a drift of length 100-m upstream of the maximal betatron functions. Such a drift is an ideal location for a Single Beam Dumper (SBD), a kicker magnet capable of extracting individual bunch trains (*i.e.*, firing at 180 Hz). Extracting the beam after the Big Bend and Decoupling and Diagnostic sections would allow the beam to pass through the collimation section and the diagnostic wire scanner regions, while preventing it from entering the sensitive detector area. Thus, even beams which are not of sufficient quality to pass through the doublet into the detector can be maximally measured for diagnostic purposes if such an extraction magnet is placed at this location.

11.4.2 Tuning Elements

The tuning elements of the beta match include the seven normal quadrupoles in the region itself; the horizontal chromatic correction sextupoles (referred to as SX1 sextupoles); a beam size monitor at the IP image in the beta exchanger; and a divergence monitor just upstream of the vertical chromatic correction section.

Beta Matching Quadrupoles

Table 11-2 summarizes the design specifications of the seven quadrupole magnets in the beta-matching section. The magnets are physically identical with one another, with a length of 1 m and an aperture of 6 mm. The pole-tip fields for 1-TeV-c.m. go to a maximum of 6.5 kGauss for luminosity operations, and 7.3 kGauss for low-divergence operations

Quadrupole Name	Pole-tip Field (kGauss)	Length (m)	Aperture (mm)
QB1	5.9	1.0	6.0
QB2	-6.5	1.0	6.0
QB3	5.6	1.0	6.0
QB4	-3.2	1.0	6.0
QB5	6.4	1.0	6.0
QB6	-1.6	1.0	6.0
QB7	3.8	1.0	6.0

Table 11-2. *NLC beta-match quadrupole specifications.*

Conditions	$\sigma_{\rm x}$ at IP Image (μm)	$\sigma_{\rm y}$ at IP Image (μm)
Monochromatic Beam	1.85	0.20
Nominal Sexts, $\sigma_{\rm E}/{\rm E} = 0.003$	31.1	0.63
Sexts Off	14.8	0.41
SX1 = 33% design	2.00	0.48
SX1 = -130% design	74.1	0.20

Table 11-3. *Beam sizes at IP image location under varying conditions.*

$(\beta_x^* = 10\,\text{cm}, \beta_y^* = 1\,\text{cm})$; the same optics at the 1.5-TeV c.m. would require 10.9-kGauss pole-tip fields, and therefore these magnets are not usable at 1.5-TeV c.m. For all IP conditions in the operating plane, 6 mm is sufficient to maintain clearances in this region: worst-case rms beam sizes are 250 microns in the horizontal and 30 microns in the vertical.

The precise tunability of the beta-match design, in terms of range of IP betatron functions which can be delivered, range of waist knobs, and orthogonality of knobs, has not been studied for this design. A detailed study of the tunability should be carried out for final acceptance of this design.

The SX1 Sextupoles

Table 11-3 shows the beam size at the IP image location in the center of the beta exchanger with different sextupole configurations. In order to achieve the monochromatic beam size in the horizontal, the main CCX sextupoles (SX1) need to be set to a value roughly 33% of their design, while achieving the monochromatic vertical size requires a value 130% of design in the opposite polarity. This in turn indicates that the CCX sextupoles require reversible power supplies, with a capability to deliver significantly more current than is required for normal operations. The requirement of tuning the two planes with different sextupole strengths is also onerous in terms of time and difficulty level. While tuning the IP Image spot is the best tune-up procedure possible, next we will discuss a short-cut to this.

IP Image and Beam Size Monitor

The IP image at the center of the Beta Exchanger is a true image, with a pure demagnification in the horizontal and the vertical between it and the IP. The demagnifications are $(8.184)^{-1}$ in the horizontal, and $(57.65)^{-1}$ in the vertical.

The NLC operating plane indicates that beam sizes from 3.5 nm to 7.0 nm may be required at the IP, which in turn requires vertical beam sizes of 200 to 400 nm at the image. The horizontal beam sizes range from 226 nm to 320 nm at the IP, indicating sizes at the image from 1.85 microns to 2.6 microns. The horizontal beam sizes achievable at the image are within the expected range of laser wire technologies, while the vertical sizes are pushing the limits of these same technologies, and entering the range of Laser-Interferometer beam size monitors.

While the measurement of the horizontal corrected size in Table 11-3 presents relatively few problems, the vertical corrected size can only be achieved with a tremendous dilution of the horizontal spot, leading to an aspect ratio of 371:1. Measurements of the vertical beam size will therefore be tremendously sensitive to installation roll of the monitor: a roll of 1 mr will result in a contribution of 74 nm added in quadrature with the 200 nm of the focused vertical spot. For this reason as well as those outlined above, it may be impractical to use the IP image monitor to measure the size of the vertical beam; rather, the optimal use might be to use the monitor to constrain the vertical waist, and use the divergence monitor (see next Section) to verify the vertical betatron function at the IP image.

Angular Divergence Monitor

In addition to beam size measurements at an image of the IP, it is useful to be able to measure the horizontal and vertical angular divergences in this area. The divergences, along with the emittances measured upstream of the beta-match region, give the best estimate of the actual betatron function at the IP image monitor, as the divergence is insensitive to virtually all aberrations.

The ideal tactic for an angular divergence measurement is to measure the beam size at a point which is dominated by the divergence of the IP or IP image, and with a known transfer matrix to the IP or IP image. In this case, such a location is immediately upstream of the vertical chromatic correction section, at the end of the beta exchanger. The beam size at this location is 206 μm (x) by 140 μm (y) in the 1-TeV-c.m. design with 10 $\times$ 0.125-mm betatron functions at the IP and 50:1 emittance ratio; the beam size at this location is not drastically different for any optics in the operating plane, allowing a conventional metal-wire scanner such as the SLC linac scanners to be employed. However, the horizontal beam size listed here is significantly enlarged by dispersion present at this location. Therefore, an additional scanner with spectrum-measurement capability is needed to complete this measurement.

Because of the difficulties in measurement at each scanner, there are two conceivable schemes for tuning the incoming beam: one which "Caps the Ts and Dots the Is (CTDI)" and one which is "Close Enough for Government Work (CEGW)."

The CTDI tune-up is as follows:

- Measure incoming phase space with diagnostic wires (Section 11.3).
- Compute match optics for desired IP/IP Image conditions.
- Tune SX1 sextupoles to 33% design, tune x waist on image monitor.
- Measure beam on spectrum monitor and SY1 wire, compute monochromatic contribution at SY1 wire, compute x divergence.
- Use divergence computation to compute betatron function at IP image, correct with magnification knob, repeat waist for verification.
- Tune SX1 sextupoles to -130% design, tune waist on image monitor.
- Measure roll of IP image monitor spot (if possible), correct (if possible).

- Repeat waist scan on IP image monitor.
- Measure vertical divergence on SY1 monitor, compute IP image betatron function, correct if necessary with magnification knob.
- Repeat waist scan.
- Tune SX1 to colliding-beam strengths, go to IP.

The CEGW algorithm is the following:

- Measure incoming phase space with diagnostic wires (Section 11.3).
- Compute match optics for desired IP/IP Image conditions.
- Tune SX1 sextupoles to 33% design, tune x waist on image monitor; use waist scan to compute divergence.
- Use size of horizontal IP image spot and/or waist scan divergence value to tune magnification knob, iterate waist.
- Set SX1 to design strength.
- Tune y waist on IP image monitor to constrain waist location, ignore minimum spot size.
- Measure vertical beam size on SY1 monitor, compute vertical betatron function at IP Image monitor; tune magnification knob to correct SY1 size to predicted value.
- Iterate waist tuning if necessary, go to IP.

The CEGW algorithm makes maximal use of the diagnostic devices to bypass measurements which are difficult or time-consuming to make. A similar tuning algorithm in FFTB converges in a matter of hours; similar speed can be expected for tuning the beta match in NLC.

11.4.3 Tolerances

At this time, no detailed study of the beta-match section had been undertaken. However, based on the optics and some understanding of the functions of various magnets, it is possible to say a few things about operational tolerances of the system.

Field Strength Tolerances

The absolute accuracy tolerances on the beta-match quadrupole strengths are quite forgiving. The quadrupoles are set by matching the measured incoming beam to the desired IP conditions. Because the incoming phase space measurements are unlikely to have a greater precision than 1%, absolute quad strength accuracy of 0.1% will cause beam mismatch to be dominated by the errors in measurement, rather than errors in the magnets. The final settings of the magnets will be determined by tuning the IP image and/or the IP itself. This may require a granularity of the field strengths at the level of 0.01% of the maximum (assuming a maximum of 9.5-kGauss pole-tip fields). Stability of the fields with time will need to be approximately the same as the requisite granularity of the field.

Multipole Content Tolerances

Because the beam is much smaller in the beta-match quadrupoles than it is in other final-focus magnets, tolerances on higher multipole content will be correspondingly looser.

Vibration Tolerances

The beta-match quadrupoles are upstream of the first sextupole of the chromatic correction sections, and therefore these magnets are unable to induce steering between the sextupoles of a section. Tolerances for vibration will therefore be determined by dispersion and steering at the IP.

11.5 Chromatic Correction and Final Transformer

This beam line extends from the beta-matching section to the IP, and it comprises horizontal chromatic correction, beta-exchanger, vertical chromatic correction, and final transformer. Located at the end of the final transformer is the final doublet, which provides the last focusing before the beams collide (see Section 11.6 for a detailed discussion of the final doublet). The chromaticity of the final doublet is similar to that at the FFTB, and about five times as large as in the SLC. If the chromaticity were not corrected, the vertical IP spot size would increase by about a factor of 100 from the design value. The chromatic correction is accomplished in two separate beam-line sections, CCX and CCY, each of which accommodates a pair of sextupoles separated by an optical transform ($-I$) that cancels geometric aberrations and second-order dispersion. The chromatic correction is very similar to the FFTB, but it is different from the SLC in which the two sextupole pairs are interleaved. An interleaved placement of sextupoles is not acceptable for the NLC, because of the large higher-order aberrations this would generate. The bending section upstream of the beta-matching section is an innovation, unfamiliar from either SLC or FFTB. It allows the final focus to operate for any beam energy between 175 GeV and 750 GeV, without change of IP position, and hence it is known as "geometry-adjustment" section.

To obtain a proper estimate of the NLC luminosity, it is important to account for all possible sources of spot-size dilutions. To limit the total dilution, a budget was established for each effect contributing to the IP spot size, and, in particular, for each low-order aberration originating between the linac and the interaction point. Within this section, we present these dilution budgets, and discuss the implied tolerances on element strengths and positions, the frequency and accuracy of aberration tuning, and also the maintenance systems which stabilize the final focus between tunings. We believe that all the requirements can be met, and that there is even a potential for a further increase of the luminosity beyond that estimated here.

11.5.1 Introduction, Parameters and Dilutions

Final-focus and IP-beam parameters for different NLC scenarios are listed in Table 11-4. These parameters, in particular the cases Ia and IIa, form the platform on which the following discussion will be based. As indicated by the table, the final-focus designs for 500 GeV and 1 TeV deliver the desired horizontal and vertical spot sizes of about 250 nm and 4–8 nm, respectively. Both nominal luminosity and spot sizes quoted include various sources of dilution which are summarized in Table 11-5. The total and subtotals in Table 11-5 are calculated according to the formulae

$$1 - d_{subtot_j} \quad = \quad \Pi_i \, \mathrm{in}_j \, (1 - d_i) \tag{11.6}$$

$$1 - d_{tot} \quad = \quad \Pi_j \left(1 - d_{subtot_j}\right) \tag{11.7}$$

where d_i denotes the different relative luminosity dilutions.

To further understanding of Table 11-5: some supplementary comments on each table subsection are given below:

- **Emittance Growth.**
  The phase of the beam-line modules with respect to the IP must be very well-regulated and tuned for the beam line to function as designed. This can be easily monitored and arranged by studying betatron oscillations through the system. As a result of this phase regulation, elements within different beam-line modules affect the beam distribution in a way that is phase-related. What is of concern is the axis of the vertical (or horizontal) emittance in the IP phase. This we have denoted in Chapter 9 by the symbol $r_{y,IP}$. (If the beam distribution is looked at in a normalized phase space, where the beta functions are factored out, the distribution is normally assumed to be round, of radius $r_y = \sqrt{\epsilon_y}$. What we want to emphasize here is that the distribution will not remain round. The collimators in the FD phase, for example, give a kick in the IP phase, and enlarge the distribution along that axis only.) Emittance growth is not only generated by wakefields but also by synchrotron radiation, chromaticity, and higher-order optics. All of the entries in this subsection of the table refer to the enlargement of the IP axis of the emittance.

- **Collision.**
  The jitter is assigned according to the jitter budget proposed in Chapter 9 The calculation of the luminosity loss assumes that for the same luminosity loss the vertical missteering can be twice as large than for a rigid Gaussian bunch by virtue of the disruption. This is well-documented from simulations even for bunch charges that are a factor of 2 smaller than the design.

  Crab-cavity phase jitter (which is jitter of the positron beam cavity with respect to the electron-beam cavity) introduces an additional component to the horizontal jitter. So do field variations of the bending magnets, and quadrupole vibrations.

- **Tuned Low-order Aberrations.**
  In particular five aberrations—waist (x and y), dispersion (x and y) and skew coupling—must be scanned and corrected in regular intervals. (We do not assume a sublime dither technique.) Other aberrations will be more stable and can be scanned less frequently.

  Based on SLC experience, we expect to tune aberrations so that only a residual 0.5% luminosity loss per aberration remains. This number could conceivably be improved by developing a sensitive luminosity monitor for beam tuning purposes. In Table 11-5, the existence of such a monitor has been assumed neither for current nor for possible columns.

  We have assumed that the tuning of the five major aberrations is performed every 15 min, and that without tuning the design stability tolerances result in a 2% additional luminosity loss per aberration after 1 h. We expect that the aberrations increase from the minimum setting tolerance (0.5% per aberration) directly after tuning to the minimum plus the product of (time between tunings/1 h) and the budgeted spot-size increase for that aberration (as determined by system stability tolerances; typically 1–2%), added quadratically since they are independent.

  For the δ-dependent aberrations (chromaticity, dispersion, and chromatic skew), we presume that a smaller minimum (namely 0.25%) can be achieved by doubling the bunch energy spread for these scans.

  Dispersion has an effect on beam size due to energy spread within the bunch (an effect which is included under table subsection Tuned Aberrations) and also on beam jitter, since the individual bunches in a train and train-to-train will have different energies (an effect which is considered in the table subsection Collision). Insofar as the dispersion arises in the beam-delivery system, these two dispersions will be identical, and so the dispersion changes leading to increased beam size may be monitored by observing the correlation of position-jitter at the IP (which is being measured for each bunch train) with the energy of the bunch train (as can be measured

for each train by observing the orbit in the final-focus system). Thus the dispersion aberration can be tuned almost continuously and does not require allocation of a growth due to time between scannings. This potential improvement has not been included in Table 11-5.

It should be noted that dispersion can be present in the linac, and since inter-bunch energy spread and intra-bunch energy spread have different sources, and arise in different locations in the linac, it may occur that the beam size and the jitter due to dispersion cannot both be tuned together. In this case, one could try to tune by introducing orbit changes in the linac. However, this effect is considered an emittance dilution that originates in the linac, and not in the beam-delivery system, and thus we have not added it to our table.

Note that we equally have not added any aberration-induced dilutions originating in the collimation system. The reason is that such dilutions can be monitored non-invasively in the skew-correction section and can be corrected on a more frequent basis.

When an aberration affects both the vertical and the horizontal spot size, only the vertical effect needs to be considered, because, when the latter is tuned, the horizontal effect will be negligible as a result of the 100/1 spot-size ratio.

The crab-cavity adjustment refers to the voltage stability of the crab cavity. A 2% effect corresponds to a 6% voltage stability. We suppose that it can be tuned to 0.5% using a voltage scan, and then that it will be stable to 0.5% between tunings.

- **Other.**
  The entry for e^+e^- arrival time assumes an rms timing difference of the two beams equal to 170 fs (corresponding to 0.2° S-Band or 50 μm). This implies that the two bunches will be $0.2\beta_y^\star$ offset when colliding at the IP.

  Under table subsection Other, we have also listed the time devoted to beam-based alignment and tuning. Losses from tuning scans are calculated under the assumption that the five major aberrations are scanned every 15 min, the control system is designed to minimize the control-time overhead, the magnet settle time is 200 ms per field change, an aberration scan consists of seven steps with different magnet settings, and that at each step a beam-beam deflection scan over 50 pulses is performed.

  The time-out for beam-based alignment presumes the alignment to be completed in 24 h and to be executed every three months.

Table 11-5 lists two budgets of luminosity dilutions. The first refers to the Current Lattice and is based on tracking; the second is called Possible and assumes:

1. A reduced $\beta_x^\star$ which compensates the increase of the horizontal spot size due to design and tuning aberrations. This is consistent with the fact that the horizontal design spot size was chosen to limit the number of beamstrahlung-photons per electron radiated during collision. We do not suppose that the collision effect can be offset in this way, since it is not related to the beam size.
2. A redesign of the final-focus system which diminishes the effect of synchrotron radiation from bending magnets and final quadrupoles, We have learned during the design process, and we have reason to believe, that we can improve the performance of these modules. In many cases we have actually demonstrated this improvement.
3. Better control and adjustment of crab-cavity phase and voltage.
4. An improvement in the chromatic properties of big bend and IP switch.

The dilutions for the nominal lattice amount to a 46.6% overall loss in luminosity, which can possibly be reduced to a total of only 25.9% by the aforementioned means.

In Table 11-4, we have considered the nominal lattice, but have attributed a 13% spot-size increase to residual uncorrected vertical aberrations which is about twice the value listed in Table 11-5 and have assumed that the increase of the horizontal spot size due to residual tuned aberrations and horizontal steering (though not that due to horizontal emittance growth!) is compensated by a reduced $\beta_x^\star$. The expected increase of the spot-size product $\sigma_x \times \sigma_y$ then amounts to 50%, for 1-TeV-c.m. energy, and translates into a total luminosity loss of about 33%, which lies between the Current and the Possible value of Table 11-5.

The momentum bandwidth of the NLC final focus, defined by a 10% increase of either spot size for a mono-energetic beam, is at least $\pm 0.6\%$ for both c.m. energies (500 GeV and 1-TeV c.m.), even assuming the most difficult parameter set of the ZDR operating plane. The final focus is very forgiving in regard to increased emittances or to orbit variations: If β_y^* is held constant, the luminosity decreases roughly as the square root of the emittance, while the momentum bandwidth remains unaffected by the larger divergence. A 0.5-sigma variation of the incoming orbit causes an average spot-size increase by less than 1%.

The length of the final-focus system, from the geometry-adjustment section to the IP, is about 1800 m. This length is independent of energy, and is determined by the requirements for operation at 1.5-TeV-c.m. energy. Optimization of the final-focus system calls for a different dispersion and thus for a different bending angle at different c.m. energies. The different bending angle helps to balance nonlinear aberrations and synchrotron radiation effects at each energy. A dedicated geometry-adjustment section at the entrance to the final focus keeps the resulting geometry change at an acceptable level and the IP position constant. The horizontal magnet displacements required during an energy upgrade from 350 GeV to 1.5 TeV do not exceed 45 cm. The following discussion mainly refers to the two final-focus systems at 500 GeV and 1-TeV-c.m. energy with equal geometry, *i.e.,* Version II in Figure 11-10.

11.5.2 Layout and Optics

General Description

The basic layout of the proposed NLC final-focus system is very similar to that of the Final Focus Test Beam. In the region between the diagnostics section (DS) and the IP, the NLC final focus is constructed from six functional modules. These are, in the order of their location: geometry-adjustment section (GAS), beta- and phase-matching section (BMS), horizontal chromatic correction section (CCX), beta-exchanger (BX), vertical chromatic correction section (CCY), and final transformer (FT). Two conventional magnets, one superconducting and one permanent quadrupole at the end of the FT (the final doublet, or, perhaps more appropriately, the final quartet) provide the last focusing before the two beams collide.

The total distance from the entrance of the GAS to the IP is about 1820 m. A schematic of the magnet configuration is depicted in Figure 11-8, which also shows the beta functions corresponding to the 500-GeV parameter set listed in Table 11-1 (Case Ia in Table 11-4). In this case, the vertical beta function has a maximum value of 190 km at the main sextupoles in the CCY and peaks at about 75 km in the final doublet. The maximum value of the horizontal beta function is 75 km, at the CCX-sextupoles. A large beta function at the sextupoles is advantageous for chromatic correction, considering the effects of synchrotron radiation, chromo-geometric aberrations, and orbit-stability tolerances [Zimmermann 1995]. The length of the CCX is about two-thirds that of the CCY, which reflects the larger horizontal beta function at the IP. Note that, per a recent design modification, the beta-matching section has been augmented by two additional quadrupoles, to allow matching of the betatron phase advance between the collimator section and the IP. These two new quadrupoles and the accompanying minor optics change of the BMS are not included in the following discussion.

	c.m. energy	0.5 TeV			1 TeV		
		Ia	Ib	Ic	IIa	IIb	IIc
Luminosity w. dilution	L (10^{34} cm^{-2} s^{-1})	0.41	0.40	0.41	0.77	0.82	0.73
Lum. w. dil. & pinch	L (10^{34} cm^{-2} s^{-1})	0.58	0.55	0.60	1.02	1.10	1.06
Luminosity w/o dilution	L (10^{34} cm^{-2} s^{-1})	0.52	0.49	0.51	1.16	1.20	1.02
Lum. w. pinch w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.75	0.71	0.78	1.63	1.72	1.56
# particles per bunch	N_b (10^{10})	0.65	0.75	0.85	0.95	1.1	1.25
# bunches	n_b		90			90	
# bunch trains per s	f		180			120	
Enhancement factor	H_D	1.41	1.40	1.48	1.33	1.35	1.46
Hor. spot size w. dil.	σ_x (nm)	264	<294	<294	231	<250	<284
Vert. spot size w. dil.	σ_y (nm)	5.07	<6.26	<7.76	4.35	<5.08	<6.52
Hor. spot size w/o dil.	σ_x (nm)	253	283	283	200	219	253
Vert. spot size w/o dil.	σ_y (nm)	4.18	5.20	6.48	3.35	3.97	5.20
Hor. IP beta function	β_x^* (mm)	8	10	10	10	12	16
Vert. IP beta function	β_y^* (mm)	0.125	0.15	0.2	0.125	0.15	0.2
Norm. hor. emittance	$\gamma\epsilon_x$ (10^{-8} mr)		400			400	
Norm. vert. emittance	$\gamma\epsilon_y$ (10^{-8} mr)	7	9	10.5	9	10.5	13.5
Hor. emittance growth	$\Delta(\gamma\epsilon_x)_{SR}$ (10^{-8} mr)		~ 23			~ 38	
Hor. IP divergence	$\sigma_{x'}$ (μrad)	31.6	28.3	28.3	20.0	18.3	15.8
Vert. IP divergence	$\sigma_{y'}$ (μrad)	33.5	34.6	32.4	26.8	26.5	26.0
Bunch length	σ_z (μm)	100	125	150	125	150	150
Crossing angle	θ_c (mr)			20			
Rms energy spread	$\delta_{\rm rms}$ (%)			≥ 0.3			
Total energy bandwidth	$\Delta E/E$ (%)		≥ 1.2			≥ 1.2	
Free length from IP	l^* (m)			2			

Table 11-4. *Interaction-point beam parameters for different NLC scenarios. Spot-size increases are scaled from NLC-Ia and NLC-IIa. Luminosity enhancement is estimated from P. Chen's scaling laws. Nominal luminosity and nominal spot sizes include dilutions due to synchrotron radiation (0–16%), high-order aberrations (7%), residual uncorrected low-order aberrations (here assumed as 13%), e^+e^- timing offset (2%), crab crossing error (2%) and vertical orbit jitter (3%); see also Tables 11-5 and 11-9. Effect of residual low-order aberrations and horizontal orbit jitter on the horizontal spot size is assumed to be compensated by a reduced β_x^*. The numbers quoted for luminosity and spot size "without dilution" (w/o) are calculated according to linear optics and do not include any of the additional effects.*

Category	Description	Current Lattice		Possible	
		Horizontal	Vertical	Horizontal	Vertical
Emittance Growth (*actually* r_{IP})		**20.6%**	**11.8%**	**—**	**8.1%**
	Collimation System	3.0%	2.0%	—	2.0%
	IP Switch, Big Bend, SCS	3.0%	1.6%	—	1.0%
	Final Focus	15.4%	7.3%	—	4.0%
	Uncorrected Aberrations	6.0%	1.5%	—	1.5%
	Synchrotron Radiation	10.0%	5.9%	—	2.5%
	In Dipoles	10.0%	2.1%	—	0.1%
	In Quadrupoles	0.0%	3.5%	—	2.0%
	In Solenoid	0.0%	0.4%	—	0.4%
	Wakes (res.-wall & geom.)	0.2%	1.34%	—	1.34%
	Protection Collimators	0.1%	0.5%	—	0.5%
	Main Collimators	0.1%	0.65%	—	0.65%
	Misalignment	0.0%	0.25%	—	0.25%
	Jitter Wake	0.1%	0.3%	—	0.3%
	Quadrupole Wake	0.0%	0.1%	—	0.1%
	IR Beampipe	0.0%	0.2%	—	0.2%
Collision		**5.1%**	**3.5%**	**4.1%**	**3.5%**
	Incoming (to BDS) Jitter	0.5%	0.8%	0.5%	0.8%
	Jitter Amplification of Coll.	0.6%	0.7%	0.6%	0.7%
	Quad Motion	0.8%	1.5%	0.8%	1.5%
	Bend Power Supply Ripple	0.7%	0.0%	0.7%	0.0%
	Crab Cavity Phase-diff. Jitter	2.0%	0.0%	1.0%	0.0%
	Dispersion (inter-bunch & centroid)	0.5%	0.5%	0.5%	0.5%
Tuned Aberrations		**3.0%**	**5.5%**	**—**	**5.5%**
	Linear	2.5%	4.0%	—	4.0%
	Waists	1.0%	1.0%	—	1.0%
	Skew Quad (2)	0.0%	1.5%	—	1.5%
	Dispersion (intra-bunch δ)	0.5%	0.5%	—	0.5%
	Crab-angle Adjustment	0.5%	0.0%	—	0.0%
	δ-dependent 'linear'	0.5%	1.0%	—	1.0%
	Chromaticity	0.5%	0.5%	—	0.5%
	Chromatic Skew	0.0%	0.5%	—	0.5%
	Higher Order	0.5%	1.5%	—	1.5%
	Normal & Skew Sextupole	0.5%	1.5%	—	1.5%
Subtotal:		**26.9%**	**19.6%**	**4.1%**	**16.2%**
Other (*down time not included!*)		**7.0%**		**6.0%**	
	e^+-e^- arrival-time diff.	2.0%		1.0%	
	Losses from tuning scans	4.0%		4.0%	
	Time-out for bm-based alignment	1.0%		1.0%	
Total:		**46.6%**		**25.9%**	

Table 11-5. *Expected luminosity losses from dilutions in the beam delivery systems for 1-TeV-c.m. energy. See also Table 11-9 for a detailed account of low-order aberrations, and Table 11-8 for the effect of synchrotron radiation and higher-order aberrations.*

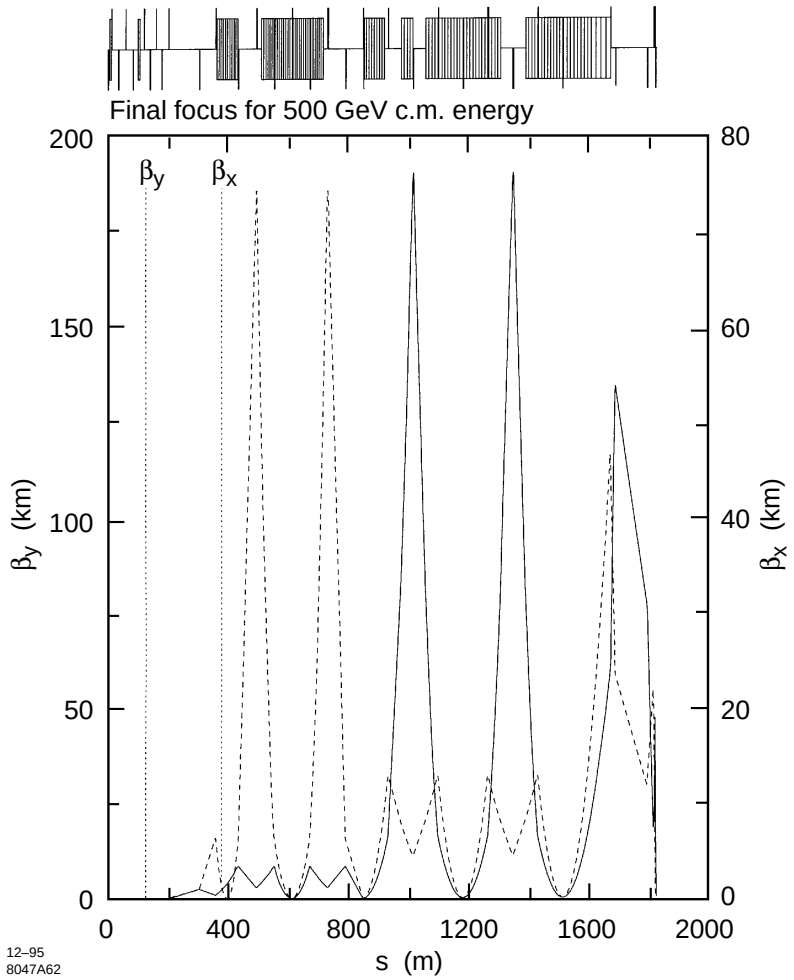


Figure 11-8. *Horizontal and vertical beta functions from BMS to IP, for the 500-GeV final focus.*

The optics at 1-TeV-c.m. energy is almost the same as that for 500 GeV. Again assuming the parameters of Table 11-4 (Case IIa), the peak values of horizontal and vertical beta functions at the CCX sextupoles are 60 km and 190 km, respectively.

The upper part of Figure 11-8 indicates that more than half of the final focus is occupied by about 100 bending magnets. These magnets generate the dispersion required for chromatic correction. Their maximum field at 1-TeV-c.m. energy is only 160 G, in order to restrict the emittance growth due to synchrotron radiation. The length of the entire system, the maximum beta functions, and the maximum dispersion (hence the bending angles) were optimized for the original design parameters, not only with regard to the effect of synchrotron radiation, but also with regard to nonlinear aberrations, magnet-vibration and field-ripple tolerances. The optimization procedure is discussed in the next section, and in [Zimmermann 1995].

The bending in CCX and CCY is in opposite directions, so that the final focus exhibits an “S”-shape geometry. For a c.m. energy of 1 TeV, the accumulated absolute bending angle, $\sum_i |\theta_i|$, is 6.3 mr, the bending angle inside the CCY alone is 0.9 mr, and, due to the `S' -geometry, the total net angle of the final focus is small: about 0.6 mr. At 350 GeV, due to shifted relative importance of aberrations and synchrotron radiation, the bending angles and the dispersion must

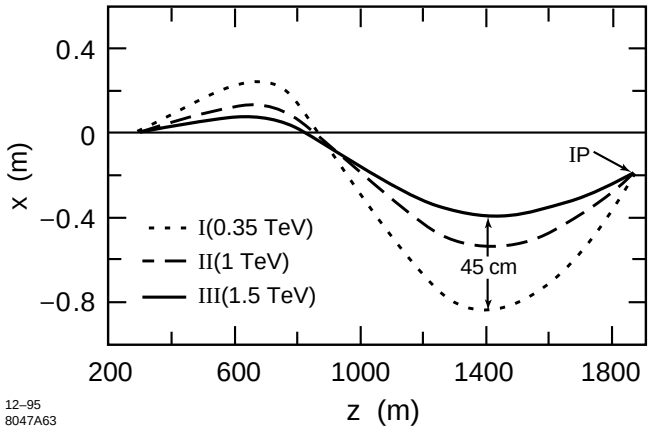


Figure 11-9. *Top view of the final-focus geometry. Magnet displacements by at most 45 cm are necessary during an upgrade from 500 (350) GeV to 1.5 TeV, while the IP-orbit angle changes by about 1.5 mr.*

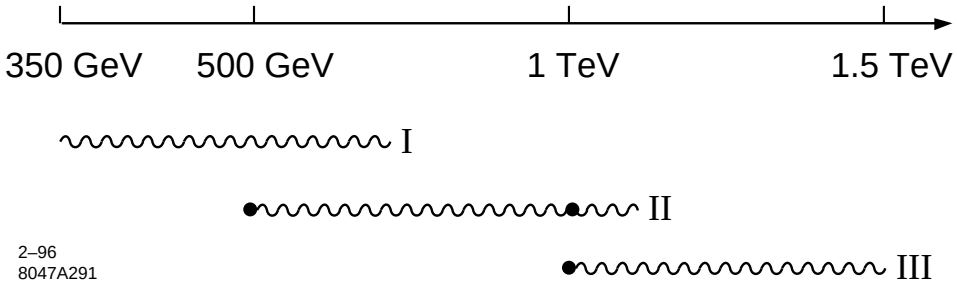


Figure 11-10. *Energy range of three different final-focus geometries. Dots show systems reported in the text.*

be increased by about 50%, with an accompanying change of geometry. The maximum horizontal displacement of magnets required for the upgrade from 350 GeV to 1.5 TeV is about 45 cm (Figure 11-9). The first geometry change occurs at about 500 GeV, the second around 1.1–1.2 TeV. The energy range covered by the three final-focus geometries is illustrated in Figure 11-10. The final-focus design for 1 TeV, including the final doublet, can operate in the entire energy range from 500 GeV to 1 TeV, and possibly beyond. Between 350 GeV and 1.5 TeV, the orbit angle at the IP varies by about 1.5 mr, while the IP position itself is independent of energy. A constant IP position is desirable, since the detector needs to be strongly coupled to the ground, to preserve the coherence of magnet motion caused by ground waves.

To protect the shielding masks in the final doublet region from hard synchrotron radiation, soft bending magnets of field strength 12 Gauss are placed over a distance of 64 m in front of the final transformer. The present soft-bend configuration was devised by S. Hertzbach and implemented by R. Helm. The soft bends deflect the beam orbit at the entrance to the final doublet (quartet) by about 8 mm horizontally, so that high-energetic photons generated in upstream bends and quadrupoles will not hit the inner bore of the last two final-doublet quadrupoles [Hertzbach 1995].

The 44 quadrupoles between the BMS and IP are typically 0.5-m long, and, for 1-TeV-c.m. energy, their pole-tip field is 3–5 kGauss. A few magnets, with larger apertures, require pole-tip fields of about 8 kGauss. The final focus also comprises between seven and 16 sextupoles which cancel chromatic aberrations. First, there are two conventional $-I$ sextupole pairs located in the chromatic correction sections. These are used to compensate the first-order chromaticity of the system. In addition, between three and 12 weaker sextupoles are interspersed in the GAS, CCX, BX, CCY, and

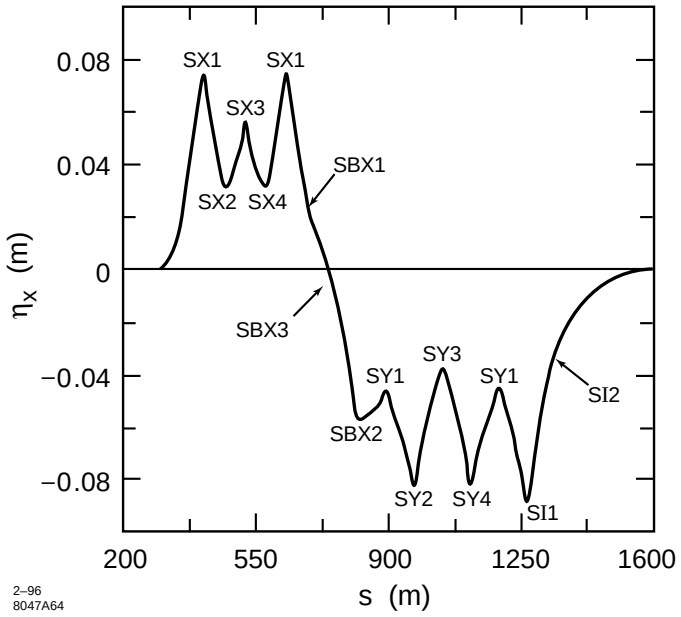


Figure 11-11. *Sextupole locations and dispersion for the 1-TeV design.*

FT; all at positions with nonzero dispersion. The sextupole locations and the dispersion function for the 1-TeV final focus are illustrated in Figure 11-11.

The sextupole strengths are determined from tracking to optimize the momentum bandwidth of the system. This application of sextupoles for bandwidth-optimization was first proposed by Brinkmann at DESY [Brinkmann 1990]. Similar to Brinkmann's early results, the momentum bandwidth of the NLC final focus is at least doubled by means of the additional sextupoles. This beneficial effect of Brinkmann-sextupoles is explained by a reduced chromatic breakdown of the $-I$ sections between the main sextupoles and also of the FT: a Taylor-map analysis of the final-focus optics reveals a significant reduction of fifth-order chromo-geometric aberrations due to the additional sextupoles (Section 11.5.3).

Optimization

The length of the 1.5-TeV final-focus system was originally optimized with regard to nonlinear aberrations, such as third-order horizontal and vertical chromaticity, and chromo-geometric terms with generator $x'^2 y'^2 \delta$, and also with regard to the effect of synchrotron radiation in the bending magnets, octupole-like aberrations from long sextupoles, magnet vibration tolerances inside the CCY, and power-supply ripple. A general optimization procedure is described in [Zimmermann 1995], and is a modified version of an earlier proposal by Irwin [Irwin 1991]. Some specific side-constraints for the actual design are not included in this optimization scheme. In particular, the optimization assumes only one operating energy and detector backgrounds are not taken into account.

The final focus has been designed such that its total length is constant in the entire energy range between 350 GeV and 1.5 TeV. The length of the system is then determined by the 1.5-TeV case, and it may be interesting to compare the design at this energy with the shortest possible, or optimum, final focus. Table 11-6 lists the theoretical optimum and the actual design values of vertical beta function and dispersion at the Y-sextupoles, sextupole strength, and length.

Parameter	1.5 TeV	
	Design	Theor. optimum
β_D^y (km)	300	80
η_D (mm)	30	24
k_D (m^{-2})	2.2	8.3
Δx (nm)	300	300
$\Delta k/k$	5×10^{-5}	8×10^{-5}
$L_{\rm tot}$ (m)	1800	900

Table 11-6. *Comparison of the 1.5-TeV final focus with a hypothetical design optimized for minimum length, assuming $\beta_x^* = 10$ mm, $\beta_y^* = 125$ μm, and $\delta \approx 3 \times 10^{-3}$. The length L_{tot} denotes the total distance from the start of the final focus to the IP, $\Delta k/k$ is the maximum field ripple of quadrupoles in the CCY, and Δx the tolerable orbit drift at the second Y-sextupole due to perturbations internal to the CCY.*

The optimization assumes that the maximum tolerable horizontal orbit variation Δx at the second Y-sextupole caused by perturbations internal to the CCY is the same as for the actual design, *i.e.*, 300 nm.

The length of the 1.5-TeV final focus appears to be about a factor of 2 larger than the theoretical optimum. Not included in the optimization, however, are the geometry-adjustment section in front of the CCX, which holds the IP position constant at all energies, the soft-bend section, and the long final transformer which proved to be essential for reducing detector background due to synchrotron radiation. The geometry-adjustment section and final transformer in the 1.5-TeV design are about 230-m and 500-m long, respectively, and account for most of the additional length, compared with the theoretical optimum.

Apertures and Beam Stay-Clear

The collimation depth in the postlinac collimation section will be about 7 σ_x, 35 σ_y and $\Delta E/E \approx 4\%$. There are also four horizontal and vertical collimators in the final-focus system, close to the main sextupoles, with a slightly larger collimation depth. The beam stay-clear in the final-focus system has to exceed this collimation depth. A side-constraint is that there should be as few as possible variations of the vacuum-chamber dimension, since these generate geometric wakefields. The beam-pipe radius r of the present design fulfills

$$r^2 \geq \max\left[(20\sigma_{\beta,x})^2 + (\eta_x \times 0.04)^2\ , (45\sigma_{\beta,y})^2\right] \quad . \tag{11.8}$$

The contribution from the 4% energy spread is small compared with the 20-$\sigma_{\beta,x}$ betatron beam size. Figure 11-12 illustrates the variation of the 20 σ horizontal and 45 σ vertical beam envelopes, right-hand side of Eq. 11.8, along the 500-GeV (c.m.) final focus, assuming normalized emittances $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m, which are 10%–20% larger than the design value. The proposed beam-pipe radius is also depicted. In most of the system the radius is constant, equal to 9.5 mm. Magnet apertures are assumed to be 1 mm larger, *i.e.*, about 10.5 mm.

The largest beam size is encountered around the CCX sextupoles. Here the beam-pipe radius needs to be increased to 18.5 mm. Larger apertures of 19–20 mm are then required for the two SX1 sextupoles, for four quadrupoles (QE1, QX3A, QX3B, and QBX1) and for nine B2 bending magnets in the CCX. A larger beam-pipe radius, 14 mm, is also necessary in the final transformer, and in the adjacent soft bending magnets. In the figure, the beam-pipe radius was chosen as 14 mm for all soft bends.

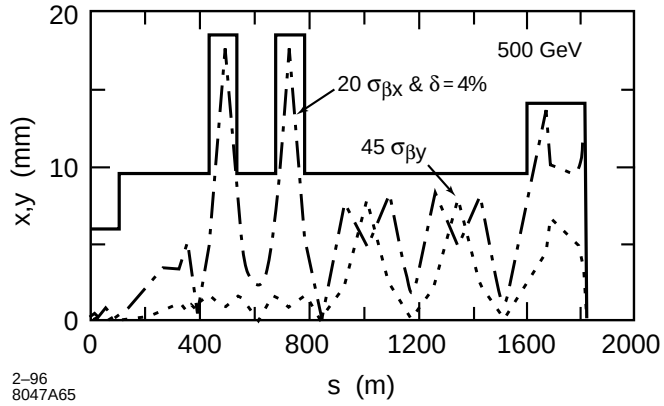


Figure 11-12. *Horizontal and vertical beam envelopes ($20\sigma_x$ and $45\sigma_y$) in the final-focus system for 500-GeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

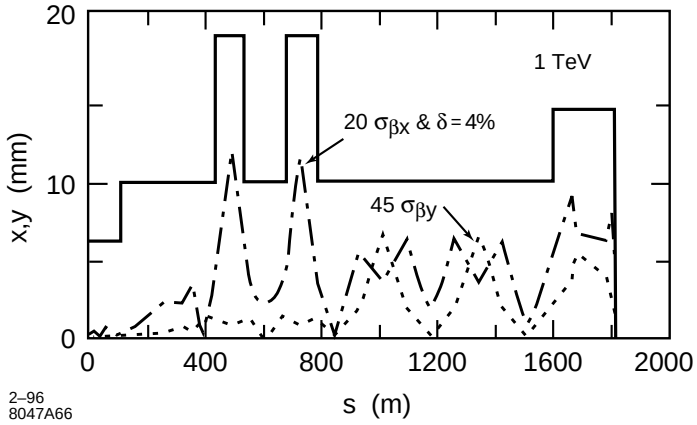


Figure 11-13. *Horizontal and vertical beam envelopes ($20\sigma_x$ and $45\sigma_y$) in the final-focus system for 1-TeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

In the 1-TeV final focus, the beam-pipe radii are still the same as for 500 GeV, and, because of the reduced beam sizes, the beam stay-clear becomes magnificent; see Figure 11-13.

An increased aperture could also be useful around the SY sextupoles, if the vertical emittance is much larger than the design value (for instance, during commissioning). In that case, the best luminosity would be achieved for constant vertical beta function (this is shown in Section 11.5.3), "Bunch Length," which implies larger IP divergence and larger aperture needs. Increased apertures in the CCY do make sense only, of course, if the aperture is not limited somewhere else, *i.e.*, in the final-doublet magnets.

The beam stay-clear at 500-GeV-c.m. energy for the 1-TeV final-doublet design is depicted in Figure 11-14. The beam-pipe radius is 11 mm inside the two conventional quadrupoles Q3 and Q2, as well as in the superconducting quadrupole Q1, and it decreases to 5 mm in the last, permanent-magnet quadrupole QA. The same doublet can operate in the entire energy range between 500-GeV and 1-TeV-c.m. energy. At 500 GeV the horizontal beam stay-clear in the final doublet is 23 σ_x; the vertical beam stay-clear is about 55 σ_y (see Figure 11-14). Again, at 1 TeV the beam

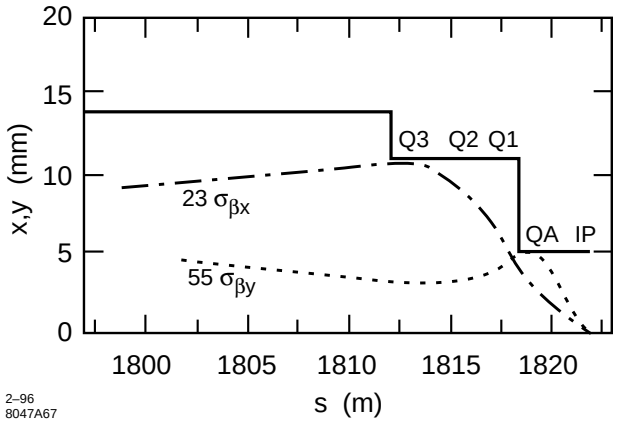


Figure 11-14. *Horizontal and vertical beam envelopes ($18\sigma_x$ and $45\sigma_y$) at the four quadrupoles of the final doublet, for 500-GeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

stay-clear of the doublet, in terms of beam size, is much larger than at 500 GeV. For a detailed discussion of the final doublet see Section 11.6.

To reduce background and to ease the collimation, the final-focus design will likely evolve towards even larger beam stay-clear. An overall vertical stay-clear of 55 σ_y at 500-GeV-c.m. energy is readily achieved by slightly increasing the sextupole apertures. Either by modifying the final doublet or by using a different doublet at 500 GeV and at 1 TeV , the vertical aperture at 500 GeV may be further increased to 60 σ_y.

In conclusion, the apertures of the final-focus design are very large, when viewed in terms of beam size and compared with the SLC (where the stay-clear is about $15\sigma_{x,y}$). A further increase of the vertical beam stay-clear from 55 σ_y to 60 σ_y would also not be difficult.

11.5.3 Performance

Beam Size at the Interaction Point

In this subsection, we discuss contributions to the spot-size from synchrotron radiation and from higher-order optical aberrations. The spot size, or luminosity, is further diluted by residual uncorrected low-order aberrations, which are described later in this section, by crab-crossing errors and by timing offsets between the two beams: A crab-crossing error arises when the phases of the two crab cavities vary with respect to each other. In that case the two beams do not collide head-on. The crab-cavity tolerances, which are discussed in Section 11.7.2, allow for a 2% luminosity loss from imperfect crab crossing. If there is a timing (or phase) error between the two bunch compressors, the collision occurs longitudinally offset from the IP waist, and the spot size at the collision point is increased by the effective waist shift. The proposed tolerance budget also assigns a 2% luminosity loss to this timing error (tolerances on bunch-compressor rf phases are discussed in Chapter 5).

The IP-spot-size increase due to higher-order chromatic aberrations and due to synchrotron radiation in bending magnets and quadrupoles is shown in Table 11-7, for the IP beam parameters of Table 11-1. Each entry in Table 11-7 was obtained by tracking 10,000 particles using the program DIMAD [Servranckx 1990]. Spot sizes for the 1.5-TeV

	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
	σ_x (nm)	σ_y (nm)	σ_x (nm)	σ_y (nm)	σ_x (nm)	σ_y (nm)
Linear	252	4.18	200	3.35	200	3.35
$\delta_{\rm rms} = 0.3\%$	264	4.32	212	3.40	208	3.55
+ s. rad. in bends	264	4.32	232	3.47	210	3.56
+ s. rad. in quads	264	4.33	232	3.59	211	3.71

Table 11-7. *RMS IP spot sizes for 500 GeV and 1 TeV, as obtained by tracking 10,000 particles with DIMAD.*

	500 GeV		1.0 TeV		1.0 TeV (1.5 TeV design)	
	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$
Chromo-geometric	3.5%	3.3%	**6.0%**	1.5%	4.0%	**6.0%**
Rad. in bends	0.0%	0.0%	**10.0%**	2.1%	1.0%	0.3%
Rad. in quads (tracking)	0.0%	0.2%	0.0%	**3.5%**	0.5%	**5.1%**
Total	**4%**	**3%**	**16%**	**7%**	**6%**	**11%**

Table 11-8. *Relative spot-size increases due to various effects, according to Table 11-7.*

final focus (Version III in Figure 11-10) operated at 1 TeV are also shown, in the far right column of the table. They are slightly better than those obtained with Version II at 1 TeV (second column from the right).

Table 11-8 compiles the relative spot size increases due to different sources as deduced from the above tracking results. At 500 GeV, the chromo-geometric blow-up is about 3.5% both horizontally and vertically. There is hardly any spot-size increase caused by synchrotron radiation.

At 1 TeV, the spot-size increase is considerably larger: 16% horizontally and 7% vertically for the 1-TeV design, and 6%/11% for the 1.5-TeV version operated at 1 TeV (Table 11-8). For the 1-TeV final focus, the largest contribution to the vertical spot-size arises from synchrotron radiation in the final doublet (Oide effect), which causes an rms spot-size increase by about 3.5%. In general, the Oide effect decreases with increasing length of QFT2/QFT3. In the proposed design, the total length of these two quadrupoles is about 4 m, and further improvement for greater length is negligible. (The Oide effect is discussed later in this section and in 11.6.4: "The Oide Effect Including Horizontal Motion".)

The horizontal blow-up at 1 TeV is primarily caused by radiation in the bending magnets (10%) and by chromo-geometric aberrations (6%), assuming a Gaussian momentum distribution of rms value 0.3%. The blow-up due to synchrotron radiation is proportional to the R_{16} matrix-element from the bending magnets to the IP. This blow-up could be reduced by either lowering the dispersion or by weakening the bending magnets. On the other hand, the chromo-geometric aberrations are enhanced by a reduced dispersion (and larger sextupole strength). Thus the present value of dispersion, at which the effects of synchrotron radiation and aberrations are roughly comparable, is about the optimum. One possibility to further reduce the residual horizontal spot-size increase due to radiation and aberrations would be to increase the length of the final-focus system.

The impact of dispersion on spot size and bandwidth can be understood by noting that most of the aberrations arise from the chromatic breakdown of the $-I$ between the main sextupoles. The aberrations affecting the vertical spot size

are then described by a Hamiltonian of the form

$$H \propto K_s^2 \eta_s^2 \left(y'^2 \delta^3 + 2\delta^2 x' y'^2 \frac{\beta_{x,s}^{\frac{1}{2}}}{\eta_s} + x'^2 y'^2 \delta \frac{\beta_{x,s}}{\eta_s^2} + \ldots \right) \quad , \tag{11.9}$$

where K_s is the sextupole strength, $\beta_{x,s}$ the beta function, and η_s the dispersion at the sextupole. The coordinates x', y' denote the normalized slopes at the IP (in units of $\sqrt{\mathrm{m}}$). The product $K_s \eta_s$ is a constant, determined by the doublet chromaticity. The Hamiltonian, Eq. 11.9, shows that most of the aberrations are reduced for increased values of η_s except for the first term which represents the third-order vertical chromaticity. Similarly, most of the important aberrations which influence the horizontal spot size are also smaller for increased dispersion. The third-order vertical chromaticity, which is independent of the dispersion, is canceled by the Brinkmann-sextupoles (Section 11.5.3, "Aberrations").

In summary, synchrotron radiation and high-order chromo-geometric aberrations lead to a total increase of the vertical rms spot size by about 3% and 7% at 500 GeV and 1 TeV, respectively. Since a large part of the blow-up at 1 TeV is due to the Oide effect, the actual luminosity loss will be smaller than quoted, because only few photons are radiated. This is elaborated in Section 11.6. Finally, the increase of the horizontal spot size due to synchrotron radiation could be compensated with a further reduction of β_x^*, but we have not made this assumption.

Momentum Bandwidth

A large momentum bandwidth of the final focus is very desirable considering beam dynamics in the linac, although a larger energy spread necessitates tighter tolerances in the final focus. The bandwidth of the final focus has been increased by introducing additional sextupoles, as suggested by Brinkmann. Examples of the achievable bandwidth are presented in Figure 11-15 and 11-16 for 500 GeV and 1 TeV, respectively, where in both cases three additional Brinkmann sextupoles have been used: one in each of CCX, CCY, and FT. A minimum of three sextupoles appears to be needed for achieving a decent bandwidth, but the further improvement in going from three to 12 Brinkmann-sextupoles is rather small. Therefore, the presently adopted design employs only three Brinkmann sextupoles.

Figure 11-15 shows the vertical and horizontal beam sizes at 500 GeV as a function of the momentum-offset Δ of a mono-energetic beam. The momentum bandwidth for a 10% blow-up of either spot size is $\pm 0.61\%$. As illustrated by Figure 11-16, at 1 TeV the bandwidth for a 10% spot-size increase is about the same. In both figures, the spot sizes are given in units of the values for zero momentum offset. If Brinkmann-sextupoles are not used, the momentum bandwidth is less than $\pm 0.35\%$.

Aberrations

In this subsection, we first discuss the importance of low-order aberrations. Although they can be tuned, these aberrations turn out to be significant for proper estimating the achievable luminosity. Towards the end of this subsection, we also describe the dominant higher-order aberrations.

The perhaps most important, yet often overlooked source of IP spot-size dilution are residual low-order aberrations. These are aberrations, such as waist shift, skew coupling, etc., for which tuning schemes exist, but which cannot be fully compensated, due to the finite accuracy of the beam-beam deflection scans, from which the IP beam size is inferred during tuning. In general, we allocate a 1–2% spot-size increase to each first-, second- and third-order aberration, including the spot-size increase occurring between two tuning scans. This appears to be a fairly conservative estimate of tuning accuracy and stability. For comparison, the actual tuning accuracy in the SLC is about 0.5% per aberration.

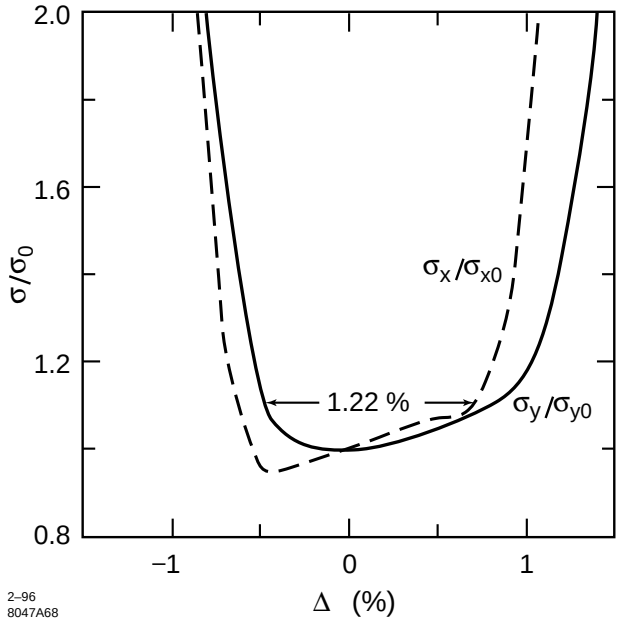


Figure 11-15. *Relative increase of spot sizes as a function of momentum-offset Δ for a monoenergetic beam at 500 GeV, using three Brinkmann-sextupoles. Effects of synchrotron radiation are not included.*

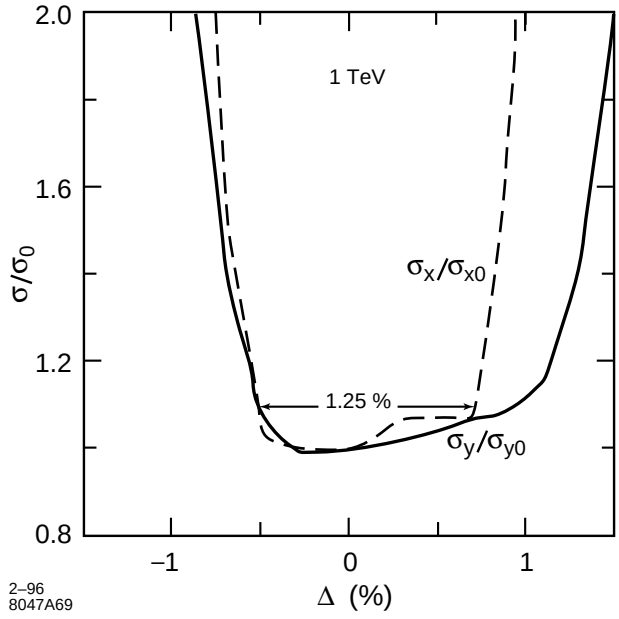


Figure 11-16. *Relative increase of spot sizes as a function of momentum-offset Δ for a monoenergetic beam at 1 TeV, using three Brinkmann-sextupoles. Effects of synchrotron radiation are not included.*

The aberration tuning in the NLC final focus can further be improved in a number of ways: A fast and precise luminosity monitor, *e.g.,* a detector for low-angle Bhabha scattering, will provide a more accurate luminosity measurement than the beam-beam deflection scans and will, therefore, allow a better fine-tuning of aberrations. For all energy-dependent aberrations, the tuning may also be bettered by temporarily increasing the energy spread. The horizontal blow-up will be compensated by reducing the horizontal IP beta function. In addition, if first- and second-order dispersion are continually monitored and corrected, using measured correlations between the beam position at the IP or at the pre-image point and the centroid energy, these two aberrations will become insignificant.

The effect of the low-order aberrations is summarized in Table 11-9 which lists all aberrations that will be tuned, including tuning procedure, tuning frequency, luminosity loss attributed to each aberration, and associated maintenance systems. The table shows that the total increase of the vertical spot size due to residual uncorrected tunable aberrations is expected to be about 13%. Thus, imperfectly-tuned vertical aberrations amount to about half the total vertical blow-up originating in the beam-delivery system; most of the rest being due to synchrotron radiation and to uncorrectable higher-order aberrations. Note that the assumed value for the average dilution per aberration is somewhat ambiguous, since it depends on tuning accuracy, tuning frequency and BPM stability, and can be improved in numerous ways. The assumptions in Table 11-9 differ from those made earlier in Table 11-5. The latter table was more optimistic and predicted a vertical blow-up due to tuned aberrations of only about 6%.

Optical aberrations of order 5 or higher (in the Hamiltonian) exist by design, and at present no tuning scheme is contemplated for any of these. Higher-order aberrations can be calculated from a Taylor map M representing the final-focus system to arbitrary order in (x, x', y, y', δ). Such a map is extracted from a standard MAD input file by methods of differential algebra [Berz 1989] (also called truncated power series algebra), using the program DESPOT [Forest] [3]. The Taylor map can be rewritten as a Dragt-Finn factorization, from which an approximate nonlinear Hamiltonian is determined. Formally we have

$$M = R\exp(-H_{\rm non}) + \text{ higher-order terms} \tag{11.10}$$

where the term R denotes the linear transfer matrix, and $H_{\rm non}$ represents the nonlinear Hamiltonian describing the aberrations. The spot-size increase caused by these aberrations is given by

$$\Delta\sigma^2_{x,y} \approx \left\langle \left(\frac{\partial H_{\rm non}}{\partial p_{x,y}} \right)^2 \right\rangle \quad . \tag{11.11}$$

The angular brackets indicate an average over the (linear) beam distribution at the IP, usually assumed as Gaussian.

An example of such an analysis is given in Table 11-10, which shows contributions to the spot size of some important fifth-order aberrations, to be added in quadrature. The largest aberration is the third-order vertical chromaticity (generator $y'^2\delta^3$), which in the absence of Brinkmann-sextupoles would cause a spot-size increase of 16–25%. In the optimized design, the effect of this aberration becomes negligible (see right column). There is still a considerable amount of third-order horizontal chromaticity, which seems to give rise to a spot-size increase of 5–10%. However, it is somewhat misleading to look at this term in isolation: to produce the desired optimum bandwidth, the third-order horizontal chromaticity has (empirically) been balanced against a nonzero first-order chromaticity which is of similar size and opposite sign. For the optimization of higher-order terms, evidently, a compromise had to be found between a large momentum bandwidth and a small chromo-geometric blow-up for a beam of certain finite momentum spread.

In summary, uncorrected lower-order aberrations cause a 6%–13% increase of the vertical spot size. The horizontal spot size may be recovered by reducing $\beta^\star_x$. The residual fifth-order chromo-geometric aberrations give rise to a vertical blow-up of 2%–10%.

[3] In the current installation, the MAD input first has to be converted into a DESPOT input with the help of the code TRACY.

Time	Name	Hamiltonian	Budget[e] (%)[a]	Tuning Knob	Diagnostic	Maint. System
1/5 s	Steering	y'	.5+.5	Corrector at FD	IP kick	FD anchor + seism. + steering
		x'	.5+.5 [b]			none
1 hr	Dispersion	$\delta y'$	N/A[d]	Weak SD Sexts. asym. hor. move	IP spot size scans	FF orbit control
		$\delta x'$	N/A [b,d]	Weak SF Sexts. asym. hor. move		
	Waist	y'^2	.5+.5	Weak SD Sexts. symm. hor. move		
		x'^2	1+1 [b]	Weak SF Sexts. symm. hor. move		
	Skew1	$x'y'$	1+1	Weak SD Sexts. asym. vert. move		
1 wk	Sextupole	$x'y'^2$	2	Sp. sext. in FT	IP spot size scans	Sext. strength & LT stability
		x'^3	2 [b]			
	Skew sextupole	y'^3	2	Sp. skew sexts. in FT		Sext roll, FD mult. & LT stability
		x'^2y'	2			
	Skew2	xy'	2	Sp. skew quad		LT stability & solenoid strength
	Chromaticity	$\delta y'^2$	1 [c]	SD sext strength		& sext. strength & roll
		$\delta x'^2$	1 [b,c]	SF sext strength		
	Chromatic Skew	$\delta x'y'$	1 [c]	CCY vert. disp.		Occasional FFS alignment studies
	2nd order disp.	$\delta^2 y'$	N/A [c,d]	QT6 and QT5 move		& LT stability
1 mo	Within range of aberration tuning knobs	All	N/A	Element movers	Beam-based alignment	FFS alignment studies & XLT stability
1 yr	Within range of bm based align. (capt. tol.)	N/A	N/A	Coarse element movers	Survey	XXLT stability

[a] .5+.5 means the budget is 1/2% for the final doublet and 1/2% for the remainder of the system; 1+1 means the budget is 1% for the final doublet and 1% for the remainder of the system; system being either the final focus (for steering) or the whole beam line from linac to IP (for all other aberrations).
[b] The pure x-aberrations can be compensated by a slight reduction of β_x^*.
[c] Chromatic aberrations can be enhanced by increasing the energy spread. We assume that these aberrations are tuned out with a doubled energy spread. The chromatic skew can be eliminated by canceling the vertical dispersion. The second-order dispersion at the IP could also be detected directly if there were a beam size monitor available at the IP.
[d] Dispersion and second-order dispersion can be continually corrected, using the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy of both beams.
[e] Perhaps, because they are stable, the aberrations in the one-week timescale could, under steady beam conditions, be better compensated using a luminosity monitor. We have not made this assumption.

Table 11-9. *Low-order aberration tuning and maintenance. This table allocates a 1% spot-size increase to each δ-dependent aberration, except for dispersion which is assumed to be perfectly compensated, and it assigns a 2% spot-size increase to most other aberrations. The total increase of the vertical spot size is 13%. Note that in Table 11-5 different assumptions about tuning accuracy, frequency and stability were made (0.5% increase per aberration including dispersion), which resulted in about half the vertical blow-up.*

$(\Delta\sigma_y^2)^{\frac{1}{2}}/\sigma_{y0}$	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
Generator	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.
$x'y'^2\delta^2$	0.31	0.21	0.27	0.29	0.52	0.63
$y'^2\delta^3$	0.52	0.02	0.58	0.03	0.78	0.12
$(\Delta\sigma_x^2)^{\frac{1}{2}}/\sigma_{x0}$	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
generator	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.
$x'^4\delta$	0.12	0.08	0.06	0.22	0.25	0.30
$x'^2\delta^3$	0.25	0.27	0.33	0.47	0.28	0.33

Table 11-10. *Effect of fifth-order nonlinear aberrations in the final-focus system, with and without Brinkmann-sextupoles for the parameters listed in Table 11-1 and $\delta_{\rm rms}$ =0.3%. The contribution to the spot size quoted has to be added in quadrature.*

Synchrotron Radiation

Besides the effect of lower- and higher-order aberrations, further increases of the spot size arise from synchrotron radiation in final doublet and bending magnets. These are the subject of this section. The Oide effect is also treated, much more comprehensively, in Section 11.6.4.

Synchrotron radiation in the final-doublet quadrupoles causes different particles to be focused at a different distance from the last quadrupole, and, thereby, gives rise to an ultimate limit on the achievable spot size. This has first been pointed out by K. Oide [Oide 1988], after whom it is called "Oide limit". Generalizing Oide's original formula for the spot-size increase [Oide 1988] to two planes (horizontal and vertical) and both quadrupoles of the final doublet, the increase of the vertical spot size due to synchrotron radiation in the last two quadrupoles is given by

$$\Delta\sigma_y^2 = \frac{55 r_e \lambda_e \gamma^5}{48\pi\sqrt{3}} \left[\int_0^{l_1+l_2} \left[\int_0^s g_y(s_1)^2 K(s_1) ds_1 \right]^2 \cdot \right.$$
$$\left. \cdot |K(s)|^3 < \left[|g_x(s)^2 x_0'^{*2} + g_y(s)^2 y_0'^{*}2| \right]^{\frac{3}{2}} y_0'^{*2} > ds \right] \quad , \tag{11.12}$$

where the angle-brackets denote an average over the bunch distribution, the variable s is the longitudinal position measured backward from the exit face of the last quadrupole, and $K(s)$ is the quadrupole strength at position s. The Green's functions g_x and g_y are depicted in Figure 11-17 for an approximation to the NLC final doublet.

The figure shows that most of the spot-size increase occurs in the second pair of quadrupoles viewed from the IP (QFT2 and QFT3), and is due to the focusing in the horizontal plane, rather than to the final vertical focusing. In general, increasing the length of QFT2 and QFT3, while keeping the integrated quadrupole field $K_2 \cdot l_2$ constant, reduces the effect of synchrotron radiation in the final doublet. The length chosen is close to the optimum, in the sense that further improvement for still greater length is negligible. With the present final-doublet design, the increase of the rms spot size due to the Oide effect is less than 0.1% at 500 GeV and about 3.5% at 1 TeV; compare Table 11-8. A modification of the final-doublet design may be required to reduce the blow-up at 1 TeV. The reader is referred to Section 11.6 for a more detailed discussion of the Oide effect.

Synchrotron radiation in the bending magnets affects the beam size at the IP in two different ways. First, radiation in dispersive regions increases the horizontal spot size because of a nonvanishing R_{16}-matrix element between dispersive bend regions and the IP. Synchrotron radiation in the bending magnets causes a horizontal spot-size increase by about 0% and 1–10% at 500 GeV and 1 TeV, respectively (Table 11-8). The blow-up at 1 TeV could be reduced by lengthening the system. Second, for particles losing energy in the final and central bending sections of the CCY,

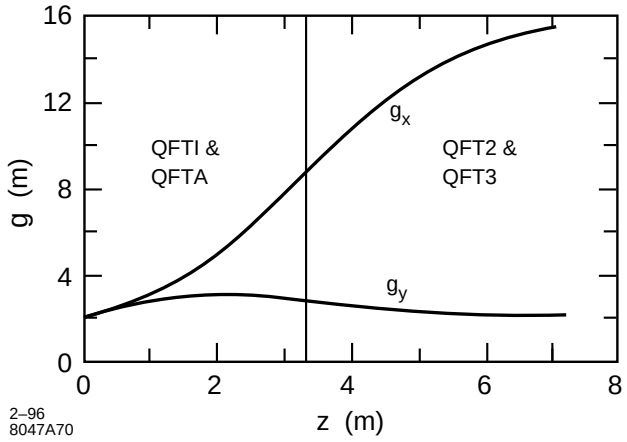


Figure 11-17. *Horizontal and vertical Green's functions as a function of longitudinal position s measured backward along the NLC final doublet. the zero point $s = 0$, refers to the exit face of the last quadrupole.*

	Energy	500 GeV	1 TeV	1 TeV[a]
Number of photons	N_γ	33	66	34
Critical energy	E_c (MeV) in bends	0.3	2.7	1.6
Critical energy	E_c (KeV) in soft bends	25	200	173
Average energy loss	$\Delta E_{\rm rad}$ (MeV)	3.3	52	17
Induced energy spread	$\Delta\delta_{\rm rms}$	0.6×10^{-5}	3.2×10^{-5}	1.2×10^{-5}
Emittance growth	$\Delta(\gamma\epsilon_x)$ (mr)	8×10^{-9}	4.1×10^{-7}	4.1×10^{-8}
Rel. emittance growth	$\Delta(\epsilon_x)/(\epsilon_{x\,0})$ (%)	0.2	9.6	1.0
Hor. spot-size increase	$\Delta\sigma_x/\sigma_{x0}$ (%) (tracking)	0.0	10.0	1.0
Vert. spot-size increase	$\Delta\sigma_y/\sigma_{y0}$ (%) (tracking)	0	2	0.3

[a] This column refers to the 1.5-TeV final focus operated at 1 TeV.

Table 11-11. *Characteristics of synchrotron radiation in the bending magnets.*

the vertical chromaticity of the final doublet is not or only partly compensated, and the particles are chromatically defocused, resulting in an increase of the vertical spot size. This second mechanism leads to a vertical blow-up by about 0.3–2% at 1 TeV (Table 11-8).

The spot-size increases due to synchrotron radiation in the final doublet and bending magnets have been listed in Table 11-8. Some related parameters such as the average energy loss $\Delta E_{\rm rad}$, the average number of photons emitted per electron N_γ, and the critical energy E_c in the bending magnets are given in Table 11-11.

We conclude that synchrotron radiation in bending magnets and quadrupoles is not important at 500 GeV and that at 1 TeV it reduces the luminosity by about 15%, assuming the same final-focus geometry as at 500 GeV (Version II in Figure 11-10). The luminosity decrease due to synchrotron radiation is less than half that, *i.e.*, 7% , if the 1.5-TeV final focus (Version III) is operated at 1 TeV.

Long-Sextupole Effect

A long sextupole magnet gives rise to an octupole-like fourth-order aberration (see, *e.g.*, [Irwin 1992]). For a certain strength, the maximum sextupole length causing a 2% spot-size increase is given by

$$l_{sext} \leq \frac{1.2}{\sqrt{3}\beta_x^s \beta_y^s \epsilon_{x,y} + \sqrt{15}\beta_{y,x}^{s\ 2}\epsilon_{y,x}} \frac{1}{k_{sext}^2} \tag{11.13}$$

where

$$k_{sext} \equiv \frac{2 l_{sext} B_T}{(B\rho)a^2} \tag{11.14}$$

denotes the integrated sextupole strength in units of m^{-2}, and the beta functions are those at the sextupole. Applying this formula to the 1-TeV final focus for normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m, we find the maximum tolerable sextupole lengths

$$l_{SX1} \leq 27\,\mathrm{m} \quad \text{and} \quad l_{SY1} \leq 23\,\mathrm{m} \quad , \tag{11.15}$$

For the same optics at 500 GeVand normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 10 \times 10^{-8}$ m, the limit on the sextupole length is

$$l_{SX1} \leq 9\,\mathrm{m} \quad \text{and} \quad l_{SY1} \leq 12\,\mathrm{m} \quad . \tag{11.16}$$

Thus, the fourth-order aberrations induced by the finite sextupole length are insignificant. Note that the 500-GeV limit is tighter, because both the emittances and the horizontal beta function are larger than at 1 TeV. By contrast, due to increased dispersion and lowered sextupole strength at 350 GeV, the maximum sextupole length for the 350-GeV optics will be larger than at 500 GeV.

At 500 GeV, the sextupole lengths may be chosen as 0.4 m, which, for the assumed magnet apertures of 20 mm and 10.5 mm, translates into a pole-tip field of 3.3 kGauss and 1.9 kGauss for SX1 and SY1, respectively. The upgrade to 1 TeV could be performed by adding more sextupoles, to increase the effective length of the two SX1 and SY1 magnets to 0.8 m. Alternatively, longer sextupoles could be installed from the beginning.

We conclude that the present design is far from the long-sextupole limit.

Bunch Length

It is interesting to study the dependence of luminosity on bunch length and vertical beta function. A larger bunch length would relax requirements on the bunch compressor or on the main linac, while an increase of the vertical beta function would reduce the spot-size blow-up due to aberrations and due to synchrotron radiation.

Figure 11-18 shows the luminosity as a function of bunch length, for three different values of β_y^* and a c.m. energy of 500 GeV. The luminosity is estimated from a scaling law for luminosity enhancement, derived by P. Chen [Chen 1993], assuming linear IP spot sizes. The figure illustrates that an increase of the bunch length from 125 μm to 140 μm would cause a luminosity loss by about 3%. A similar luminosity loss occurs if the vertical IP beta function is raised from 125 μm to 150 μm.

The conclusion here is that both bunch length and IP beta function can be changed over a wide parameter range with only marginal loss of luminosity.

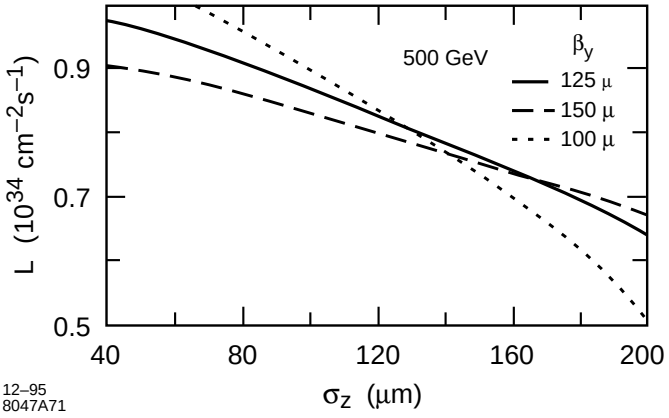


Figure 11-18. *Luminosity at 500-GeV-c.m. energy as a function of the bunch length, for three different values of β_y^*, 90 bunches with 0.65×10^{10} particles per bunch, and a repetition rate of 180 Hz; other parameters as listed in Table 11-1.*

IP Effect	Magnitude	Mover Symmetry	Mover Plane	$\Delta/\ \mu$m at SX1	$\Delta/\ \mu$m at SY1
Skew	$r_{32} = 0.20$	symmetric	y	0.85	0
x-waist	1.6 mm	symmetric	x	1.66	0
y-waist	25 μm	symmetric	x	0	0.28
x-dispersion	19 μm	asymmetric	x	4.30	0
y-dispersion	1.1 μm	asymmetric	y	1.07	0

Table 11-12. *IP tuning with CCX and CCY main-sextupole movers. The minimum displacement requirements, Δ, in the right two columns and the magnitude, in the second column, are for a 2% IP single spot size increase ($E_{\rm beam} = 250$ GeV, $\beta_x^* = 8$ mm, $\beta_y^* = 125\ \mu$m, $\gamma\varepsilon_x = 5 \times 10^{-6}$ m-rad, $\gamma\varepsilon_y = 5 \times 10^{-8}$ m-rad, $\sigma_\delta = 0.3\%$).*

11.5.4 Tuning

Control over IP horizontal and vertical dispersion, waist position, demagnification, chromaticity, and linear coupling, is provided by orthogonal magnet- or mover-knobs similar to those suggested by Irwin for the SLC [Walker 1993, Zimmermann 1995b] or the FFTB. Laser wires will measure the beam size at several locations in the final focus. Valuable experience will be gained from a first laser wire installed at the SLC [Ross 1994]. Decoupling, match of the incoming dispersion and emittance-measurements are performed upstream of the CCX (*e.g.*, in big bend, SCS and DS). The BMS-quadrupoles are used to correct a mismatch in phase space of the incoming beam and to adjust the demagnification at a pre-image point of the IP in the BX (the tuning on the pre-image point is described in Section 11.4), or at the IP itself.

Symmetric or asymmetric displacements of the main-sextupole pairs correct waist shifts, coupling and dispersion at the IP. A mover-based tuning is attractive, since a precise adjustment and control of sextupole positions is required regardless. A summary of main-sextupole mover requirements, their effect on the IP beam and minimum step-size specifications is given in Table 11-12. In all cases the SX1 or SY1 sextupole pairs are moved either symmetrically or asymmetrically.

The minimum mover step size should be $\sim$0.9 μm and $\sim$0.3 μm for the SX1 and SY1 sextupoles, respectively. The accuracy of the corrections is determined by mover precision, and can be improved by adding movable weak sextupoles

adjacent to the main ones. The minimum mover step size requirements can be increased a factor of ten by using a main to weak sextupole strength ratio of 9:1. This also allows the main sextupoles to remain aligned so that chromaticity adjustments do not couple to IP dispersion, skew or waist effects. The main-sextupole mover range should be $\pm$300–500 μm ($\pm$3 mm for the weak sextupoles).

Alternatively, instead of movable weak sextupoles, one could correct waist, dispersion and skew coupling with three pairs of normal and skew quadrupoles, as are being used in the SLC. An advantage is that field changes of such tuning quadrupoles could be implemented much faster than the equivalent sextupole moves. The maximum pole-tip fields required to correct a 100% spot-size increase are of the order 10–30 G at a 2-cm radius. A magnet-time constant of 50–100 ms would be sufficient to limit the luminosity loss due to tuning: assuming that seven aberrations per beam are tuned once per hour, that each aberration tuning consists of seven beam-beam deflection scans at different quadrupole settings, and that a settle time of four time constants is allocated to each quadrupole-field change, the total luminosity loss due to time spent on aberration tuning is estimated to be about 1%.

A correction limit is set by additional aberrations generated by using sextupole movers or tuning quadrupoles in the CCX or CCY to correct optical errors induced by the final quadrupole. For example, a 100-μr rolled final quadrupole (QFT1) generates skew which is correctable by moving the SX1 main-sextupole pair vertically by 60 μm. However, the vertical beam size remains 4% larger than nominal due to the non-local correction. For this reason it will be useful to include a small air-core skew quadrupole near the final doublet. A skew quadrupole of 10-cm length and 50-mm pole radius will require $\sim$4-Gauss pole-tip field step size and $\sim$500-Gauss range at 250 GeV/beam to correct up to 250-μr QFT1 rolls. Similarly, some vernier on the final quadrupole gradients of $\sim$2% will be useful for correcting large doublet induced waist errors.

Some preliminary tuning simulations have been performed on the final focus system simply to test the tuning concept. The incoming beam at the start of the BMS was unrealistically distorted ($\beta_x = 2\beta_{x0}$, $\Delta\alpha_y = 1$, $\Delta\eta_x = 10$ mm, $\Delta\eta_y = 1$ mm) and a final quadrupole gradient error of 0.3% was introduced. This degrades the luminosity by three orders of magnitude. Assuming an IP beam size and divergence measurement with precision of 5%, and using the BMS quadrupole strengths and main-sextupole movers, the final beam sizes achieved were within 10% of nominal. More realistic simulations will be performed in the future which include magnet misalignments, rolls and gradient errors. The beam-based alignment of quadrupoles and sextupoles will be similar to that performed in the SLC [Emma 1992] or FFTB [Tenenbaum 1995]. To monitor and maintain the alignment and orbit, conventional strip-line BPMs are used as well as high-resolution radio-frequency beam-position monitors (rf BPMs) [Hartman priv] which could be integrated into the structure of about 10 quadrupoles in the CCY and the FT. Alternatively each quadrupole and sextupole could be equipped with these rf BPMs, but this is not strictly necessary. The rf BPMs can detect small dispersion-generating drifts of magnet-positions and orbits over a minute timescale (Section 11.5.6, “CCY and FT Orbit Feedback.”)

11.5.5 Tolerances

Tolerance requirements are very tight for most parts of the NLC design. This is especially true for the final-focus system, in which the beta functions are largest and which has to generate and collide beams of minuscule transverse sizes. It is, therefore, necessary to evaluate all tolerances and to demonstrate that they can be met and how .

Variations of the incoming beam parameters, magnet motion, and changes of the magnetic fields will all affect the luminosity. Each change, drift, or vibration needs to be smaller than a certain tolerance value, in order that the design luminosity can be achieved. The timescale over which a tolerance has to be met depends on the type of aberrations generated and also on the tuning frequency or on the damping time of an automated feedback loop.

In accordance with Table 11-9, we distinguish between the following stability categories and associated timescales:

vibration/ripple:	$t <$ 1/5 s
stability/drift:	$t <$ 1 hr
long-term stability:	$t <$ 1 week
very long term stability:	$t <$ 1 month
extremely long term stability:	$t <$ 1 year

Each timescale corresponds to a tuning frequency. All tolerances studied in this section are specified relative to these timescales. Specifically, the tolerances and sensitivities discussed in this section pertain to vibration and ripple (steering tolerances), to stability (skew, waist, and dispersion effects), and to long-term stability (second-order dispersion, chromaticity, geometric aberrations).

Unless noted otherwise, all tolerances presented in this section will refer to a c.m. energy of 1 TeV , normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 10 \times 10^{-8}$ m, and to IP spot sizes $\sigma_y \approx 3.6$ nm, $\sigma_x \approx 225$ nm. Both emittances and spot sizes are about 10% larger than the linear design values quoted in Table 11-1.

In the following, we first determine how sensitive the IP spot size is to incoming orbit variations and to increased emittances. We then look at how displacing a single magnet affects the IP beam position and beam size. Next, the field stability tolerances are discussed for bending magnets, quadrupoles and sextupoles, in Section 11.5.5, "Power Supply Tolerances." In Section 11.5.5, "Scaling of Tolerances," we briefly describe how the tolerances scale with beta function, emittance and energy. In Section 11.5.5, "Tolerance Budget and Luminosity Loss," the overall tolerance budget and the resulting luminosity loss are described, including the different contributions to various aberrations. If the magnet motion is due to ground waves, the individual displacements are not uncorrelated, and the tolerances for incoherent magnet motion do not apply. Instead the wavelength-dependent response of the final-focus system has to be calculated. The luminosity loss due to ground motion is computed in Section 11.5.5. "Ground Motion". Ground motion also plays a role in requirements on repetition rate of both IP spot-size tuning (waist correction, etc.) and beam-based alignment. These questions will be addressed in Section 11.5.5. "Capture Tolerances and Tuning Frequency". Collimator tolerances are discussed in Section 11.5.5, "Collimator Tolerances". Next, thermal stability and the required vacuum pressure are discussed in Sections 11.5.5, "Temperature", and 11.5.5, "Vacuum". Finally, some beam-loss mechanisms other than gas-scattering are described.

We will distinguish between sensitivities and tolerances. By sensitivity we mean the amount of displacement or field change of a single magnet which causes either a 0.1 σ_x (0.2 σ_y) steering error or a 1% spot-size increase. While the sensitivity can be calculated for every magnet, tolerances are assigned. The tolerances are chosen so as to restrict the total luminosity loss from each aberration, due to all magnets, to 1–2%. Thus tolerances translate directly into engineering specifications on power supplies or magnet supports, while the sensitivity has no immediate practical interpretation. The two notions are, however, closely related, since a tolerance is roughly equal to the sensitivity divided by the square root of the number of magnets.

Incoming Jitter Tolerance

Figure 11-19 shows the relative increase of the horizontal and vertical IP spot-size caused by a change of the incoming horizontal and vertical beam orbit. The orbit change is quoted in units of the design beam size (or divergence) at the entrance of the final focus, and the resulting blow-up represents an average over both position and slope changes of either sign.

The figure demonstrates that a horizontal or vertical orbit change by 0.5 σ would lead to an increase of the vertical or horizontal IP spot size by less than 1%. This sensitivity to orbit changes appears tolerable.

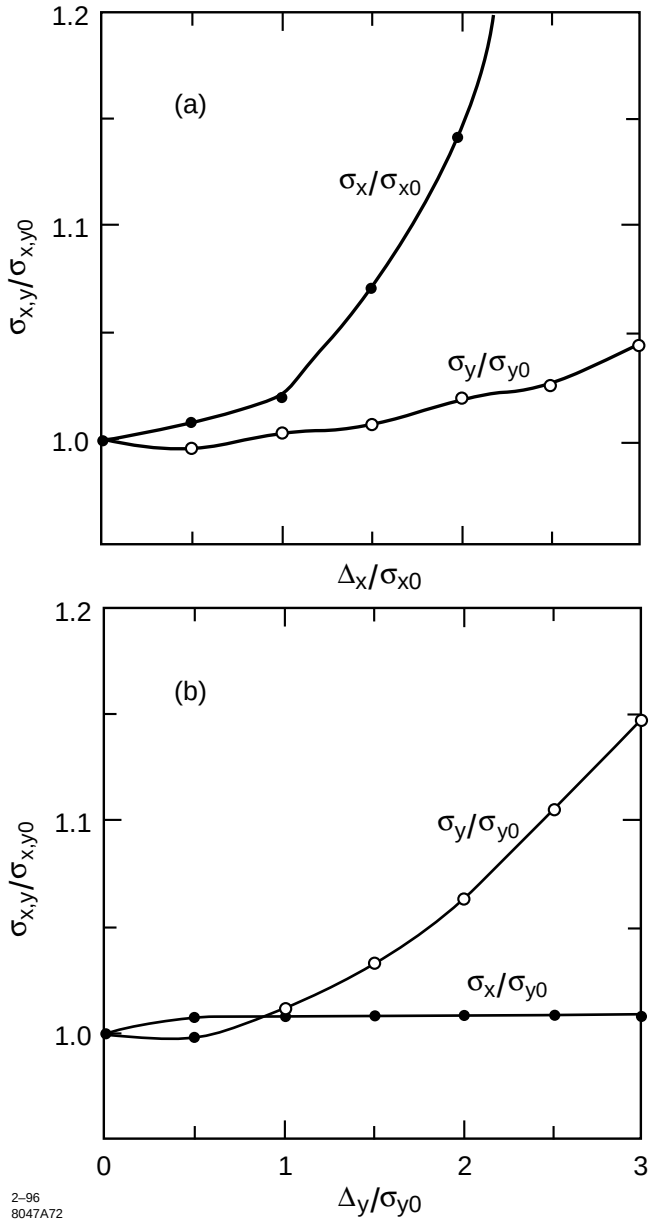


Figure 11-19. *Relative increases of the IP spot sizes as a function of horizontal or vertical incoming orbit variation in units of the design beam size. Effects of synchrotron radiation are not included.*

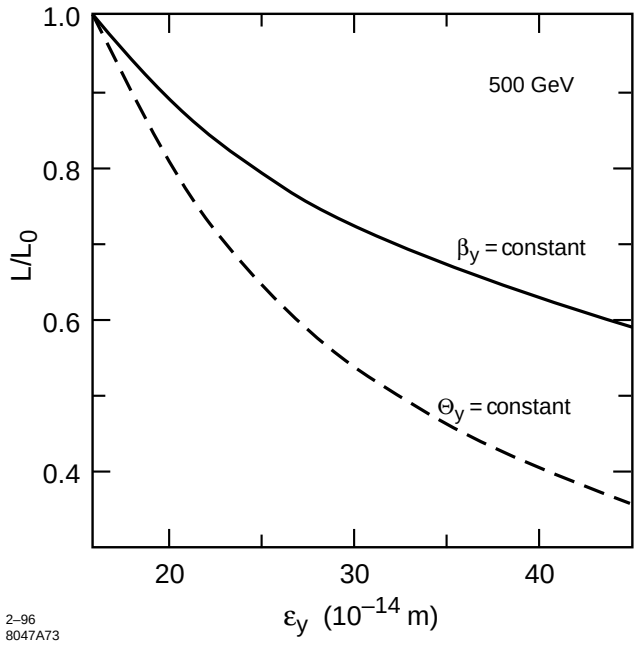


Figure 11-20. *Relative loss of luminosity for larger vertical emittances, keeping either IP beta function or divergence constant.*

Sensitivity to Emittance Increase

Figure 11-20 illustrates the relative loss in luminosity, disruption and hourglass effect not included, at 500 GeV as a function of the vertical emittance, for either constant β_y^* or constant θ_y^*. It is clearly advantageous to keep the beta function constant, and the importance of the chromo-geometric aberrations, relative to the linear spot size, does not appear to increase for the emittances considered. The reason is that the Lie generators of the dominant aberrations, $x'y'^2\delta^2$ and $y'^2\delta^3$ (compare Table 11-10), are proportional to y'^2 and thus increase in the same way as the linear spot size. If the vertical emittance is three times larger than the design, the luminosity is reduced by about 40%. The bandwidth for twice the vertical emittance is still the same as the design value.

A similar result is found if both the vertical and the horizontal emittances are increased together. When both emittances are doubled, to $\gamma\epsilon_x = 10^{-5}$ m and $\gamma\epsilon_y = 17.5 \times 10^{-8}$ m, the horizontal and vertical spot sizes for a 0.3% energy spread at 500-GeV-c.m. energy are 416 nm and 7.16 nm, respectively, if the beta functions are held constant (8 mm and 125 μm). This corresponds to a luminosity loss by about a factor of 2, so that again contributions from the nonlinear aberrations do not significantly increase. The bandwidth does not deteriorate, as it is shown in Figure 11-21.

Tolerances on Magnet Vibration and Position Drift

Steering Tolerance. *a) Quadrupole Vibration:* According to the jitter-budget proposed by J. Irwin (Table 9-3), the maximum tolerable incoming vertical orbit-jitter is about $0.44\sigma_y$ corresponding to a 2.4% loss of luminosity. The contributions to the IP position jitter from vertical and horizontal vibrations of magnets in the final focus (not including the final doublet) shall result in no more than an additional 0.5% luminosity reduction each.

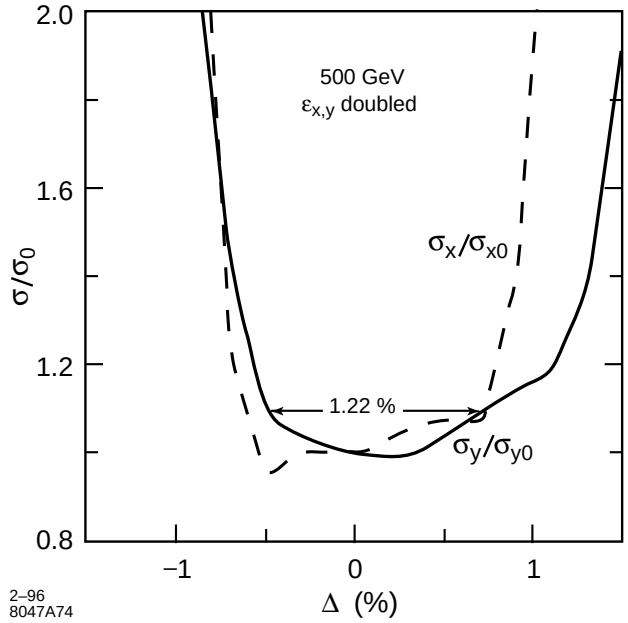


Figure 11-21. *Momentum bandwidth of the 500-GeV final-focus system, for increased emittances,* $\gamma\epsilon_x = 10^{-5}$ *m,* $\gamma\epsilon_y = 17.5 \times 10^{-8}$ *m, and constant IP beta functions,* $\beta_x^* = 8$ *mm,* $\beta_y^* = 125$ μ*m.*

The tolerance on the quadrupole vibration amplitude is thus calculated by requiring that the resulting change of the horizontal (vertical) IP beam position is not larger than $0.1\sigma_x$ ($0.2\sigma_y$). The vertical tolerance is looser by a factor of two, due to the large vertical disruption.

For a single quadrupole, the maximum vibration amplitude Δx, Δy is given by [Roy 1992]

$$\Delta x \quad < \quad \frac{\sigma_x^*}{10 R_{12} k_q} \tag{11.17}$$

$$\Delta y \quad < \quad \frac{\sigma_y^*}{5 R_{34} k_q} \tag{11.18}$$

where R_{12} and R_{34} are the R-matrix elements from the quadrupole to the IP and k_q denotes the integrated quadrupole strength. Sensitivities to both horizontal and vertical quadrupole vibrations are listed in Tables 11-13 and 11-14. Note that these numbers apply to each magnet individually, and that the actual vibration tolerance is given in Table 11-19. The tightest sensitivities (again not including the final doublet) for a 0.5 % luminosity loss, due to a single magnet, correspond to vibration amplitudes of about 70 nm horizontally and 14 nm vertically, for about seven magnets at 1-TeV-c.m. energy (Table 11-14, and Figures 11-23 and 11-24).

The sensitivity for skew-quadrupole vibrations can be calculated in a similar way. Some resulting tolerances for the correction skew-quadrupoles in the SCS can be found in Table 11-28.

b) Roll of Bending Magnet: If a bending magnet vibrates in the x-y plane, it can steer the beam vertically at the IP. The maximum allowed roll-angle θ_{xy}, for a $0.2\sigma_y$ change of the IP beam position, is

$$\theta_{xy} \leq \frac{\sigma_y^*}{5 R_{34}^{bIP} \theta} \tag{11.19}$$

where σ_y^* denotes the vertical IP beam size, R_{34}^{bIP} the (3,4)-R matrix element from the bending magnet to the IP, and θ the bending angle. The tolerances for most magnets are fairly tight (Table 11-16 and Figure 11-31), eg a change in roll by 2 μr of a single bend magnet in the BX can reduce the luminosity by 1%. This is a jitter tolerance, since steering errors at the IP will be corrected in less than 0.2s by a fast IP collision feedback, akin to that of the SLC.

Dispersion. A displaced or rolled magnet can generate dispersion at the IP. If the dispersion is tuned and corrected in hourly intervals, the corresponding tolerances are position-drift tolerances, which for some magnets may be difficult or expensive to meet. Fortunately, dispersion can be corrected continuously, if one utilizes the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy for both beams. For this purpose, the pulse-to-pulse energy variation could be intentionally increased, by lowering the gain of one or more energy feedback loops.

a) Quadrupole-Position Drift: A displaced quadrupole increases the IP spot size due to dispersion, which is generated both directly by the chromatic kick from the displaced quadrupole and by the orbit change in the downstream elements. The spot-size increases by 1% or less, if the quadrupole-position drift Δx or Δy satisfies

$$\Delta x < \frac{\sigma_x^*}{5\sqrt{2}\delta \mid -R_{12}+T_{126}\mid k_q} \tag{11.20}$$

$$\Delta y < \frac{\sigma_y^*}{5\sqrt{2}\delta \mid -R_{34}+T_{346}\mid k_q} \tag{11.21}$$

where δ is the rms momentum spread ($\delta \approx 0.3\%$), and T_{126}, T_{346} denote the T-matrix elements (in TRANSPORT notation) from the quadrupole to the IP. For magnets inside or behind the CCY, the largest contribution to these two coefficients arises from the uncompensated chromaticity of the final doublet.

At 1 TeV, the maximum tolerable vertical dispersion at the IP is 160 nm, corresponding to a 1% blow-up of the vertical spot size for an intrabunch energy spread of $\delta \approx 0.3\%$. The vertical displacement corresponding to a 1% luminosity loss is about 40 nm for the first two quadrupoles in the FT (QFT5 and QFT6) and for QEI4, further upstream, and it is 140 nm for the two quadrupoles in the center of the CCY (Table 11-14). These sensitivities correspond to an orbit change by about 110 nm at the end of the first doublet magnet (QFT3), which needs to be detected and corrected by a fast orbit feedback. The orbit variation to be measured is small compared with the vertical rms beam size of 60 μm. Nevertheless, simulations show that orbit correction with the desired accuracy is possible in less than 2 s, based on four 1-μm-resolution BPMs. These simulations are presented in Section 11.5.6, "CCY and FT Orbit Feedback." The feedback response time could be further improved by using rf BPMs, which have much higher resolution.

Horizontal displacement sensitivities due to dispersion are 140 nm or larger.

b) Sextupole-Position Stability: Sextupole motion at a location with nonzero dispersion generates dispersion at the IP. The corresponding maximum motion for a 1% spot-size increase is

$$\Delta x_s^\eta < \frac{\sqrt{2}\sigma_x^*}{5k_s\delta\eta_s R_{12}} \tag{11.22}$$

$$\Delta y_s^\eta < \frac{\sigma_y^*}{5\sqrt{2}k_s\delta\eta_s R_{34}} \tag{11.23}$$

where δ denotes the rms momentum spread, $R_{12,34}$ are the R-matrix elements from the sextupole to the IP, $\sigma_{x,y}$ the IP beam sizes, η_s is the nominal dispersion at the sextupole, and k_s the integrated sextupole strength.

In case the dispersion is not corrected continually, this is both a vibration and a stability tolerance, which has to be met over 30–60 minutes.

c) Bend Roll: Also a rolled bending magnet generates dispersion at the IP. The bend roll causing a 1% luminosity loss is given by

$$\theta_{xy}^{b} \leq \frac{\sigma_y^*}{5\sqrt{2}|T_{346} - R_{34}|\theta\delta} \tag{11.24}$$

where θ is the bend angle, δ the rms momentum spread, and R_{34}, T_{346} denote the R- and T- matrix elements from the bending magnet to the IP.

Skew Coupling and Waist Shift. *a) Sextupole-Position Stability:* In addition to dispersion, sextupole motion causes waist shift or skew coupling, and also thereby increases the IP spot size. The corresponding maximum motion for a 1% luminosity loss is:

$$\Delta x_s^w \quad < \quad \frac{1}{5\sqrt{2}k_s \max(\beta_x, \beta_y)} \tag{11.25}$$

$$\Delta y_s^s \quad < \quad \frac{1}{5\sqrt{2}k_s\sqrt{\beta_x\beta_y}}\sqrt{\frac{\epsilon_y}{\epsilon_x}} \tag{11.26}$$

where ϵ_x, ϵ_y denote the emittances and k_s the integrated sextupole strength. Values for the 15 sextupoles in the 500-GeV final focus are listed in Table 11-15. The tightest sensitivities for a 1% luminosity loss are about 500 nm horizontally and vertically for the two main Y-sextupoles (Figures 11-28 and 11-29). Again the actual tolerances that have to be achieved are tighter than those quoted, in this case by roughly a factor of two assuming uncorrelated sextupole motion.

b) Quadrupole-Position Stability (or Drift): The same skew coupling or waist shift is generated when a quadrupole moves inside the CCX or CCY steering the beam off-center through the second sextupole. The maximum drift amplitudes due to this effect for a 1% luminosity loss are:

$$\Delta y_q \quad \leq \quad \frac{\sqrt{2}\Delta y_s^s}{R_{34}^{qs}k_q} \tag{11.27}$$

$$\Delta x_q \quad \leq \quad \frac{\sqrt{2}\Delta x_s^w}{R_{12}^{qs}k_q} \tag{11.28}$$

where Δx_q and Δy_q denote the motion tolerances, due to waist shift or skew coupling, R_{34}^{qs} and R_{12}^{qs} the R-matrix elements from the quadrupole to the next sextupole, and k_q the integrated quadrupole strength. The term Δy_s^s is the tolerance for sextupole movement giving rise to a skew aberration, and Δx_s^w is the tolerance for sextupole movement giving rise to a waist aberration. For a few quadrupoles in the CCY, the vertical (horizontal) displacement sensitivity due to this effect can be as tight as 110 (500) nm, for a spot size increase by 1% (Figures 11-25 and 11-26 present sensitivities for 1% luminosity loss).

c) Quadrupole Roll: A quadrupole rolling in the x-y plane may generate skew coupling at the IP. The tolerance on the x-y roll angle θ_{xy} for a 1% spot size increase is given by

$$\theta_{xy} \leq \frac{1}{10\sqrt{2}k_q\sqrt{\beta_x\beta_y}}\sqrt{\frac{\epsilon_y}{\epsilon_x}} \tag{11.29}$$

where β_x, β_y are the horizontal and vertical beta function at the quadrupole, and k_q is the integrated quadrupole strength. Roll sensitivities for all quadrupoles are listed in Table 11-14. The tightest actual tolerances are those for the final-doublet magnets QFT3–QFT1: about 1.5 μr in 1 h (compare Figure 11-27).

d) Bend Roll: A rolled bending magnet may steer the beam vertically at the next sextupole, thus generating skew coupling. This effect gives rise to a second drift tolerance on the bend roll angle, in addition to that arising from

vertical dispersion. The bend roll drift causing a 1% luminosity loss due to skew coupling is

$$\theta^b_{xy} \leq \frac{\Delta y^s_s}{\sqrt{25} R^{bSX}_{34} \theta} \tag{11.30}$$

where θ is the bend angle, Δy^s_s the drift tolerance for the downstream sextupole due to skew coupling, and R^{bSX}_{34} the R-element between bending magnet and sextupole.

Second-Order Dispersion. *a) Quadrupole Alignment:* The maximum tolerable second-order dispersion at the IP is about 35 μm, for a 1% luminosity loss and an rms energy spread of $\delta_{\rm rms} \approx 0.3\%$. If no tuning scheme for this aberration is implemented, the number above translates into a vertical alignment tolerance for quadrupole magnets in the final focus. Second-order dispersion is primarily generated by interaction of a chromatic kick from a displaced quadrupole with the doublet chromaticity. In the absence of a correction scheme, the absolute alignment tolerance for a single final-focus quadrupole due to generated second-order dispersion would be

$$\Delta y < \frac{\sigma^*_y}{10\ \delta^2\ |T_{346}|\ k_q}. \tag{11.31}$$

The tightest sensitivities of about 12 μm are those for the first two magnets in the FT (QFT6 and QFT5). Sensitivities for all quadrupoles are given in Tables 11-13 and 11-14.

One possibility for compensating the second-order dispersion is by vertical moves of quadrupoles QFT6 and QFT5. A quadrupole move will also create first-order dispersion, which is of similar magnitude as the second-order dispersion and which, therefore, will dominate the spot size during a scan. Specifically, we have the following relation between first- and second-order dispersion $\eta^{(1)}_y$, $\eta^{(2)}_y$ at the IP and the vertical quadrupole positions Δy_{QFT5}, Δy_{QFT6}:

$$\eta^{(1)}_y = 3.8 \times \Delta y_{QFT5} - 3.0 \times \Delta y_{QFT6} \tag{11.32}$$

$$\eta^{(2)}_y = -4.0 \times \Delta y_{QFT5} + 1.7 \times \Delta y_{QFT6} \tag{11.33}$$

In principle, the two quadrupoles can be moved simultaneously such that only second-order dispersion is generated, but no first-order dispersion. As an example, if QFT5 is moved by -18 μm and QFT6 by -23 μm, a second-order dispersion of 35 μm is produced, and the first-order dispersion is not changed. However, in order to realize such a pure $\eta^{(2)}_y$ tuning knob, the mover positions would need to be controlled with an unrealistic precision of 30 nm. Otherwise the measured spot sizes during an $\eta^{(2)}_y$-scan are still affected by residual first-order dispersion. Hence, the second-order dispersion cannot be tuned by a simple scan, but instead an iterative procedure is necessary: after each QFT5/QFT6 move the first-order dispersion is corrected by the standard procedure using sextupole movements in the CCY, before the vertical spot size can be measured. Such a 2-dim. scan (or a "scan of scans") would be a new type of procedure, not used at the SLC.

This complication can be avoided by utilizing the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy, as it has already been proposed for correcting the regular dispersion. In this case, the second-order dispersion can be continually tuned, and it then causes no significant spot-size dilution.

Power Supply Tolerances

Steering Tolerance. *a) Bending Magnets:* A field change of a bending magnet steers the beam horizontally at the IP. To give a horizontal steering error smaller than $0.1\ \sigma_x$, the relative field-ripple of the bending magnets must satisfy

$$\frac{\Delta\theta}{\theta} \leq \frac{\sigma^*_x}{10\ R^{bIP}_{12}\theta} \tag{11.34}$$

where θ denotes the nominal bending angle, and R_{12}^{bIP} the R-matrix element from the bending magnet to the IP. Sensitivities for bending magnets in the different bending sections are listed in Table 11-16 (Figure 11-30). The actual tolerances which need to be achieved are quite tight, on the order of 3×10^{-6}. This is a "jitter" tolerance, which has to be met over 1/5 s.

Waist Shift. *a) Bending Magnets:* Similar to the effect of horizontally drifting quadrupoles (Eq. 11.27), a field change of the bending magnets inside the CCX or the CCY steers the beam horizontally off-center through the next sextupole, which causes a waist shift at the IP. The maximum allowed field ripple for a 1% luminosity loss is given by

$$\frac{\Delta\theta}{\theta} = \frac{\Delta x_s^w}{R_{12}^{bs}\theta} \tag{11.35}$$

where Δx_s^w denotes the horizontal vibration tolerance due to waist shift of the next sextupole (SX1b or SY1b), θ the nominal bending angle, and R_{12}^{bs} the R-matrix element from the bending magnet to the sextupole. Some numbers may be found in Table 11-16. Assuming 10-m-long bend sections inside the CCX or the CCY are fed by independent power supplies, the tolerable relative field ripple is about 2×10^{-5}. The tolerance decreases to 4×10^{-6}, if all magnets are powered as a string. This is a stability tolerance (timescale 1 hr), unless there is an orbit stabilization system such as discussed in Section 11.5.6: "CCY and FT Orbit Feedback."

b) Quadrupoles: The change of a quadrupole field also causes a waist shift at the IP. For an associated spot-size increase smaller than 1%, the maximum field change is written [Roy 1992]

$$\frac{\Delta k_q}{k_q} = \frac{1}{5\sqrt{2}\max(\beta_x, \beta_y)k_q} \tag{11.36}$$

where k_q is the integrated quadrupole strength, and $\beta_{x,y}$ the beta function at the quadrupole. The sensitivities for the quadrupoles of CCX, CCY, and FT are listed in Table 11-17. The tightest sensitivities (except for the final doublet) for a 1% luminosity loss are about 10^{-4}. Again, the final tolerance that actually needs to be achieved will be smaller than the quoted numbers, each of which refers to ripple in a single magnet only.

Dispersion. Field changes of magnets may give rise to dispersion at the IP. It is contemplated to continually monitor and correct the IP dispersion, using the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy of both beams, so as to render the dilution due to dispersion as insignificant.

a) Quadrupoles: In addition to causing a waist shift, a field change of quadrupole magnets at locations with nonzero dispersion also generates horizontal dispersion at the IP. The relative field change for a 1% loss of luminosity due to the induced dispersion is given by

$$\frac{\Delta k_q}{k_q} = \frac{\sigma_x^*}{5\sqrt{2}\delta k_q \eta_x^q R_{12}} \tag{11.37}$$

where k_q is the integrated quadrupole field, R_{12} the R-matrix from the quadrupole to the IP, and η_x^q the value of the dispersion at the quadrupole. Sensitivities are listed in Table 11-17.

The relative field-drift sensitivity of four CCX quadrupoles is 4×10^{-4} for a 1% luminosity loss due to generated dispersion. Their sensitivity due to induced waist shift is about twice as tight.

b) Bending Magnets: Horizontal dispersion is also generated by a field change of a bending magnet which alters the downstream orbit. The change of bend angle $\Delta\theta$ causing a 1% luminosity loss is given by

$$\Delta\theta \leq \frac{\sigma_x^*}{5\sqrt{2}|T_{126} - R_{12}|\delta} \tag{11.38}$$

where θ denotes the bending angle, δ the rms momentum spread, σ_x^* the horizontal IP spot size, and T_{126}, R_{12} the T- and R-matrix elements from the bend to the IP (in TRANSPORT notation).

Chromaticity. *a) Sextupoles:* The sextupole field-drift sensitivity due to the induced chromaticity follows from

$$\frac{\Delta k_s}{k_s} \leq \frac{1}{5\sqrt{2}k_s \max(\beta_{s,x}, \beta_{s,y})\eta_{s,x}\delta} \tag{11.39}$$

where $\beta_{s,x,y}$ denotes the horizontal or vertical beta function at the sextupole, $\eta_{s,x}$ the dispersion and k_s the integrated sextupole strength in units of m^{-2}. Sensitivities are not very tight (Table 11-18).

Geometric Aberrations. *a) Sextupoles:* The sextupole field-drift sensitivity due to induced third-order geometric aberrations is approximately

$$\frac{\Delta k_s}{k_s} \leq \frac{1}{5\sqrt{2}k_s \sqrt{\beta_{s,x}\epsilon_x} \max\left(\frac{2}{\sqrt{3}}\beta_{s,x}, \beta_{s,y}\right)} \tag{11.40}$$

where $\beta_{s,x}$ and $\beta_{s,y}$ denote the horizontal and vertical beta function at the sextupole, and k_s the integrated sextupole strength. The resulting tolerances are loose and similar to those imposed by chromaticity (Table 11-18).

Magnet	Steering		Spot Size			
			disp. and skew		2nd o. disp.	roll
	Δx (nm)	Δy (nm)	Δx (μm)	Δy (μm)	Δy (μm)	θ_{xy} (mrad)
QDD	1700.	150.	7.	0.14	33.	4.9
QDF	600.	650.	11.	0.26	61.	7.6
QDD	660.	440.	37.	0.30	70.	4.3
QDF	400.	220.	6.	1.92	470.	2.5
QDD	640.	100.	6.	0.49	110.	2.2
QDF2	340.	80.	2.	0.20	46.	1.6
QDD2	2300.	50.	12.	0.09	22.	3.0
QDF2	240.	90.	4.	0.15	35.	1.3
QDD	640.	150.	6.	0.18	42.	4.0
QDF	550.	650.	3.	0.36	85.	10.0
QDD	1400.	440.	5.	0.51	120.	8.4
QDF	9200.	220.	4.	7.54	1600.	4.7
QDD	1700.	100.	7.	0.41	95.	2.7
QDF	600.	130.	8.	0.37	87.	1.9
QDD	660.	90.	65.	0.20	47.	1.8
QDF	400.	150.	12.	0.29	69.	22.3
QDD	640.	150.	9.	0.23	54.	4.0
QDF2	340.	720.	3.	0.35	81.	6.4
QDD2	2300.	240.	12.	0.49	116.	7.2
QDF2	240.	160.	11.	13.16	2290.	1.6
QDD	640.	100.	11.	0.52	120.	2.2
QDF	550.	130.	5.	0.52	120.	2.6
QDD	1400.	90.	6.	0.29	67.	3.6
QDF	9200.	150.	4.	0.45	104.	4.3
QDD	1700.	150.	7.	0.37	86.	4.9
QDF	600.	650.	7.	0.87	204.	7.6
QDD	660.	440.	17.	2.62	632.	4.3
QG1	430.	240.	122.	2.41	545.	2.6
QG2	1200.	70.	11.	0.36	82.	2.8
QG1	2300.	250.	3.	1.73	395.	4.5
QG2	1700.	3000.	8.	1.36	320.	6.8
QB1	2100.	1200.	39.	8.71	1993.	13.0
QB2	2700.	950.	147.	15.65	3450.	8.8
QB3	880.	370.	199.	19.44	3762.	1.3
QB4	280.	70.	32.	5.19	947.	0.2
QB5	84.	50.	9.	3.75	668.	0.1

Table 11-13. *Vibration, drift and roll sensitivity of quadrupoles between the SCS and BMS in the 1-TeV final focus; each number corresponds to a horizontal (vertical) position error of $0.1\sigma_x$ ($0.2\sigma_y$), or to a 1% spot-size increase, respectively.*

Magnet	Steering		Spot Size			
			disp. and skew		2nd o. disp.	roll
	Δx (nm)	Δy (nm)	Δx (μm)	Δy (μm)	Δy (μm)	θ_{xy} (μrad)
QE2	430.	89.	280.0	9.8	1600.	200.
QE2	400.	89.	240.0	9.8	1600.	200.
QE1	76.	97.	26.0	7.7	1400.	63.
QX3	76.	97.	26.0	7.7	1400.	63.
QX2	250.	57.	4.0	2.3	830.	120.
QX2	260.	58.	4.0	2.3	830.	130.
QX1	30000.	6800.	1.0	3.5	1300.	800000.
QX1	15000.	3400.	1.0	3.5	1300.	300000.
QX2	260.	58.	3.0	1.8	670.	130.
QX2	250.	57.	3.0	1.8	680.	120.
QX3	76.	98.	28.0	9.4	1600.	63.
QBX1	76.	97.	32.0	8.1	1400.	63.
QBX2	250.	57.	8.0	1.2	260.	120.
QBX2	260.	58.	7.0	1.2	260.	130.
QBX3	30000.	6800.	2.0	2.0	480.	800000.
QBX4	22000.	4900.	3.0	2.9	680.	480000.
QBX5	260.	58.	2.0	8.3	4800.	91.
QBX5	260.	56.	2.0	8.4	5200.	88.
QBX6	440.	17.	46.0	6.8	2400.	45.
QY3	440.	17.	47.0	7.0	2300.	45.
QY2	260.	57.	0.6	0.7	2500.	89.
QY2	260.	59.	0.6	0.7	2400.	92.
QY1	43000.	9600.	1.0	0.2	52.	—
QY1	22000.	4800.	1.0	0.2	52.	480000.
QY2	260.	58.	0.6	0.3	90.	90.
QY2	260.	56.	0.6	0.3	93.	88.
QY3	440.	17.	46.0	8.4	1800.	45.
QEI1	440.	17.	96.0	5.9	3100.	45.
QEI2	260.	57.	2.0	0.3	88.	89.
QEI2	260.	59.	2.0	0.4	85.	92.
QEI3	43000.	9700.	3.7	0.1	25.	—
QEI4	85000.	19000.	14.0	0.4	99.	—
QFT6	50.	11.	1.0	0.1	13.	9.
QFT5	70.	8.	1.5	0.1	11.	9.
QFT4	120.	17.	48.0	0.3	68.	20.
QFT3	12.	3.	2.2	0.7	85.	3.
QFT2	17.	2.	3.2	0.7	600.	3.
QFT1	21.	1.	4.6	0.1	79.	4.
QFTA	37.	1.	9.0	0.1	67.	9.

Table 11-14. *Vibration, drift, and roll sensitivity of quadrupoles between the CCX and IP in the 1-TeV final focus; each number corresponds to a horizontal (vertical) position error of $0.1\sigma_x$ ($0.2\sigma_y$), or to a 1% spot-size increase, respectively.*

Magnet	Δx (μm)	Δy (μm)
SX1	3.0	1.9
SX3	260000.0	46000.0
SX1	3.0	1.9
SY1	0.5	0.5
SY3	460000.0	81000.0
SY1	0.5	0.5
SI2	320000.	57000.0

Table 11-15. *Sextupole vibration sensitivity for the 1-TeV final focus due to induced skew coupling or waist shift, for a 1% spot-size increase.*

Name	$\Delta\rho/\rho$ (10^{-5}) (steering)	$\Delta\rho/\rho$ (10^{-4}) (beam size)	θ_{xy} (μrad)
BG	16.5	—	96.0
B1A	6.5	—	15.0
B1B	5.2	—	11.0
B2	1.7	2.3	11.0
B3A	3.5	—	8.0
B3B	3.5	—	1.7
B4	3.3	0.8	3.4
B5A	5.7	—	6.7
B5B	4.4	—	10.0
B5C	10.0	—	23.0

Table 11-16. *Field ripple and roll sensitivity for the bending magnets. Each number corresponds to a steering error by $0.1\sigma_x$, $0.2\sigma_y$ or to a spot-size increase by 1%, respectively.*

Magnet	Waist Shift $\Delta k/k$ (10^{-3})	Dispersion $\Delta k/k$ (10^{-3})	Total $\Delta k/k$ (10^{-3})
QE2	2.10	8.5	2.08
QE2	2.20	7.6	2.07
QE1	0.20	0.4	0.18
QX3	0.20	0.4	0.18
QX2	1.40	3.5	1.30
QX2	1.40	3.7	1.30
QX1	11000.00	240.0	240.00
QX1	4200.00	120.0	120.00
QX2	1.40	3.8	1.30
QX2	1.40	3.6	1.30
QX3	0.20	0.5	0.18
QBX1	0.20	0.5	0.18
QBX2	1.40	5.0	1.34
QBX2	1.40	5.0	1.37
QBX3	11000.00	72000.0	11000.00
QBX4	5900.00	25000.0	5800.00
QBX5	1.00	1.9	0.90
QBX5	1.00	1.9	0.86
QBX6	0.09	4.0	0.09
QY3	0.09	4.0	0.09
QY2	1.00	1.3	0.77
QY2	1.00	1.3	0.80
QY1	16000.00	460.0	460.00
QY1	5900.00	230.0	230.00
QY2	1.00	1.3	0.80
QY2	0.96	1.3	0.80
QY3	0.09	4.0	0.09
QEI1	0.09	4.0	0.09
QEI2	0.98	1.4	0.80
QEI2	1.00	1.4	0.80
QEI3	16000.00	890.0	890.00
QEI4	23000.00	1800.0	1800.00
QFT6	0.10	330000.0	0.10
QFT5	0.05	720000.0	0.05
QFT4	0.15	—	0.15
QFT3	0.04	200000.0	0.04
QFT2	0.02	390000.0	0.02
QFT1	0.01	990000.0	0.01
QFTA	0.02	—	0.02

Table 11-17. *Field ripple sensitivity for quadrupoles between the CCX and IP due to induced waist shift or due to horizontal dispersion generated at the IP. Each listed value corresponds to a 1% spot-size increase.*

Name	$\Delta k_S/k_S$ (10^{-3})		
	Chromaticity	Geom. aberrations	Total
SX1	12.0	4.7	4.4
SX3	—	—	—
SX1	12.0	4.7	4.4
SY1	2.9	3.4	2.2
SY3	—	—	—
SY1	2.9	3.4	2.2
SI2	—	—	—

Table 11-18. *Field ripple sensitivity due to chromaticity or due to geometric aberrations for the normal sextupoles in the CCX, BX, CCY, and FT: each number listed corresponds to a 1% (horizontal or vertical) beam-size increase.*

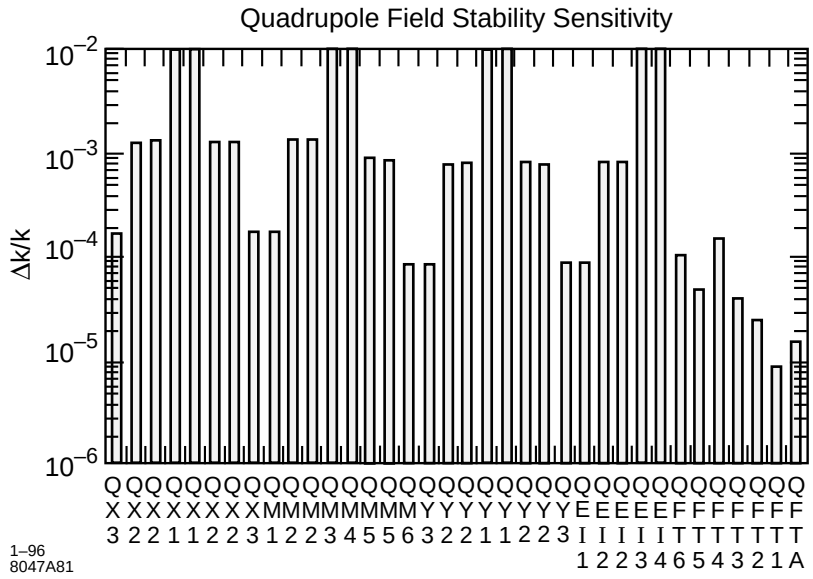


Figure 11-22. *Field-stability sensitivity of quadrupoles in the BMS, CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

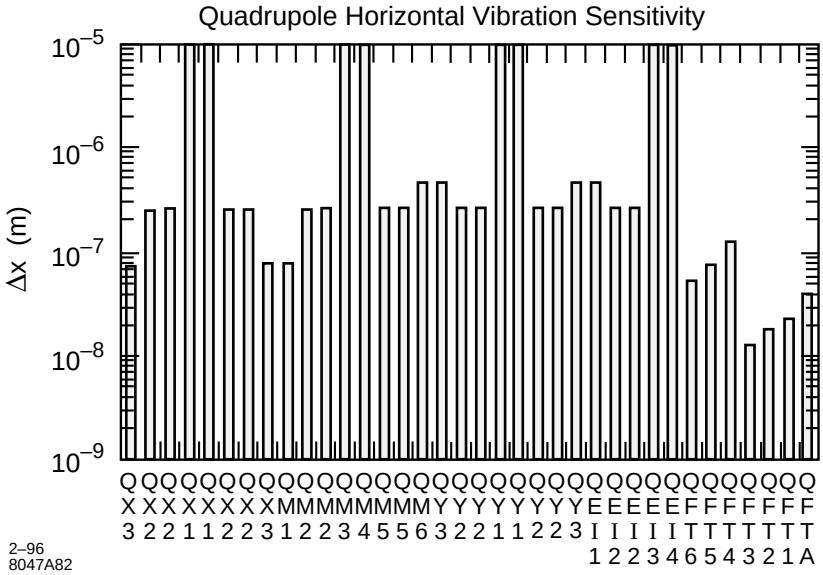


Figure 11-23. *Horizontal vibration sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

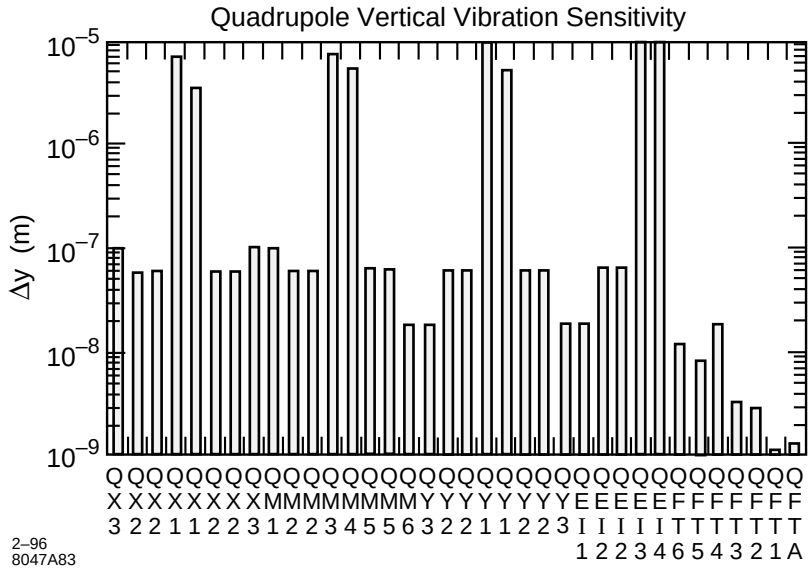


Figure 11-24. *Vertical vibration sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

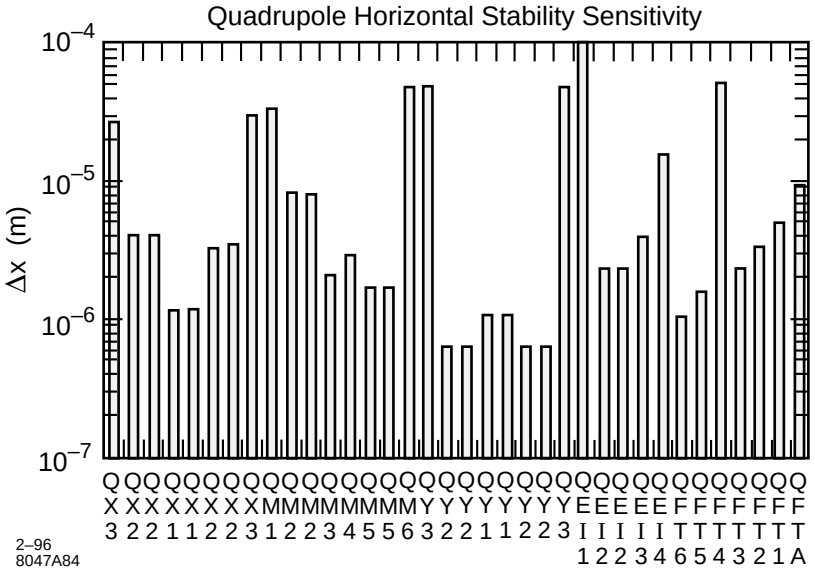


Figure 11-25. *Horizontal stability (or position-drift) sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

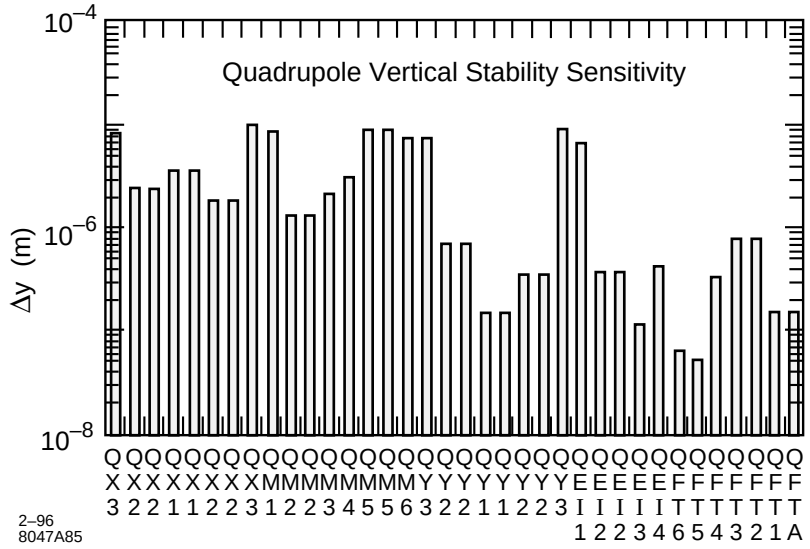


Figure 11-26. *Vertical stability (or position-drift) sensitivity of quadrupoles in CCX, BX, CCY and FT; each corresponding to a luminosity loss of 1%.*

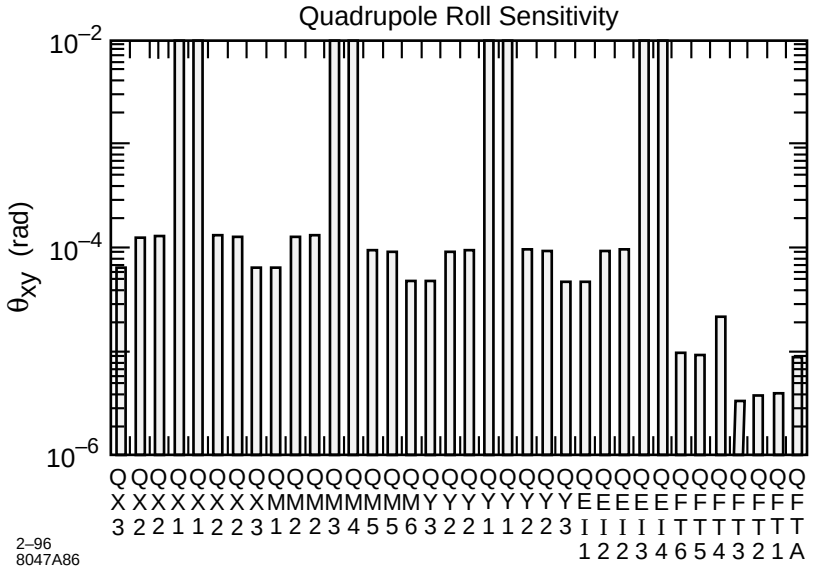


Figure 11-27. *Roll sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

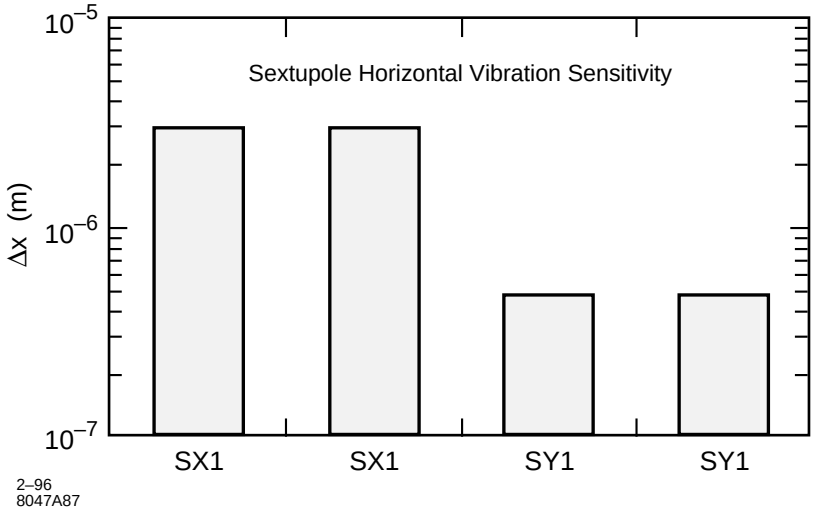


Figure 11-28. *Horizontal vibration sensitivity of main sextupoles in the final focus; each corresponding to a luminosity loss of 1%.*

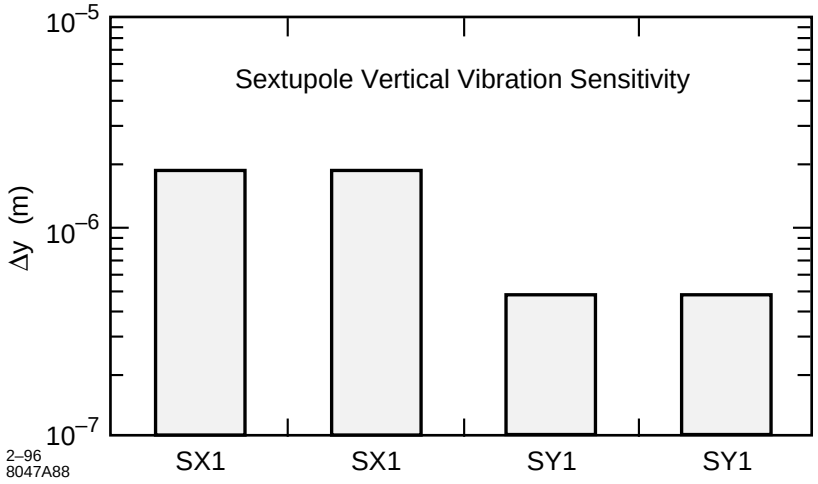


Figure 11-29. *Vertical vibration sensitivity of main sextupoles in the final focus; each corresponding to a luminosity loss of 1%.*

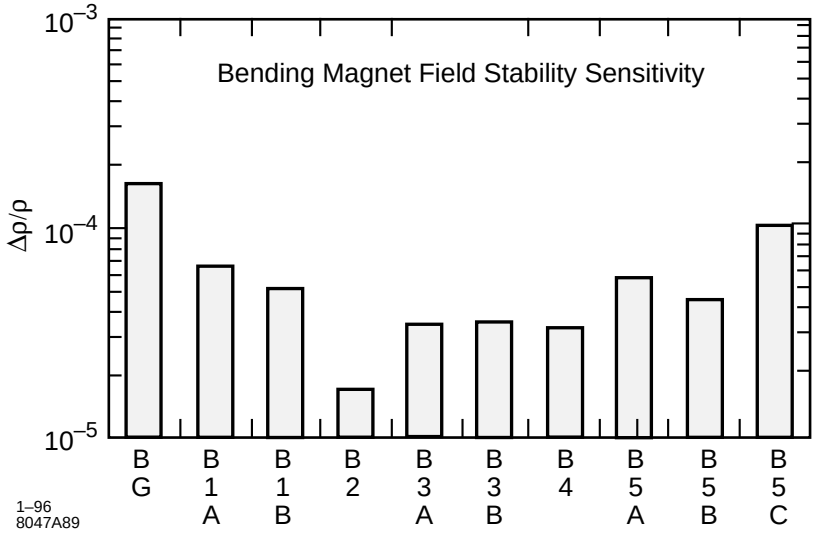


Figure 11-30. *Field-stability sensitivity for bending magnets in the final focus, assuming uncorrelated field ripple; each corresponding to a luminosity loss of 1%.*

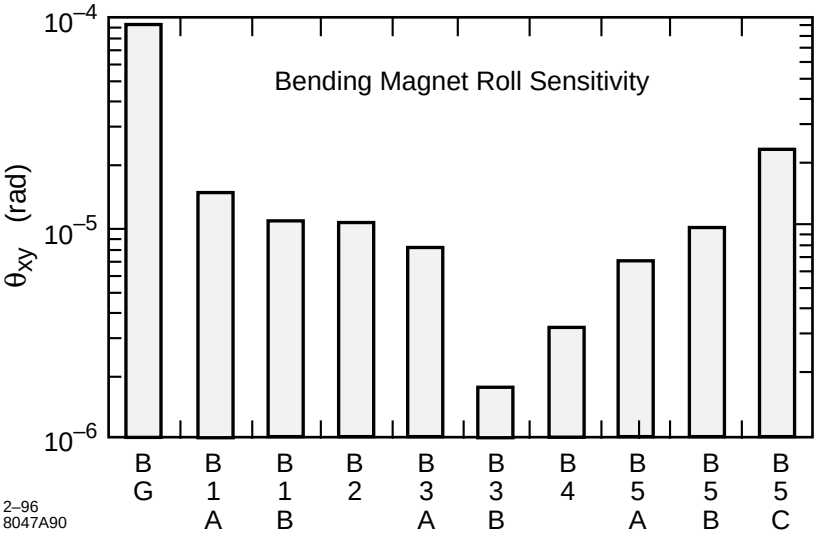


Figure 11-31. *Roll sensitivity of bending magnets in the final focus; each corresponding to a luminosity loss of 1%.*

Scaling of Tolerances

All vibration and stability (drift) tolerances presented in this section refer to a particular parameter set for 1-TeV-c.m. energy (Table 11-1). Most of the sensitivities scale directly with the IP beam size, independently of whether the latter changes due to an emittance variation or due to a different beta function. An exception are quadrupoles in front of the BMS, whose sensitivity is not affected by a change of the IP beta function. From a tolerance point of view, the operation at 500 GeV is equal to an emittance increase. Thus, almost all tolerances at 500 GeV will be a factor of $\sqrt{2}$ looser than those quoted in this chapter.

Tolerance Budget and Luminosity Loss

Beam-beam deflection scans will be used for correcting low-order aberrations which affect the IP spot size, such as waist shift, skew coupling etc.. Because of limited resolution of the deflection scans, these corrections cannot be perfect, but they are accurate to about 1%, implying a corresponding spot-size increase. Since this spot-size increase is dictated by the tuning method, it sets a lower limit on practicable tolerances, *e.g.,* for element strength changes and position drifts. Between two scans of the same aberration, the spot-size increase due to this aberration is assumed to grow from 1% to about 3%, because of magnet field and position drifts. On average (over time) the residual spot-size increase is then expected to be on the order of 2% for each tunable aberration, although various possibilities exist to reduce this number to 0.5% (see discussions in Section 11.5.3 and also Table 11-5). Thus, for consistency, a set of tolerances on magnet motion and field stability was chosen which roughly assigns a 2% spot-size increase caused by all components between linac and IP, or 1% generated in the final focus alone, to each such perturbation. Figure 11-32 illustrates that, for the tolerances chosen, various different effects may contribute by similar amounts to each aberration.

To facilitate design and cost estimates, we have grouped the magnets into different classes. The same tolerances then apply to all magnets of one class. Specifically, the quadrupoles are divided into four classes. The majority of the final-focus magnets is in Class 1. Four quadrupoles in the CCY with tighter tolerances form Class 2. Three quadrupoles at the entrance of the final transformer, with still tighter tolerances, are in Class 3, and the final-doublet magnets define Class 4. The sextupoles are split into two classes: main sextupoles and Brinkmann-sextupoles. The bending magnets are not further subdivided. A list of the different magnet types, along with relevant tolerances and the number of magnets of each type, is presented in Table 11-19. Table 11-20 describes the individual tolerances for each magnet class and how they can be met. The overall loss of luminosity corresponding to these magnet tolerances is 11%; see Table 11-21. This number assumes that dispersion is corrected at hourly intervals only, and not continually in which case the luminosity loss would be 1–2% smaller. Taking into account other residual aberrations, such as skew sextupole components and additional sources of aberrations, the estimated total loss of luminosity due to uncorrected tunable aberrations is about 13% (Table 11-9).

Ground Motion

Transverse quadrupole displacements due to ground motion may lead to a serious reduction of luminosity if the two beams are steered off collision faster than feedback systems are able to correct the IP beam position. There are several different approaches to treat the effect of ground motion on the final-focus system. Here, we base the analysis on a two-dimensional power spectrum following Juravlev and co-workers [Juravlev 1995]. A complete treatment of ground motion can be found in Appendix C.

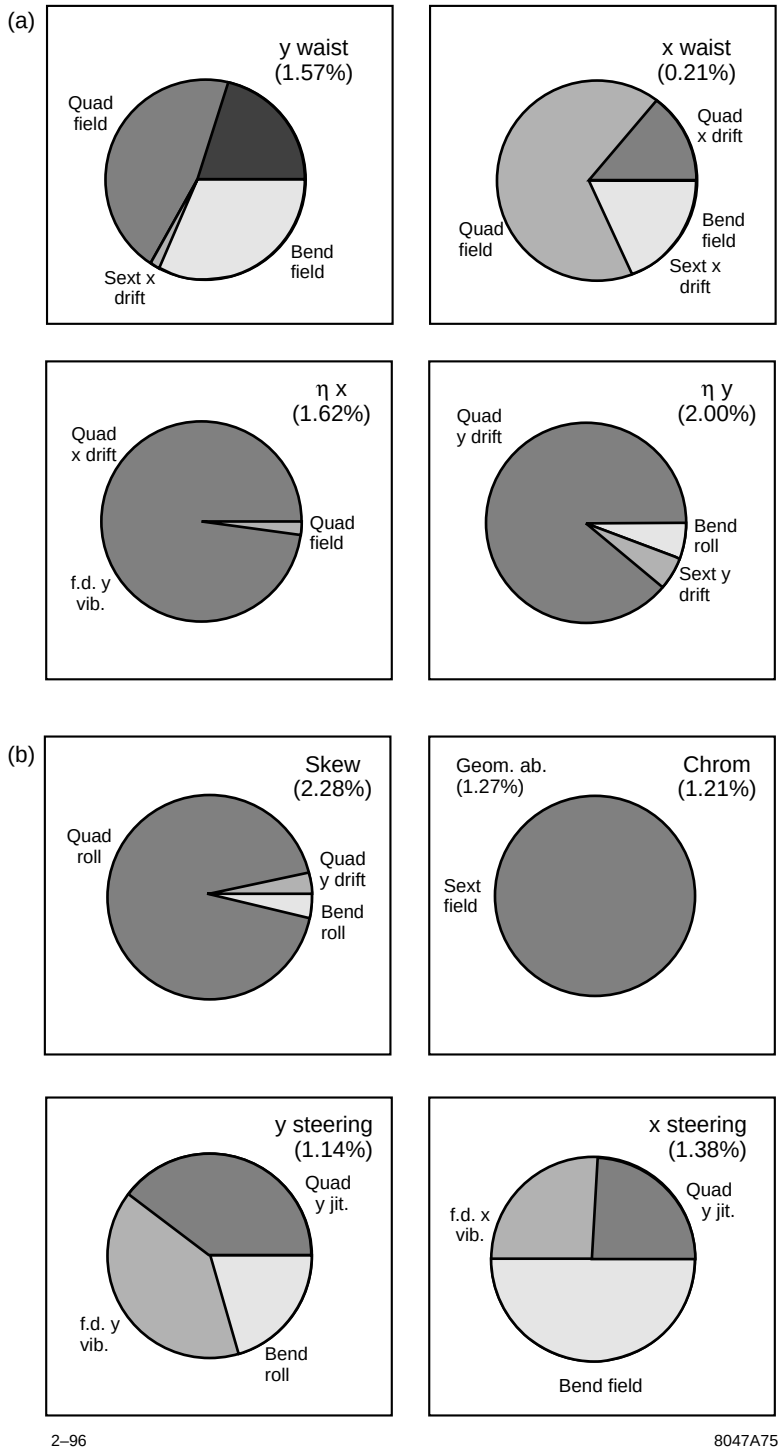


Figure 11-32. *Sources of important aberrations in the 1-TeV final-focus design, for the tolerances of Table 11-19. Numbers in parentheses refer to the overall luminosity loss due to a particular aberration.*
f.d.: final doublet; jit.: jitter; field: field ripple or drift.

Type	P #	Radius a (mm)	Length L (m)	Drift (10^{-5}) $\Delta K/K$	Vibration[a] (nm) Δx	Δy	Drift[a] Δx (nm)	Δy (nm)	θ_{xy} (μr)
Quad 1	33	10.50	0.5/1.0	3.3	6.8	7.3	144.0	33.0[b]	11.3
Quad 2	4	10.50	0.5/0.5	1.8	54.0	4.8	15000.0	2000.0	15.7
Quad 3	3	10.50	1.0/1.0	1.6	9.0	3.4	480.0	21.0[b]	3.4
Quad 4	4	4.00	1.5/2.0	0.4	3.6	0.7	1600.0	98.0	1.8
Skew q. 1	4	6.00	0.4/0.4	100.0	50.0	30.0	15000.0	15000.0	500.0
Sext. 1	4	10.50	0.4/0.4	220.0	50.0	50.0	50.0	50.0	20.0
Sext. 2	12	10.50	0.4/0.4	900.0	1000.0	1000.0	10000.0	10000.0	20.0
Bend 1	100	—	5.0/16.0	0.3	100.0	100.0	500.0	500.0	0.5

[a] With respect to ground.

[b] The vertical drift tolerance will be looser at least by a factor of 10, if the dispersion at the IP is continually corrected utilizing the correlation of beam-beam deflection and beam energy.

Table 11-19. *Magnet types and tolerances for 1-TeV-c.m. energy. Tolerances result in the luminosity loss of Table 11-21.*

If the vertical betatron phase advance from the entrance point e to the IP is a multiple of π, the offset of the two beams at the IP $\Delta y \equiv y_{\rm right} - y_{\rm left}$ due to an arbitrary vertical displacement $y(s)$ of the final-focus magnets is given by

$$\Delta y = -\sum_{i,\rm right} k_i R_{34}^{i\to IP} y(s_i) + R_{33}^{e\to IP} y(s_{e,\rm right}) + \sum_{i,\rm left} k_i R_{34}^{i\to IP} y(s_i) - R_{33}^{e\to IP} y(s_{e,\rm left}) \equiv \sum_j \mu_j y(s_j) \quad (11.41)$$

where $R_{34}^{i\to IP}$ denotes the (3,4) TRANSPORT-matrix element from the ith magnet to the IP, $R_{33}^{e\to IP}$ is the (3,3) matrix element from the entrance of the final focus to the IP, and k_i the integrated quadrupole strength of quadrupole i. For simplicity, we have replaced the $\pm k_i R_{34}$ and $\pm R_{33}$ by the generalized lattice parameters μ_j. The subindex i runs over only one side of the final focus system, while the subindex j runs over both sides. If we square the sum in Eq. 11.41, we will find mixed expressions of the form $y(s_i)y(s_j) \equiv y(s_i)y(s_i + \Delta s_{ji})$, whose expectation value over position s and over time t is given by

$$< y(s_i)y(s_j) >_{s,t} = \lim_{S\to\infty} \frac{1}{ST} \int_{-\frac{S}{2}}^{\frac{S}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} y^*(s_i,t) y(s_i + \Delta s_{ji}, t)\, ds\, dt \quad (11.42)$$

$$= \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty \frac{dk}{2\pi} P(\omega,k) \cos(k\Delta s_{ji}) \quad (11.43)$$

where $\Delta s_{ji} \equiv s_j - s_i$. The term $P(\omega,k)$ denotes the two-dimensional power spectrum (regarding frequency and wave number) of the ground motion. The expectation value for the change of the beam-offset squared after time T can be written

$$< (\Delta y(t) - \Delta y(t+T))^2 >_t = 2 \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty \frac{dk}{2\pi} G(k) P(\omega,k)(1 - \cos(\omega T)) \quad . \quad (11.44)$$

The function $G(k)$ describes the response to a harmonic displacement of quadrupoles as a function of wavelength. It equals the average squared ratio of the beam-beam offset at the IP and the ground-motion amplitude at that wavelength, and is called the lattice response function. In terms of the lattice parameters μ_j and the positions s_j, this function can be written as:

$$G(k) \equiv \sum_{j1,j2} \mu_{j1}\mu_{j2} \cos(k(s_{j1} - s_{j2})) = 4 \left(\sum_j \mu_j \sin k s_j \right)^2 \quad (11.45)$$

Ripple:	
Quad1	Conventional
Quad2	10^{-5} commercially available, standard FFTB supply
Quad3	10^{-5} commercially available, standard FFTB supply
Quad4	Permanent magnet QA: SmCo with Erbium, see Sec. 11.6. Superconducting Q1: under design Q2 and Q3: if conventional, need special supply
Skew q1	Conventional power supply
Sext1	Conventional power supply
Sext2	Conventional power supply
Bend1	Special supply
Vibration:	
Quad1	8 nm, FFTB support: ground to quad diff. 1 nm
Quad2	5 nm, FFTB support: ground to quad diff. measured as 1 nm
Quad3	3.5 nm, FFTB support: ground to quad diff. 1 nm
Quad4	1 nm, optical anchor, plus seismometer on ground, to keep in line, See IR section
Skew q1	Conventional mount
Sext1	Conventional mount
Sext2	Conventional mount
Bend1	Conventional mount
Drift:	
All	Rf BPMs or orbit stabilization system. See Sec. 11.5.6. Thermal stabilization of tunnel ($<$0.1 μm/hr achieved in FFTB)
Drift (roll):	
All	Rf BPMs or orbit stabilization system. See Sec. 11.5.6.

Table 11-20. *Tolerances for magnet types from Table 11-19 and how they can be achieved.*

The lattice response function $G(k)$ for the present NLC final-focus design is shown in Figure 11-33. For large wave numbers k (above 0.6 m^{-1}) the response function $G(k)$ is about constant, equal to ten. The oscillations are due to the discrete distances between the supports of the four final-doublet magnets. The calculation assumes that all magnets are supported beneath their center.

For small values of the wave number k, that is for $k < 0.01$ m^{-1}, $G(k)$ increases as the sixth power of k. There are three reasons for this: First and second, each side of the final-focus optics fulfills the Irwin sum rules [Irwin 1995a]:

$$-\sum_i k_i R_{34}^{i\to IP} + R_{33}^{e\to IP} = 1 \tag{11.46}$$

$$-\sum_i k_i R_{34}^{i\to IP} s_i + R_{33}^{e\to IP} s_e = 0, \tag{11.47}$$

where Eq. (11.46) expresses that a constant displacement of the entire system, including incoming beam orbit, has no effect on the beam-beam offset at the IP, and Eq. 11.47 means that a displacement of constant slope also does not affect the beam-beam offset. Note that, in Eq. 11.47, the betatron phase advance from the final-focus entrance e to the IP was

Name	Tolerance	$\Delta L/L$ (%)
Quads	hor. vibr.	0.33
	vert. vibr.	0.50
	hor. stab.	1.00
	vert. stab.	1.00
	field ripple	0.50
	roll	1.10
	Total:	**4.23**
Final Doublet	hor. vibr.	0.36
	vert. vibr.	0.50
	hor. stab.	0.95
	vert. stab.	1.00
	field ripple	0.38
	roll	0.96
	Total:	**3.90**
Sextupoles	hor. vibr.	0.03
	vert. vibr.	0.05
	field ripple	2.40
	Total:	**2.50**
Bends	field ripple	0.69
	field stability	0.67
	roll vibration	0.24
	roll drift	0.13
	Total:	**1.71**
Overall	**Total:**	**10.9**

Table 11-21. *Overall luminosity loss corresponding to imposed magnet tolerances. Reference design luminosity (without pinch) is* 0.96×10^{34} *cm*$^{-2}$ *s*$^{-1}$*. Numbers correspond to tolerances listed in Table 11-19.*

assumed to be a multiple of π, but this assumption is not essential. The two conditions (11.46) and (11.47) imply at least a k^4 behavior at small k. In addition and third, the final-focus system consists of several paired π-modules, *i.e.*, of identical modules with phase advance π on either side of the IP. For large wavelengths, the effects of these modules cancel each other, giving rise to the observed k^6 asymptotics. In an intermediate wave number range, between 0.01 and 0.6 m^{-1}, the response function increases about linearly with k (Figure 11-33).

The frequency-dependent correlation of ground motion at different positions can be characterized by

$$R(\omega, \Delta s) \equiv \int_0^\infty \frac{dk}{2\pi} \mu(\omega, k)(1 - \cos k\Delta s) \tag{11.48}$$

where

$$\mu(\omega, k) \equiv \frac{P(\omega, k)}{P(\omega)}, \tag{11.49}$$

with

$$P(\omega) \equiv \int \frac{dk}{2\pi} P(\omega, k) \tag{11.50}$$

representing the local power density of the ground motion. The function $R(\omega, \Delta s)$ can be extracted from recent ground motion measurements at SLAC, performed with two STS-2 seismic sensors on the linac tunnel floor. The

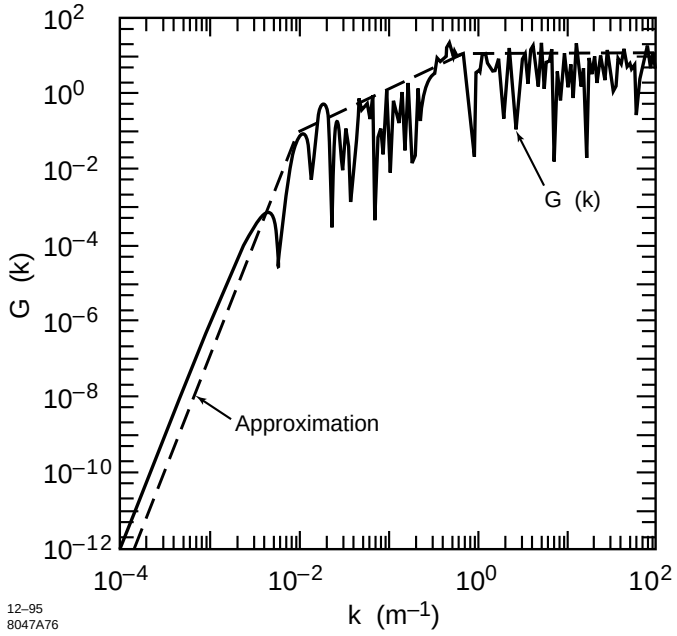


Figure 11-33. *Lattice response function for the NLC final-focus system.*

measurements suggest that R is parameterized quite well by the expression [Adolphsen 1995]

$$R(\omega, \Delta s) \approx 1 - J_0\left(k(\omega)\Delta s\right) \tag{11.51}$$

where J_0 denotes the zeroth-order Bessel function, $k(\omega) = \omega/v(\omega)$ with

$$v(\omega)\ [\,\mathrm{m\ s^{-1}}] \approx 450 + 1900 \exp\left(-\frac{\omega}{4\pi}\right) \tag{11.52}$$

and SI units are used throughout. The quantity $v(\omega)$ may be interpreted as the velocity of ground waves at frequency $f = \omega/(2\pi)$. The correlation function $R(\omega, \Delta s)$ is converted into the two-dimensional power spectrum $P(\omega, k)$ via Fourier transform:

$$P(\omega, k) = 4P(\omega) \times \int_0^\infty (1 - R(\omega, L)) \cos(kL) dL \cdot \tag{11.53}$$

or

$$P(\omega, k) = \begin{cases} \frac{4}{\sqrt{k(\omega)^2 - k^2}} P(\omega) & \text{if} \quad k(\omega) > k \\ 0 & \text{else} \end{cases} . \tag{11.54}$$

Interestingly, the same functional dependence is obtained if the ground motion is modeled by an ensemble of isotropic plane surface waves, for which each frequency is associated with a certain wavelength; longer wavelengths corresponding to lower frequencies [Irwin 1995b].

According to measurements at various locations [Juravlev 1995], a reasonable approximation to $P(\omega)$ for a "quiet" site is

$$P(\omega)[\,\mu\mathrm{m}^2/\mathrm{Hz}] \approx \frac{15.6 \cdot 10^{-3}}{\omega^4}. \tag{11.55}$$

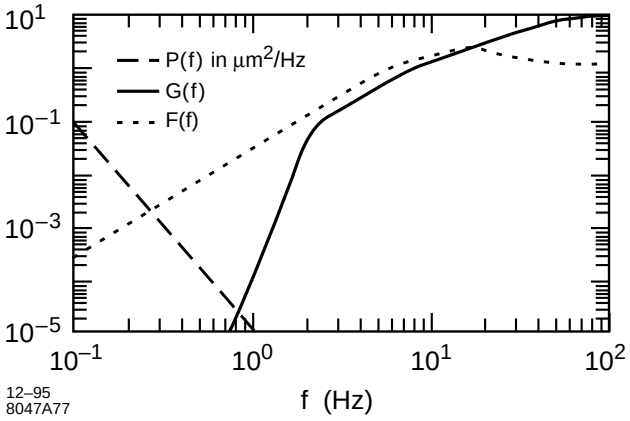


Figure 11-34. *Three functions which determine the rms beam-beam separation due to plane-wave ground motion: $F(f)$ — the feedback response [Hendrickson 1995]; $P(f)$ — the local power density; $G(f) \equiv \tilde{G}(2\pi f)$—defined by Eq. 11.57. The integral over the product of these functions gives the square of the rms beam-beam separation.*

where the angular frequency ω is given in s^{-1}, and only frequencies $\omega > 0$ are considered. The power density $P(\omega)$ decreases rapidly at high frequencies, while at low frequencies the response to ground motion is suppressed by the lattice-response function $G(k)$ in conjunction with the dispersion relation, Eq. (11.54). A further reduction may be achieved by orbit-feedback systems.

The ground-motion spectral density, Eq. (11.54), determines the rms separation $\Delta y_{\rm rms}$ of the two beams at the IP, as caused by the plane-wave ground motion. The rms separation can be written as

$$\Delta y_{\rm rms}^2 = \int_0^\infty \frac{d\omega}{2\pi} \tilde{G}(\omega) P(\omega) F(\omega), \tag{11.56}$$

where $F(\omega)$ denotes the feedback response, $P(\omega)$ is the power density, Eq. (11.55), and $\tilde{G}(\omega)$ represents the lattice response function converted into the frequency domain. It is defined as an integral of the product of $G(k)$ and the k-dependent part of $P(\omega, k)$ over k, *i.e.*,

$$\tilde{G}(\omega) \equiv \int_0^{k(\omega)} G(k) \frac{4}{\sqrt{k(\omega)^2 - k^2}} \frac{1}{2\pi}\, dk \quad . \tag{11.57}$$

The three functions $P(\omega)$, $\tilde{G}(\omega)$, and $F(\omega)$ are illustrated in Figure 11-34. The shape of $\tilde{G}(\omega)$ shows that the effect of ground motion is strongly suppressed by the lattice response. Numerical integration yields a value for the rms beam-beam separation of

$$\Delta y_{\rm rms} \approx 0.2 \text{ nm} \qquad \text{(with feedback)} \tag{11.58}$$

$$\Delta y_{\rm rms} \approx 0.3 \text{ nm} \qquad \text{(without feedback)} \tag{11.59}$$

and, thus, the plane-wave ground motion is insignificant, provided the magnet centers move exactly as the ground beneath them. If plane-wave ground motion were the only source of orbit-perturbation, no orbit-feedback would be needed at a quiet site. Above 6 Hz, the power spectrum of the SLAC measurements is considerably larger than that of Eq. (11.55), due to resonances of the linac support structures and also due to cultural noise. Assuming the SLAC power spectrum, one finds an expected rms beam-beam separation of about 1.1–1.3 nm, which still seems acceptable, but which is considerably larger than the 0.3 nm for a quiet site.

Some authors have argued that at low frequencies a component of ground motion exists which cannot be cast into the above framework, and which is of pure random character [Juravlev 1995]. They characterize this part of the ground motion by a so-called ATL law, according to which the change of the average squared relative displacement of two points is proportional to the distance between the points and to time. (An entirely different interpretation of ground motion as a systematic process was suggested by R. Pitthan [Pitthan 1995].) The two-dimensional spectral density describing the ATL law can be written

$$P_{ATL}(\omega, k) = \frac{4A}{\omega^2 k^2} \tag{11.60}$$

where $k > 0, \omega > 0$ is assumed, and $A \approx 10^{-8}$–10^{-5} μm^2 s^{-1} m^{-1} depending on location and on relevant timescale. The SLAC measurements over a timescale of seconds indicate $A < 6 \times 10^{-7}$ μm^2 s^{-1} m^{-1}.

Folding in the feedback response function, it is possible to determine the rms beam-beam separation due to ATL-like ground motion in the frequency range 0–0.01 Hz, where the ATL law might be applicable. For $A = 10^{-6}$ μm^2 m^{-1} s^{-1}, this rms separation is found to be 15 pm. Assuming as a worst case, that the size of uncorrelated motion is equal to the noise floor of the STS seismometer, the rms separation in the frequency range of 0.01–6 Hz is estimated to be no larger than 242 pm, and the contribution from frequencies above 6 Hz is about 124 pm at a quiet site. In total then, the rms separation from uncorrelated ATL-like ground motion does not exceed 0.3 nm, and the entire separation due to all ground motion (ATL-like and plane-wave motion) at a quiet site is about 0.4 nm, which would result in a luminosity loss of only 0.1%. The conclusion is that one can use the ground (bedrock) as a reference for stabilization.

Thus, magnet supports need to be designed which neither amplify nor damp the ground motion, but couple the magnet firmly to the ground. As discussed previously, the tightest tolerances on uncorrelated quadrupole-to-ground vibration above 5 Hz (vibration at lower frequency is compensated by the orbit feedback) are 1 nm (rms) for the final doublet and 3–4 nm for a few other quadrupoles (Table 11-19). These tolerances correspond to a total luminosity loss of about 2%. At DESY (SBLC-TF), relative rms vibrations above 1 Hz were measured to be smaller than 1 nm, which would meet all the NLC tolerances. Quadrupoles at the FFTB were found to vibrate by about 4 nm with respect to the ground, excited mainly by bad cooling pumps. These vibration amplitudes would still satisfy the tolerance criteria for all NLC magnets other than the final doublet. For the latter a special stabilization system based on an optical anchor and piezo-electric movers has been devised, which is described in Chapter 12.

Capture Tolerances and Tuning Frequency

The magnet pre-alignment tolerances required to steer the first beam through the final focus onto the dump will be loose. They are essentially determined by the magnet-mover ranges. A pre-alignment with an accuracy of 500 μm is expected to be sufficient. This is easily achieved by standard alignment techniques, which provide a typical accuracy of 100–200 μm. Once the beam reaches the NLC main dump, the actual alignment is performed using beam-based techniques, similar to those employed at the SLC or FFTB [Emma 1992, Raimondi 1993, Tenenbaum 1995].

More interesting is the question how often the beam-based alignment has to be performed, and how accurate it should be. The answer to this depends on the maximum range of the tuning elements, and on the IP spot-size increase due to ground motion. The recent progress in understanding ground motion (see previous section and Appendix C), and new work on tuning schemes shed light on this problem. A related question refers to the required frequency of aberration tuning and whether it is acceptable.

In complete analogy to the steering effect of ground motion, lattice-response functions may be defined for vertical dispersion and skew coupling, in an obvious way. The only difference is that, in this case, the two final foci on either side of the IP are treated independently. More precisely, the lattice-response functions are of the form

$$G_\delta(k) = \sum_{i,j} \mu_i^\delta \mu_j^\delta \cos(k \Delta s_{ij}) \tag{11.61}$$

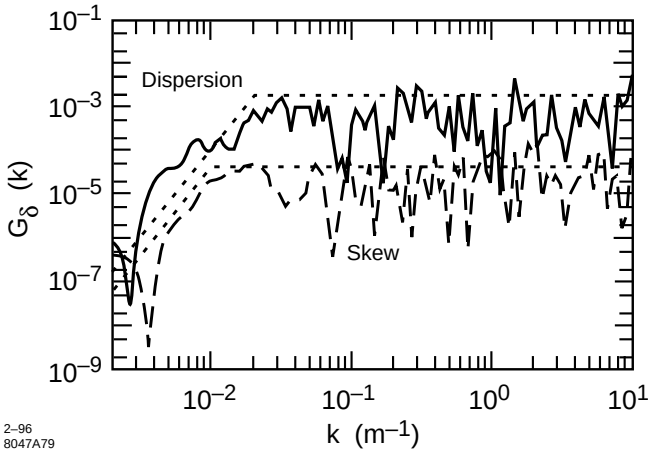


Figure 11-35. *Lattice response functions for vertical dispersion and skew coupling. Dotted lines represent approximations which are employed in the later analysis.*

where the subindices i, j run over only one side of the IP only, and the multipliers $\mu^{\delta}_{i,j}$ characterize the IP spot-size increase due to a displacement of quadrupole (or sextupole) i. The term Δs_{ij} denotes, as before, the distance between the two magnets i and j. The lattice response functions for dispersion and skew coupling are depicted in Figure 11-35, along with simple approximations which are used in the following studies. If the wavelengths are small, *i.e.*, at large k, the response is about constant, since the magnets move incoherently. For small k, the lattice response is strongly suppressed and decreases as the 4th power of k. As for the steering effect discussed earlier, this suppression represents the invariance of the luminosity with regard to displacements and tilts of the entire beam line. Unlike the steering, however, there is no cancellation between the two sides of the IP or between paired $-I$ modules, and this is the reason why the response here decreases as the 4th power of k, instead of k^6.

The spot-size increase due to ATL ground motion as a function of time T is described by the equation

$$< \Delta\sigma_y(T)^2 >= 2 \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi} G_\delta(k) P_{ATL}(\omega, k)(1 - \cos(\omega T)) \tag{11.62}$$

where $G_\delta(k)$ denotes the lattice-response function for dispersion or skew coupling (Figure 11-35). Figure 11-36 presents the increase of the spot size due to dispersion as a function of time, for three different magnitudes of ATL-type ground motion.

Even though this figure does not assume any orbit stabilization other than a beam-centroid correction at the IP, there would still be several minutes time to correct the dispersion. In reality, feedback systems which stabilize the orbit will reduce the generated dispersion. If the frequency-response curve of a typical SLC-style orbit feedback is included in the calculation, an ATL coefficient $A = 10^{-6}$ μm^2 m^{-1} s^{-1} is assumed, and the product spectrum is integrated over frequencies below 0.01 Hz, the spot size is found to reach a maximum value after about 100s and then to stay constant. The spot-size increase at that moment is less than a picometer, to be added in quadrature. Thus the effect of ATL-like ground motion on the IP spot size appears to be insignificant, assuming the orbit is stabilized by a typical SLC-style feedback system. The frequency at which the different aberrations need to be tuned, in order to keep the average luminosity loss per aberration below 1–2%, is thus not determined by the ground motion, but by the stability of the feedback BPMs. If the BPM stability were the same as measured in the FFTB, the aberrations would have to be tuned less than once per hour. Since the NLC BPMs will be located underground, and temperature and humidity much better controlled, the BPM stability is expected to surpass that in the FFTB.

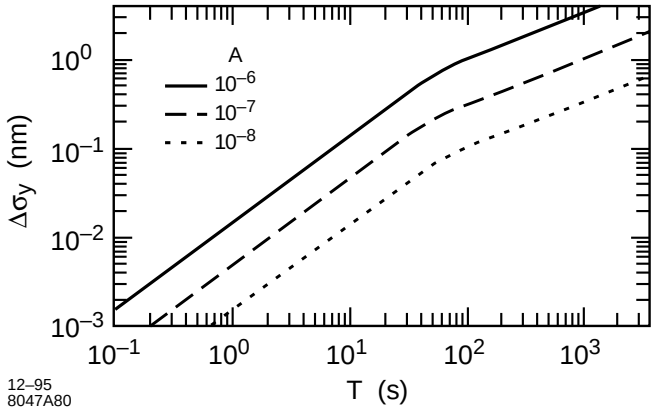


Figure 11-36. *Incoherent spot-size increase due to vertical dispersion, caused by ATL ground motion, as a function of time, for three different values of A (in units of μm^2 s^{-1} m^{-1}), and assuming $\delta \approx 3 \times 10^{-3}$.*

Collimator Tolerances

As a supplement to the two-phase collimation section upstream of the big bend, a last collimation of the final-doublet betatron phase is performed in the final focus itself. Its purpose is to reduce the number of electrons which may hit the final doublet to below ten, so that the detector background remains acceptable. The final-focus collimators remove the beam halo which is newly generated by gas scattering, bremsstrahlung and inverse Compton scattering downstream of the collimation section and by edge scattering in the IP-phase collimation. The ideal position for final-focus collimators are the high-beta points in CCX and CCY. Therefore, the two horizontal and two vertical collimators will be located close to the main horizontal and vertical sextupoles, respectively.

In addition to these four tungsten collimators, the final focus accommodates almost 200 wide-aperture titanium spoilers. These are placed upstream of each quadrupole and each dipole section, and protect the magnets against damage by a missteered beam.

Collimators do not only improve the background, but, as SLC experiments have demonstrated [Bane 1995], they can also adversely affect the beam quality by generating resistive-wall and geometric wakefields. To reduce the geometric wake, most or all of the NLC collimators will be tapered at a shallow angle, as discussed in Chapter 9, Adjustable collimators are usually flat and consist of two jaws (top and bottom, or left and right). Protection collimators are round.

Specifically, there are four main effects of collimators:

- If the incoming beam executes a betatron oscillation, the collimator wakes will cause an amplification of the oscillation in the orthogonal betatron phase. This effect is called jitter amplification, and occurs even when the collimators are perfectly aligned. If the collimator is located in the final-doublet phase, the position change at the IP Δy caused by the collimator is

$$\frac{\Delta y}{\sigma_y^*} = \xi \frac{y'}{\sigma_{y'}} \tag{11.63}$$

  for an initial betatron oscillation of size y', where all coordinates and rms values refer to the IP. The coefficient ξ is the relative jitter amplification, and is expressed in terms of collimator and beam parameters as

$$\xi \equiv \frac{\beta_y^{\rm coll} 2Nr_e}{\gamma\sigma_z}\left[\left(8 < f_r >< g_r > \frac{(b^2-g^2)(b-g)}{b^3}\right)^{\frac{1}{2}}\left(\frac{\lambda\sigma_z}{g^6}\right)^{\frac{1}{4}} + \frac{L}{g^3} < f_r > \sqrt{\lambda\sigma_z}\right] \tag{11.64}$$

where $< f_r > \approx 0.3$, $< g_r > \approx 0.82$, g is the half-aperture, b the beam-pipe radius, $\lambda = \rho/(120\pi)$ (ρ resistivity in Ωm), and L the length of the flat part of the collimator.

- In addition to jitter amplification, there is also an emittance growth arising from the variation of the wakefield kick across the bunch. Typically, this effect is smaller than the former.
- If the collimator jaws (of a flat collimator) are misaligned with respect to the center of the beam, they will also cause emittance growth, even if the incoming orbit is unperturbed. This emittance growth gives rise to an alignment tolerance. For a spot size increase by 2% or less the alignment tolerance of a single collimator is

$$\Delta y \approx \frac{\sigma_y^{\rm coll}}{4.5\,\xi} \quad . \tag{11.65}$$

- Finally, even if a flat collimator is perfectly aligned, and the beam unperturbed, the collimator still delivers a quadrupole wake-field to the beam, changing the focusing of particles in the bunch-tail halo and causing an emittance dilution. The magnitude of the quadrupole wakefield kick from a flat collimator is about a third of that of the dipole wake. Thus, for an expected orbit jitter larger than 0.5 σ_y, the quadrupole effect is small compared with the dipole kick.

The main horizontal and vertical final-focus collimators are located at the high beta points in CCX and CCY, respectively, where $\beta_x \approx 70$ km or $\beta_y \approx 200$ km. If we assume $N = 10^{10}$ particles per bunch, and $\sigma_z = 100\,\mu$m, the jitter amplification ξ for a 7-cm-long (about 20 r.l.) W collimator, and collimation at 17 σ_x or 43 σ_y, is 0.06 (0.01) in $y(x)$. The vertical alignment tolerance, Eq. (11.65), for a 2% spot-size increase caused by the two vertical collimators is 1.6 mm; the horizontal alignment tolerance for the CCX collimators is even larger. Thus, the required alignment accuracy will not be determined by emittance growth, but by the collimation efficiency. The latter probably demands that the vertical collimators are aligned to within 150 μm ($1\sigma_y$).

Temperature

The tolerances on position and field strength translate into tolerances on the temperature stability of magnets and magnet supports. The tolerances on medium-term stability are considerably relieved by the existence of an orbit stabilization system (*i.e.,* feedback) which corrects orbit distortions resulting from quadrupole drifts, bend-field changes and bend rolls.

In the FFTB, quadrupole motion due to temperature variation was measured to be smaller than 100 nm/h. This stability is an order of magnitude better than what is needed for the NLC.

Vacuum

The required vacuum pressure is determined by the tolerable background in the detector. The following discussion shows that the largest contributions to the vacuum-related background are due to bremsstrahlung. At an average pressure of 10 nTorr, less than 10 particles per bunch train will hit apertures in the final-doublet region. The background from these particles needs to be compared with that expected from synchrotron radiation generated in the bending and quadrupole magnets, and with that caused by beamstrahlung and pair-creation at the IP.

Coulomb Scattering. If an electron undergoes hard Coulomb scattering on a residual-gas nucleus, it may hit the pole faces or inner beam pipe of the final doublet. Following Irwin (Chapter 9), the relative number of scattered

particles hitting the doublet is given by

$$\frac{\Delta N}{N} = 2\pi\rho\left(\frac{r_e}{\gamma a_D}\right)^2\left(\sum_i Z_i^2\right)\left[\int_{SX1b}^{QFTA} ds\ R_{12}^2(s) + \int_{SX2b}^{QFTA} ds\ R_{34}^2(s)\right] \tag{11.66}$$

where $R_{12}(s)$ and $R_{34}(s)$ denote the R-matrix elements from point s to the entrance face of QFTA, r_e the classical electron radius, and a_D the inner radius of QFTA (assumed as 5 mm). The integrations are performed between the last sextupole in CCX and CCY, respectively,—these are the collimator positions—and QFTA. The sum over i runs over the different atoms of a molecule, and $\rho \approx 3.2 \times 10^{13}\mathrm{m}^{-3}p/$ nTorr is the molecule density. Evaluation yields

$$\left[\int_{SX1b}^{QFTA} ds\ R_{12}^2(s) + \int_{SX2b}^{QFTA} ds\ R_{34}^2(s)\right] \approx 7 \times 10^6\ \mathrm{m}^3. \tag{11.67}$$

Assuming carbon monoxide molecules and a beam energy of 500 GeV, we find

$$\frac{\Delta N}{N} \approx 4 \times 10^{-14}\frac{p}{\mathrm{nTorr}} \quad . \tag{11.68}$$

If we allow 10 electrons in each bunch train of 6×10^{11} particles to hit the doublet in the absence of final-focus collimators, a pressure of 400 nTorr is sufficient. The collimators in the final-focus system will intercept most of the scattered particles.

Bremsstrahlung. An electron may also get lost due to bremsstrahlung in the field of a nucleus. The cross section for an energy loss between ΔE_1 and ΔE_2 is [Piwinski 1985]

$$\sigma_{\mathrm{brems}} \approx \sum_i \frac{4r_E^2 Z_i^2}{137}\frac{4}{3}\ln\left(\frac{\Delta E_2}{\Delta E_1}\right)\ln\left(\frac{183}{Z_i^{\frac{1}{3}}}\right) \quad . \tag{11.69}$$

For CO molecules and an energy loss between 6% and 25%, we find $\sigma_{\mathrm{brems}} \approx 2$ barn. The fraction of particles suffering this energy loss is given by

$$\frac{\Delta N}{N} = \rho L \sigma_{\mathrm{brems}} \approx 6.4 \times 10^{-12}\frac{p}{\mathrm{nTorr}} \tag{11.70}$$

where $L \approx 1000$ m. In order that less than 10 particles per train lose 6–25% of their energy, we need to achieve a pressure of 4 nTorr. This value is two orders of magnitude smaller than that calculated before for particle loss due to hard Coulomb scattering. However, a large fraction of the electrons lost by bremsstrahlung do not hit the final-doublet apertures and will not contribute to the detector background. As a conservative estimate, we assume a pressure of 10 nTorr to be sufficient.

Inelastic Scattering. For simplicity, let us confine ourselves to hydrogen molecules, *i.e.*, we consider electron-proton collisions. In this case, using the Weizsäcker-Williams approximation, the differential cross section for leptoproduction $d\sigma_{ep}/dy$ reads [ZEUS/H1]

$$\frac{d\sigma_{ep}}{dy} = \sigma_{tot}^{\gamma p}(W_{\gamma p})\frac{\alpha}{2\pi}\frac{1+(1-y)^2}{y}\ln\frac{Q^2_{\max}}{Q^2_{\min}} \tag{11.71}$$

where $y \equiv 1 - E'/E$ is the photon energy in units of the incident electron energy. Approximating $\sigma_{tot}^{\gamma p} \approx 150\ \mu$barn, $Q^2_{\max} \approx 4E^2$, $Q^2_{min} = (m_e c^2 y)^2/(1-y)$, and integrating over photon energies from 500 keV to 250 GeV, we derive an upper bound for the relevant total cross section:

$$\sigma_{ep} \leq 180\ \mu\mathrm{barn} \quad . \tag{11.72}$$

Over a distance l of 100 m, the fraction of electrons undergoing inelastic reactions with the residual gas is then

$$\frac{\Delta N}{N} \approx \sigma_{ep} l \rho \leq 6 \times 10^{-17} \frac{p}{\text{nTorr}} \tag{11.73}$$

or $\Delta N \leq 0.003$ per bunch train for a pressure of 100 nTorr—a very small number.

Other Sources of Background

Inverse Compton Scattering. Electrons (or positrons) do not only scatter off residual-gas atoms, but can also suffer inverse Compton scattering on thermal radiation photons. The latter effect can limit the beam lifetime in high-energy electron storage rings [Telnov 1987], and it has been measured at LEP [Bini 1991] and at HERA [Lomperski 1993]. In this section, we estimate its importance for the NLC final-focus system.

The density of thermal photons increases as the third power of the temperature. For a temperature of 300 K, it is about

$$n_\gamma \approx 5 \times 10^{14}\ \text{m}^{-3}, \tag{11.74}$$

and thus equal to the density of residual-gas atoms at a pressure of 17 nTorr. The total Compton cross section can be expressed in terms of the dimensionless parameter x [Telnov 1987].

$$x \equiv \frac{4 E \omega_0}{m_e^2 c^4} \cos^2 \frac{\alpha_0}{2}, \tag{11.75}$$

where α_0 denotes the angle at which photons and electrons collide in the laboratory frame ($\alpha_0 = 0$ means head-on collision), E the beam energy, and ω_0 the photon energy. Assuming 500 GeV beam energy, an average photon energy of $\bar{\omega}_0 = 0.07$ eV and $\alpha_0 = 0$, we find $x \approx 0.56$, and a total Compton cross section not much different from the Thomson cross section $\sigma_t \approx 0.7$ barn. This cross section is an order of magnitude smaller than that for beam loss due to bremsstrahlung on residual-gas nuclei. However, the maximum energy loss of a Compton-scattered electron is quite large,

$$\left.\frac{\Delta E}{E}\right|_{\max} = \frac{x}{1+x} \approx 36\%, \tag{11.76}$$

which implies that almost all scattered electrons are lost so rapidly, that they do not make it to the final doublet. The total number of particles lost per bunch train is estimated as

$$\Delta N \approx \sigma_T L n_\gamma N \approx 36. \tag{11.77}$$

Here $N \approx 6 \times 10^{11}$ is the number of particles in the bunch train, and $L \approx 2000$ m the length of the final-focus system. The number of electrons suffering inverse Compton scattering on thermal photons is at least comparable to, if not larger than those lost due to bremsstrahlung or Coulomb scattering off the residual gas.

Touschek Effect. The scattering of two particles inside the same bunch off each other may lead to a particle loss due to the introduced change of longitudinal momentum. This effect limits the beam lifetime in many electron storage rings. It was first observed at AdA, in 1963 [Bernadini 1963].

The number of electrons scattered to energies larger than $\eta \equiv \left(\frac{\Delta E}{E}\right)_{\max}$ is given by [Walker 1987]

$$\Delta N = \frac{r_e^2 n_b N_b^2 L}{\gamma^2 \left.\frac{\Delta E}{E}\right|_{\max} 2\pi^{\frac{1}{2}} \sigma_x \sigma_y \sigma_z} J(\eta, \delta q) \quad . \tag{11.78}$$

Here, δq denotes the rms transverse momentum in units of $m_e c$, r_e the classical electron radius, n_b the number of bunches, N_b the number of particles per bunch, and L the length of the beam line considered. The loss rate is proportional to the second power of N_b, and inversely proportional to the square of the energy and to the beam size. An approximation to the function $J(\eta, \delta q)$ for $\delta q/\eta \geq 100$ is given in Eq. (7) of ref. [Walker 1987]. For the parameters of interest here, J is nearly equal to 1.

In the NLC final focus at 500-GeV-c.m. energy, transverse beam size and transverse momentum are $< \delta q >= 6.4\ \mu\text{rad}\times\gamma \approx 3.2$ and $< \sigma_x\sigma_y > \approx 148\ \mu\text{m}^2$. Assuming $N_b \approx 8.5 \times 10^9$ and $n_b = 90$, the number of scattered particles per bunch train is:

$$\Delta N \approx 2 \text{ for } \Delta E/E \geq 0.5\%, \qquad \text{and} \qquad \Delta N \approx 0.1 \text{ for } \Delta E/E \geq 2\% \tag{11.79}$$

This is a rather small effect.

11.5.6 Feedback and Stability

Overview of Feedback Systems

Several orbit feedback systems are envisioned to stabilize the beam orbit throughout the beam delivery system. In addition to feedbacks in the collimation section, the big bend, and the diagnostics section, there will be a "launch" feedback correcting the orbit in the beta-match section and CCX, and a feedback controlling the CCY and FT orbit. Each feedback system will use orbit readings from at least four BPMs to correct the orbit by moving quadrupoles of by adjusting dipole correctors. The CCY- and FT-orbit feedback, which is the most important of these feedback loops, is discussed in the next section.

A fast IP collision feedback is essential for correcting beam position and preserving collisions at the IP. This feedback will be very similar to the IP feedback at the SLC, which uses fast dither coils to maintain head-on collisions. The required dither changes are deduced from the measured beam-beam kick angle and the slope of the beam-beam deflection curve. The main difference to the SLC is that, because of the smallness of the required corrector strengths, the NLC feedback will use electrostatic vertical dither coils rather than magnetic. In addition, it may be possible to use a device similar to the crab cavity for compensating orbit changes along the bunch train, should that be desired.

Regarding control of emittance or spot size, a slow automatic skew-correction feedback is being contemplated, which may be based on DS wire scans and SCS skew correctors. Such a feedback will only be necessary if there are significant changes to the incoming x-y coupling over a few hours, and it would be straightforward to implement.

CCY and FT Orbit Feedback

The IP beam size is extremely sensitive to orbit variations which originate within the final focus. For example, a vertical displacement of QFT3 (see location on Figure 11-37; note that the design and the magnet names have changed since this was written and drawn!) by 44 nm kicks the vertical orbit in QFT1 by ∼200 nm which, for 0.3% rms energy spread, generates enough vertical dispersion at the IP to increase the beam size by 2%. A similar effect occurs for quadrupole displacements within the CCY where the orbit is steered off in the second CCY sextupole (SY1B). For these reasons it will be important to monitor and correct local orbit variations in the CCY and FT to a very high precision. The following describes a conceptual monitor and feedback system for CCY and FT local orbit distortions. Throughout, it is assumed that the IP beam-beam steering errors introduced by these orbit variations are corrected with a much faster (∼30 pulse) IP collision feedback which is not addressed here.

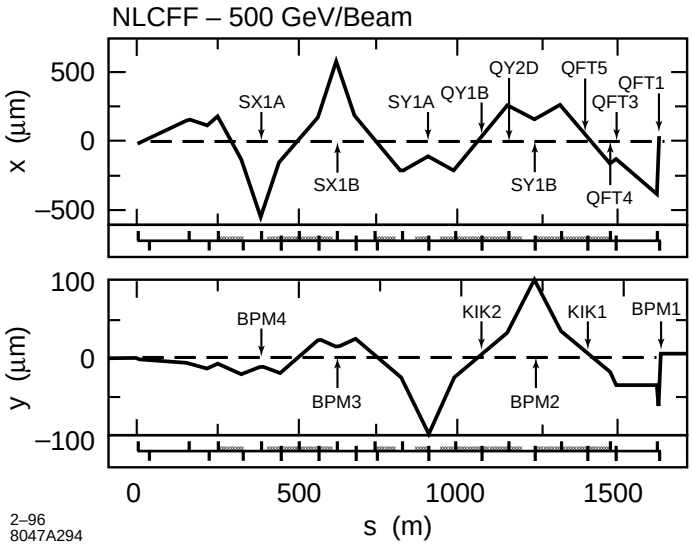


Figure 11-37. *Horizontal and vertical one-sigma IP angle betatron oscillation through the final focus (i.e., at the IP:* $\Delta x = 0$, $\Delta y = 0$, $\Delta x' = \sigma'_x$, $\Delta y' = \sigma'_y$*).*

The vertical beam size at the final doublet is quite large in comparison to the required orbit monitoring precision (60 μm compared with 200 nm). Therefore, local orbit distortion monitoring will be made in the presence of a large background signal of incoming trajectory jitter—which has much less impact on the IP beam size. The feedback system must suppress this incoming jitter signal reliably so that non-jitter related y-BPM reading variations of $\sim$200 nm are discernible. In order to accomplish this the feedback system must monitor BPMs in the CCX, CCY and FT. In the simplest case, two BPMs are placed in the CCX—each one adjacent to an SX1 sextupole (SX1A and SX1B)—in order to monitor both betatron and energy jitter (BPM-4 and BPM-3). A third BPM is placed in the CCY adjacent to the SY1B sextupole (BPM-2) to monitor CCY induced orbit changes, and a fourth BPM is placed just upstream of QFT1 (BPM-1) to monitor FT induced orbit changes. More BPMs may be added for redundancy.

A simplifying characteristic of the final focus is that only the IP angle betatron phase is visible with BPMs throughout the CCX, CCY and FT—a one-sigma IP angle oscillation generates a 60-μm BPM-1 reading while a one-sigma IP position oscillation generates a $\sim$3-nm BPM-1 reading. Figure 11-37 shows a one-sigma betatron oscillation at the IP angle phase in x and y, while Figure 11-38 shows a one-sigma oscillation at the IP position phase. The invisibility of the IP position betatron oscillation is clearly demonstrated.

For this reason each BPM reading can be broken down into a component of incoming IP angle jitter, energy jitter (for horizontal plane or vertical dispersion error), a CCY orbit kick, an FT orbit kick and a static offset. All other sources of BPM reading variations need to be $<$200 nm (BPM-1) over the scale of a few hours (time between re-tuning). A simple algorithm can then be devised where the four (or more) BPMs are used in a linear combination to determine the CCY- and FT-induced orbit changes so that corrections may be applied in both the CCY and the FT. The linear combination is determined by using the normal trajectory jitter to calibrate the BPM to BPM coefficients (assuming local orbit distortions are constant over the $\sim$200 pulses necessary for calibration). Each ith BPM is then fitted to the form

$$y_i(j) = \eta_i \cdot \delta(j) + \alpha_i \cdot y_\beta(j) + c_i \quad , \tag{11.80}$$

where $\delta(j)$ is the fractional energy deviation of the jth pulse, y_β is the incoming betatron amplitude (proportional to IP angle), and the three coefficients, α_i, η_i and c_i are the fit results. The energy deviation and betatron amplitude can then, given their $-I_{4\times4}$ transfer matrix separation, be conveniently parameterized as follows (actually any linear

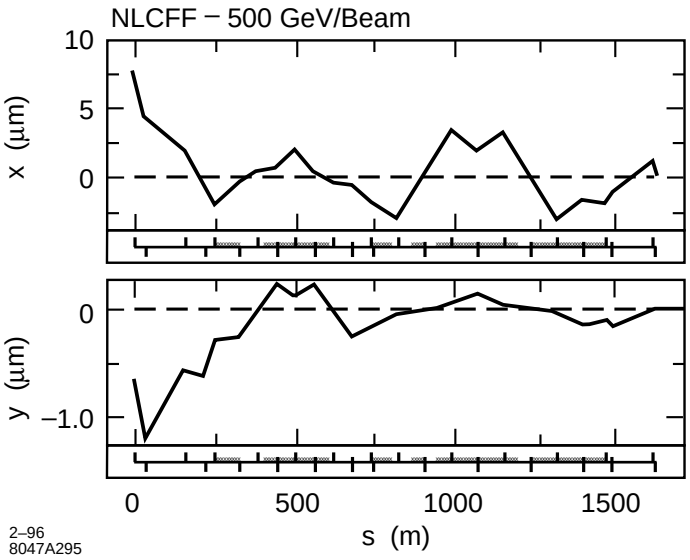


Figure 11-38. *Horizontal and vertical one-sigma IP position betatron oscillation through the final focus (i.e., at the IP:* $\Delta x = \sigma_x$, $\Delta y = \sigma_y$, $\Delta x' = 0$, $\Delta y' = 0$*).*

combination will suffice),

$$\delta(j) \equiv \frac{y_4(j) + y_3(j)}{R_{16}} \quad , y_\beta(j) \equiv \frac{y_4(j) - y_3(j)}{2} \quad . \tag{11.81}$$

where R_{16} ($\sim$150 mm) is a linear transfer matrix element between BPM-4 and 3.

This calibration is then performed after each IP re-tune. After a time necessary to generate local orbit distortions (must not be less than $\sim$10 seconds) a calculation is made using $\sim$200 more pulses and the same coefficients.

$$\begin{bmatrix} y_4 - c_4 \\ y_3 - c_3 \\ y_2 - c_2 \\ y_1 - c_1 \end{bmatrix} = \begin{bmatrix} \eta_4 & \alpha_4 & 0 & 0 \\ \eta_3 & \alpha_3 & 0 & 0 \\ \eta_2 & \alpha_2 & 1 & 0 \\ \eta_1 & \alpha_1 & \alpha_1/\alpha_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta \\ y_\beta \\ \Delta y_2 \\ \Delta y_1 \end{bmatrix} \tag{11.82}$$

Here the previously determined coefficients—assumed constant over the hour timescale—make up the matrix and the critical orbit distortion results, Δ_{y2} and Δ_{y1}, represent the local orbit distortions at BPM-2 (the second CCY sextupole SY1) and BPM-1 (the last quadrupole QFT1). This treatment assumes that all local orbit distortions are along the IP angle phase. This is generally true for the quadrupoles with the tightest tolerances (see Table 11-14). Corrections are then applied in order to maintain the CCY and FT orbit with respect to the incoming orbit. This is the only point in the algorithm where some knowledge of short sections of the optics is required. Static optical errors and BPM gain errors have been removed in the calibration process.

Simple simulations of this feedback have been run which include random BPM noise, IP position and angle jitter, energy jitter, BPM gain errors, BPM non-linearities and BPM offsets. Table 11-22 summarizes the simulation conditions.

For the simulation, two CCY quadrupoles and two FT quadrupoles (QY1B, QY2D, QFT4 and QFT3) were each displaced by 1 μm which, for QFT3, is >20 times tolerance. In the case of adjacent and opposite strength quadrupoles such as QFT3 and QFT4 the misalignments chosen were the worst case combination (*i.e.,* QFT3 $\Delta_{x,y} = +1\,\mu$m and

Number of pulses (cal. & cor.)	IP pos. rms jitter ($\sigma^*_{x,y}$)	IP ang. rms jitter ($\sigma^*_{x',y'}$)	Energy rms jitter (σ_δ)	BPM rms noise (μm)	BPM rms gain errors (%)	BPM non-linearity (% at 1 mm)	BPM rms offsets (μm)
200	0.5	0.5	0.25	1	5	2	250

Table 11-22. *Simulation conditions for CCY and FT local orbit distortion feedback system. With these conditions the feedback simulation successfully restored the IP beam size to within 2% for two CCY and two FT quadrupoles drifting 5 times their tolerance in x and y.*

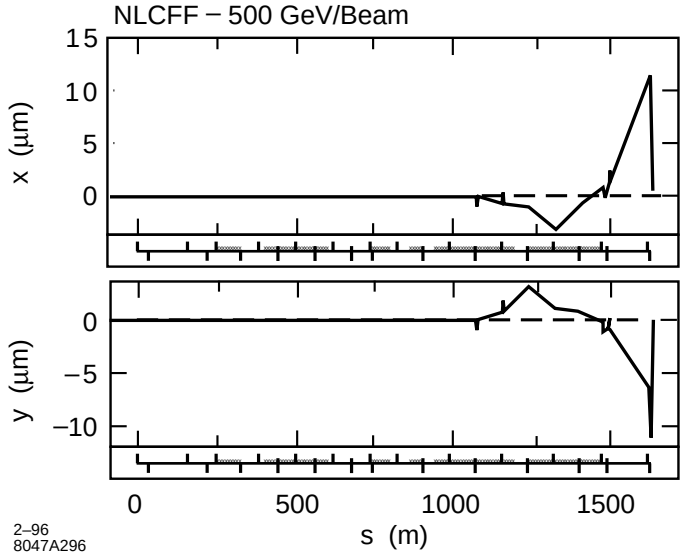


Figure 11-39. *Horizontal and vertical orbit through final focus after 1-μm alignment drift of four quadrupoles (before feedback correction). The luminosity is degraded by a factor of 10.*

QFT4 $\Delta_{x,y} = -1\,\mu$m). These sudden, uncompensated misalignments dramatically reduce the luminosity by a factor of ∼10 (assuming this drift occurs in only one of the two final-focus beam lines). In this case the vertical IP beam size increases to ∼13 times larger than nominal (2.5 nm) while the horizontal is only 6% larger. For the simulation, correctors were placed with one x and y pair (KIK2) at QY1B and one pair (KIK1) at QFT5 (each $\sim n\pi$ from the IP). Only one correction interval was applied (*i.e.*, no iterations were allowed). For these conditions the feedback successfully restored the luminosity to within 1% of nominal. Since there are only two correctors per plane and four misaligned quadrupoles the orbit is not everywhere restored. It is restored primarily at QFT1 (BPM-1) and SY1B (BPM-2) which, for this drift magnitude, is adequate. Figure 11-39 shows the difference orbit through the final focus, including the effect of the four displaced quadrupoles, before feedback correction is applied. Figure 11-40 shows the orbit after feedback correction. The horizontal feedback for this case is hardly necessary. However, further simulations have shown the horizontal feedback to work well at scales of up to ∼10 times the horizontal drift tolerances for these quadrupoles (10–30 μm).

The corrector strengths here are too small to be reliably applied by a simple dipole magnet (∼0.2 Gauss for a 10-cm length dipole). However, at this timescale (>10 seconds) a quadrupole mover could be used. In this case a 19-μm and 1.4-μm vertical displacement of QFT5 and QY1B, respectively will provide the necessary correction.

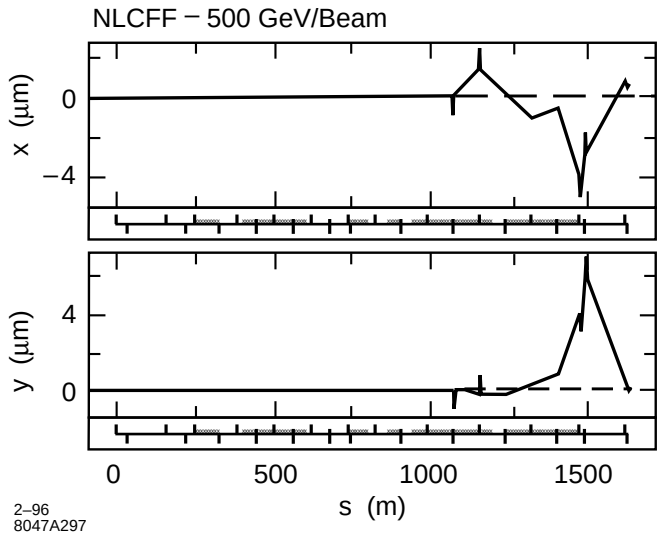


Figure 11-40. *Orbit through final focus after feedback correction. The feedback primarily restores the orbit at the final doublet and second y-sextupole. With more drifted quadrupoles than correctors, some orbit errors remain. The luminosity is restored to within 1% of nominal.*

Note that a BPM with 1-μm rms random noise was used successfully in this case. However, BPM offsets need to remain constant to $<$200-nm drift over a few hours which is probably the most challenging requirement and may be somewhat inconsistent with a 1-μm resolution BPM.

If high-resolution rf BPMs are required (a few nanometers), the dynamic range of these BPMs is limited, so that, in order to determine and correct magnet motion in the final focus, their readbacks must be averaged over a subset of orbits for which the incoming orbit jitter is small.

For these studies we have chosen to place only one corrector (per plane) in the CCY and one in the FT. It may also be possible to use more correctors so that the corrections are more closely associated with their original kick. This scheme probably allows a wider correction range, however it places much more stringent requirements on BPM resolution requirements in order that individual quadrupole drifts become measurably separable. This possibility has not been studied.

Finally, the linearity of the final-focus optics over both energy and betatron variations has been verified to be adequate (*i.e.,* $<$200-nm orbit change at BPM-1) over energy deviations of up to $\pm$0.25% and betatron oscillations more than twice the beam size. Therefore, if sampled orbit data is cut at $\pm$0.1% in energy deviation and $\pm$1-sigma in betatron jitter, the optics is completely linear.

The feedback system explored here looks promising. In the limited conditions simulated here the extraordinarily tight alignment drift tolerances in the final transformer can be greatly loosened with a local orbit feedback system by more than a factor of 20 which sets the quadrupole drift scale tolerances at the more livable level of 1 μm rather than 50 nm. Some of the important remaining issues not addressed here are summarized below.

- BPM offsets and gains need to be stable to $<$200 nm over a few hours.
- Tails in the beam distribution must not introduce a systematic position error.
- Optics need to be stable enough over a few hours (can be measured).

- BPMs must be insensitive to particle backgrounds.

11.5.7 Operations and Controls

Control System and Software

In general, the NLC control system will be based on the SLC experience and will adopt many grown features of the SLC control system. In certain aspects, however, it needs to surpass the SLC control system. Three features are particularly desirable: The control system has to be

- flexible, *i.e.,* accommodate to any unforeseen or changing demands,
- easy to modify and to augment, *i.e.,* allow accelerator physicists and operators to write applications themselves on the fly, and
- fast, *i.e.,* contain the luminosity loss due to tuning, magnet trimming, etc.

The last item is very important. As an illustration, let us consider the time required for scanning an aberration: if we only count a minimum number of 50 pulses per beam-beam deflection scan and take a settle time of 200 ms equal to four magnet time constants for each magnet trimming, we find that a seven-step aberration scan can be performed in less than 5 s. Since during an aberration scan the luminosity is close to zero, a tuning time of 5 s results in an average luminosity loss of 1.0%, assuming 10 scans per hour. To contain the luminosity loss, the processing time of the control system during a scan should be shorter than the 5-s time needed for tuning and steering. At the SLC, a tuning scan presently takes about a minute, which is almost entirely devoted to processing by the control system. The NLC control system needs to allow for a much faster tuning speed; if necessary this can be accomplished by postponing any analysis or data processing to after the aberration scan is completed and the tuning quadrupoles or sextupoles have been reset to their nominal value.

11.5.8 Components

Main Magnets

The skew correction, diagnostic and geometry sections contain 31 normal and four skew quadrupoles. Parameters are listed in Tables 11-23 and 11-28, respectively, along with power-supply ripple, vibration, and slow drift (stability) tolerances. BMS, CCX, BX, CCY, and FT comprise 40 normal quadrupoles, whose parameters are listed in Table 11-24, and the four quadrupoles of the final doublet, in Table 11-25. Field-stability sensitivities for these magnets are graphically displayed in Figure 11-22, vibration sensitivities in Figures 11-23 and 11-24, stability (or position-drift) sensitivities in Figures 11-25 and 11-26, and, finally, roll sensitivities in Figure 11-27.

Most of the quadrupoles are 0.5-m-long, and their pole-tip field is 3–7 kGauss at 1 TeV, for a typical half-aperture of 10.5 mm. Four quadrupoles in the CCX need a larger half-aperture of 19.5 mm to ensure sufficient beam stay-clear. The largest field gradients are those in the final doublet. Assuming that the half-aperture of the four final-doublet quadrupoles decreases from 11 to 5 mm (Figure 11-14), the required pole-tip fields are about 11 kGauss for the first two, conventional quadrupoles, 34 kGauss for the superconducting magnet QFT1, and 13.5 kGauss for the last quadrupole, QFTA, which is a permanent magnet (see also Section 11.6 and Chapter 12).

At least seven sextupole magnets are foreseen for correcting the horizontal and vertical chromaticity and for increasing the momentum bandwidth of the system. Parameters are given in Table 11-26. The main-sextupole vibration sensitivities are depicted in Figures 11-28 and 11-29. Note that the tolerances for the Brinkmann-sextupoles are extremely loose. The main CCX sextupoles SX1 and SX1b need reversible power supplies, since they will be used to minimize the vertical or horizontal spot size at the pre-image point during initial tune up.

More than half of the final focus is occupied by bending magnets (Table 11-27). Field-stability and roll sensitivities for 7–16-m-long sections of bending magnets are shown in Figures 11-30 and 11-31, respectively. The field-ripple sensitivities appear fairly tight for most of them. However, these ripple sensitivities correspond to jitter tolerances, which have to be met only over a time period of about 1 s, since thereafter orbit-stabilization feedbacks will correct the steering errors.

All sextupoles, quadrupoles, and dipole sections are installed on remotely controlled movers and fed by independent power supplies to facilitate beam-based alignment, orbit feedback and tuning.

Steering and Tuning Elements

All magnets in the final focus are mounted on remotely-controlled movers, which is exploited for beam-based alignment, for orbit stabilization (feedback), and for steering and tuning procedures. Regardless, there is still need for certain special correction elements: one additional skew quadrupole upstream of the final doublet will be used for correcting the residual y–x' coupling at the IP. Two normal and two skew sextupoles in the final transformer allow complete control over those second-order geometric aberrations which affect the vertical spot size. In addition, four small movable tuning sextupoles are foreseen adjacent to the main sextupoles. For this purpose, used FFTB sextupoles could be employed.

A magnetic dither coil is needed for horizontal beam-beam deflection scans. For vertical beam-beam scans and vertical orbit-feedback at the IP, an electrostatic dither is envisioned, since only tiny deflection angles are required. In addition, two horizontal and two vertical steering correctors are installed at each bend function (for instance, between the two main sextupoles of the CCY) to correct for bend-field drifts and bend rolls.

The tuning and correction elements are summarized in Table 11-29.

Diagnostics, Protection and Tune-Up Elements

The region between the entrance of the SCS and the IP comprises 75 main quadrupoles, 15 sextupoles, and five skew quadrupoles. This probably requires the installation of 75–95 conventional beam position monitors with wide dynamic range (*i.e.,* a few mm) for global alignment, initial tuning, and to diagnose "flyers." In addition about 10 rf BPMs [Hartman priv] will be integrated into the structure of some critical magnets to allow stabilization of the average orbit. This is not strictly necessary (see discussion of FT and CCY feedback in this chapter), but the rf BPMs promise very high resolution (*i.e.,* 10 nm) and correspondingly faster feedback response. A particular rf BPM will be installed at the IP pre-image point in the beta-exchanger. This BPM will detect betatron oscillations in the IP phase and, thereby, will allow to correct the beam-beam deflection scans for pulse-to-pulse orbit variations. Performance and reliability of exemplary rf BPMs is being tested at the FFTB.

Beam sizes and emittances are measured in the diagnostic section by means of six laser-wire scanners, each equipped with laser wires at three different angles. For redundancy, an equal number of conventional wire scanners will be installed. The conventional wires can be recycled from the SLC and/or FFTB. They will be useful only during tune up with single bunches and enhanced emittances. A further laser wire is used to infer the beam divergence at the pre-image point of the IP in the BX section (see the discussion in Section 11.4), and a last laser wire, located upstream

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K(10^{-4})$	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QDD	7.00	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD2	4.96	7.	2.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD2	4.96	7.	2.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG1	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG2	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG1	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG2	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00

[a] (nm), with respect to ground.

Table 11-23. *Normal quadrupoles in skew correction, diagnostic, and geometry sections.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-5})	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QB1	2.09	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB2	0.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB3	0.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB4	2.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB5	6.54	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QE2	2.71	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QE2	2.71	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QE1	7.89	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX3	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX1	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX1	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX3	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX1	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX3	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX4	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX5	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX5	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX6	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QY3	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY1	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY1	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY3	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QEI1	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QEI2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI3	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI4	0.76	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QFT6	4.12	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40
QFT5	4.01	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40
QFT4	2.51	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40

[a] (nm), with respect to ground.

Table 11-24. *Normal quadrupoles in the CCX, BX, CCY, and FT.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-5})	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QFT3	11.20	11.	2.00	.42	3.60	.70	1600.	98.00	1.80
QFT2	11.20	11.	2.00	.42	3.60	.70	1600.	98.00	1.80
QFT1	34.70	11.	1.50	.42	3.60	.70	1600.	98.00	1.80
QFTA	13.50	5.	1.50	.42	3.60	.70	1600.	98.00	1.80

[a] (nm), with respect to ground.

Table 11-25. *The quadrupoles of the final doublet. QFTA is a permanent magnet, QFT1 superconducting, and QFT2 and QFT3 are conventional.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-3})	Vibration/drift Δx	Δy
SX1	6.66	20.00	0.40	2.20	50.	50.
SX3	1.80	10.50	0.40	9.00	1000.	1000.
SX1	6.66	20.00	0.40	2.20	50.	50.
SY1	3.71	10.50	0.40	2.20	50.	50.
SY3	1.02	10.50	0.40	9.00	1000.	1000.
SY1	3.71	10.50	0.40	2.20	50.	50.
SI2	1.47	10.50	0.40	9.00	1000.	1000.

[a] (nm), with respect to ground.

Table 11-26. *Normal sextupoles in the CCX, BX, CCY, and FT.*

Name	#	B [G]	L (m)	θ (μrad)	$\Delta\rho/\rho$ (Δx) (10^{-6})	$\Delta\rho/\rho$ ($\sigma_{x,y}$) (10^{-6})	$\theta_{x,y}$ (μrad)
BG	2	538.7	5.00	0.5	2.8	15.	0.5
B1A	5	157.5	7.00	0.5	2.8	15.	0.5
B1B	5	157.5	7.00	0.5	2.8	15.	0.5
B2	28	157.5	7.00	0.5	2.8	15.	0.5
B3A	8	157.5	8.28	0.5	2.8	15.	0.5
B3B	4	157.5	9.19	0.5	2.8	15.	0.5
B4	24	112.5	10.00	0.5	2.8	15.	0.5
B5A	12	65.0	10.00	0.5	2.8	15.	0.5
B5B	8	65.0	11.46	0.5	2.8	15.	0.5
B5C	4	12.0	16.00	0.5	2.8	15.	0.5

Table 11-27. *Bending magnets in the final focus. In reality, each bending magnet listed here represents 3–5 shorter magnets on a common power supply, and possibly common support.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-4})	Vibration/drift Δx (nm)	Δy (nm)
SQ1	4.	6.	0.4	10.	50.	30.
SQ2	4.	6.	0.4	10.	50.	30.
SQ3	4.	6.	0.4	10.	50.	30.
SQ4	4.	6.	0.4	10.	50.	30.

[a] With respect to ground.

Table 11-28. *Skew quadrupoles in skew correction section.*

Type	#	Location
Geom. sextupoles	4	FT
Tuning sextupoles	4	CCX, CCY
Skew quad	1 (4)	FT (SCS)
Magnetic dither	1	at IP
Electrostatic dither	1	at IP
Steering correctors (x or y)	20	at each bend function

Table 11-29. *Steering and tuning elements in the final focus.*

or inside of the final doublet, measures the IP beam divergence. Finally, a laser-interferometer is foreseen for the pre-image point itself. It needs to be determined if a variant of such a monitor could also be installed inside the detector close to the IP.

Two insertible single-bunch stoppers will be convenient during commissioning and for tune up. One is located behind the DS and facilitates coupling correction prior to bringing beam through the final focus. The second insertible stopper, in the final transformer, allows tuning of the final focus, using single bunches with enlarged emittance at 10-Hz repetition rate, before the nominal beam is sent through final doublet and detector. A conventional profile monitor will be placed in front of each beam stopper.

Magnet sensors in each quadrupole and each dipole detect fast field changes and are part of the machine-protection system. Machine protection is also served by two toroids which measure beam losses in the final-focus region. At the high-beta points close to the main sextupoles, four tungsten collimators are installed. These collimators are important for controlling the detector background. There are also about 176 1/4 r.l. Ti spoilers which will protect quadrupoles and bends from missteered beams. Finally, three muon spoilers located at appropriate positions ensure that the muon background remains acceptable [Keller 1993]. Muon spoilers are treated in Chapter 12.

Tables 11-30 and 11-31 summarize the diagnostics and protection elements, respectively, between the SCS and the IP.

11.5.9 Summary

The proposed final-focus design delivers spot sizes not too far from the linear values, in the entire ZDR operating plane, and it has a sufficiently large momentum bandwidth of 1.2–1.3%, both at 500 GeV and at 1 TeV. Motivated by SLC experience, considerable effort was devoted to evaluating and budgeting all possible sources of spot-size dilution. The budgets for tolerances and aberrations are an integral part of the NLC design. Dedicated stabilization systems and tuning procedures throughout the system ensure that most tolerances can easily be met and maintained.

Type	#	Resolution	Comment
Conventional BPMs	84	1 μm	at each magnet
Rf BPMs	10	10–100 nm	at critical locations
Laser-wire scanners	8	10 % of beam size	3 wires at each scanner
Laser-interferometer	1 (or 2)	10 % of beam size	in BX (and close to IP)
Conv. wires	8	20 % of beam size	
Toroids	2	1 % of beam current	in SCS and FT
Profile monitor	2	30 % of beam size	in front of dump

Table 11-30. *Diagnostics elements between the SCS and IP.*

Type	#	Comment
Collimators W	4	at sextupoles
Prot. collimators 1/4 r.l. Ti	176	one per quad and dipole
Insertible single-bunch stopper	1	in FT and after DS
Muon spoilers	3	at strategic locations
Magnet sensors	176	in each quadrupole and dipole

Table 11-31. *Protection and tune-up elements between SCS and IP.*

The requirements on ground motion and on magnet-to-ground vibrations are less severe than what has already been achieved at the FFTB. A possible exception is the final doublet, whose vibration tolerances are comparable to measured quadrupole-vibration amplitudes. Therefore, the final doublet will be further stabilized by means of an optical anchor and a seismometer. In general, the final-focus system is very tolerant with regard to parameter changes, and it lends itself to an uncomplicated upgrade from 350 GeV to 1.5-TeV-c.m. energy.

Outstanding questions that still need some work include a simulation of beam-based alignment, and a description of commissioning and operation. None of these items is expected to be difficult. Further studies may also be devoted to the odd-dispersion design proposed by Oide, which could offer cost savings due to its smaller number of magnets and shorter length without compromising the performance. Tunability, tolerances, and operational flexibility of this scheme, in comparison with the present design, remain to be evaluated.

11.6 The Final Doublet

It is the function of final-focus systems to demagnify the beam. Systems using only quadrupoles are limited by a fundamental, unavoidable chromatic aberration in each plane. In present generation final-focus systems, this aberration is compensated by adding sextupoles at dispersive points in the beam line (K. Brown). The amount of chromatic aberration from quadrupoles, to be compensated by the sextupoles, remains a fundamental parameter determining the characteristics of the system, including length, optical functions, and tolerances. One wishes to start with a quadrupole configuration that has the smallest possible chromaticity.

Given a specification for the demagnification, one can show that the smallest chromaticities arise from placing strong quadrupoles as close as physically possible to the focal point. The strength of the quadrupoles is determined by magnet technology and aperture considerations. The free distance from the first quadrupole to the focal point is denoted by the symbol L^*. Factors determining L^* involve background and detector solid-angle considerations.

The function of the final strong quadrupoles then is to match rays having divergences that, on the one side are determined by the proximity of the focal point, and on the other side, are characterized by the remainder of the final focal system. Since the focal system modules are usually much larger than L^* (hundreds of meters compared to one or two meters), the function of the final quadrupoles is simply to focus to a point those particles whose trajectories are parallel to the beam direction.

The simplest such system consists of two quadrupoles and is referred to as the final doublet. The systems described in this section function as a doublet though, for reasons that will be clarified, it is beneficial to use four separate magnets, which we will call the final quartet. Since from an optical point of view the important parameters will be the horizontal and vertical chromaticities, we will seek final doublet designs that meet other criteria and limitations, keeping the chromaticities as small as possible.

The other criteria to be met include:

- L^* is large enough to meet detector solid-angle and background requirements.
- The magnetic material is suitable for a large solenoidal field environment (3 T).
- The magnetic field strength is able to meet stability requirements.
- The magnets can be constructed, supported, and monitored so as to meet alignment tolerances.
- The system satisfies geometric constraints arising from crossing exit and entrance beam lines.
- Synchrotron radiation is below the Oide limit.
- The system has acceptable energy adjustability.

The purpose of this section will be to show how chromaticity varies with doublet parameter choices, and how one may meet the above criteria with a minimum impact on chromaticity. Then, the important tolerances of the system will be explored, and some operational guidelines presented.

Since the final quartet in the beam line will not exactly focus parallel to point, they will vary somewhat from the designs presented in this chapter. If one wishes to know the parameters exactly, consult the lattice "decks." The purpose of this chapter is to clarify the principles underlying the design of the final doublet (quartet), and to give a sense of the trade-offs that were made in arriving at the present design, and could be made if further changes in quartet parameters were desirable to meet additional or modified criteria.

11.6.1 Doublet Parameters

For the purposes of studying the final doublet (FD) as a system independent of the final-focus system, it is a very good approximation to assume that its function is to focus parallel rays to a point at the IP. This can be confirmed by looking at the plots of $\sqrt{\beta(s)}$ for final-focus systems. The ratio of the slope of this function before and after the FD is less than 0.1, both horizontally and vertically.

To begin we will limit pole-tip fields for final doublet quadrupoles to 1.2 T, (the permanent-magnet material of choice, SmCo with Erbium, for a ratio of inner to outer radius of 1/4, has a pole-tip field of 1.35 T), and assume the pole-tip radius is greater than 3 mm. The latter limit arises from resistive-wall wakefields. This means for the quadrupole $Q1, \kappa_1 = B_{T1}/(a_1 B\rho) \leq 120/E(\text{GeV}) = \{0.48, 0.24, 0.16\}\text{m}^{-2}$ for beam energies of {250, 500, 750} GeV respectively. In general, $\kappa_2 < \kappa_1$, since a somewhat larger Q2 aperture is usually desirable, but more importantly

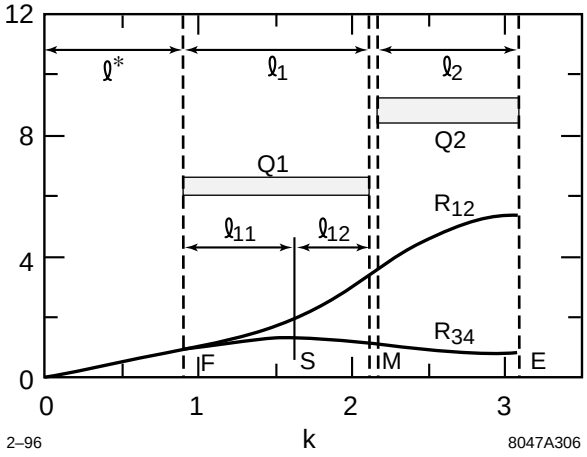


Figure 11-41. *A diagram of a model final doublet.*

the contribution to the Oide effect that arises from horizontal betatron motion through the final doublet can be reduced by taking a smaller κ_2.

Given κ_1, κ_2, and the free length, L^* (from the IP face of Q1 to the IP), the lengths L_1 and L_2 of Q1 and Q2 are determined by the parallel-to-point focusing condition. (There is one other parameter, the free space between Q1 and Q2 which we will denote by d. For purposes of this discussion we take $d = 0$.) Since κ_1 will be chosen as large as possible, and its limit is being fixed by magnet technology and resistive-wall limits, we use $k_1 = \sqrt{\kappa_1}\,\mathrm{m}^{-1}$ to scale all lengths and define $\ell^* = k_1 L^*,\ \ell_1 = k_1 L_1,\ \ell_2 = k_1 L_2$, and $\tilde{k} = k_2/k_1$.

There are now two parameters determining the system: L^* and $\tilde{\kappa}$. A sketch of a model final doublet is shown in Figure 11-41.

The IP end of Q1 is denoted by F (front face), the split between Q1 and Q2 by M (middle), and the entrance to Q2 by E (entrance). Rays have been drawn with unit slope at the IP. This implies that these rays will be the R_{12} and R_{34} functions for horizontal and vertical motion respectively. The magnet Q1 has been split into two parts at S where the R_{34} function is a maximum, defining lengths ℓ_{11} and ℓ_{12}. Note that with $k_1 \approx \{0.7, 0.5, \text{and } 0.4\}$ and $L^* = 2\,\mathrm{m}$, we have $l^* \approx \{1.4, 1.0, \text{and } 0.8\}$ for beam energies $E = \{250, 500, \text{and } 750\}$ GeV, respectively.

11.6.2 An Analytical Model of the Doublet

Element Lengths and Strengths

Letting $x = R_{12}$ and $y = R_{34}$ be the rays originating from the IP, the condition $y/y' = \ell^*$ at F implies $\ell_{11} = \cot^{-1}\ell^*$ and, at S, $y_{\max} = \sqrt{(\ell^{*2}+1)}$. We have

$$\ell_{11} = \cot^{-1}\ell^* \approx \frac{\pi}{4} - \frac{(\ell^*-1)}{2} \quad \text{for} \quad \ell^* \approx 1 \quad . \tag{11.83}$$

At boundary M, we may derive that the ratio $x/x' = (1+m)/(1-m)$ where $m = (\ell^*-1)/(\ell^*+1)\exp(-2\ell_1)$. From the nature of the y function in the ℓ_{12} region, we can make a rough estimate of $\ell_{12} \approx \pi/6$ and adding this to

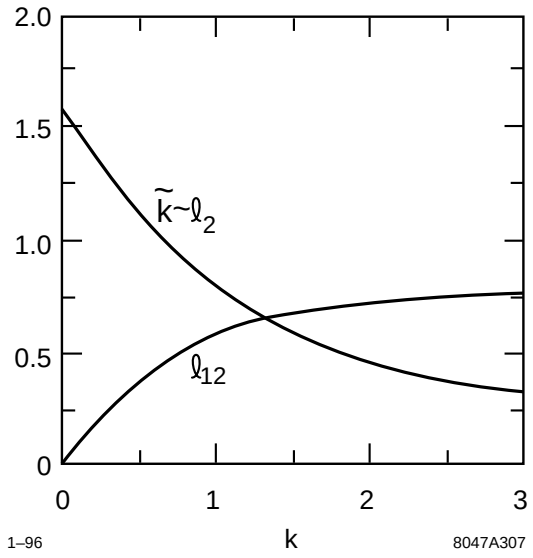


Figure 11-42. *The functions $\ell_{12}(\tilde{k})$ and $\tilde{k}\ell_2(\tilde{k})$.*

$\ell_{11} \approx \pi/4$ from above, find $\ell_1 \approx 1.3$, implying an estimate for m < 0.01 even for $\ell^* = 1.4$. (Even if $\ell_{12} \to 0$ which can occur for very small k, we have $m < 0.05$.) In other words, for all parameters under consideration, x/x' is within a few percent of unity at M. This fact makes it possible to find an analytical model for the doublet by setting $x/x' = 1$ at M.

Since for $\tilde{k} = 1$, and m small, $\ell_2 = \pi/4 = 0.79$, it follows that at M $y/y' = \coth(\pi/4) = 1.52$, hence $\ell_{12} = 0.58$. For general $\tilde{k}$, and m small, $\tilde{k}\ell_2 = \cot^{-1}\tilde{k}$. At M, $y/y' = \coth(\tilde{k}\ell_2)/\tilde{k} = \cot(\ell_{12})$ determines ℓ_{12}.

We will define a function $c(\tilde{k}) = \coth(\tilde{k}\ell_2) = \coth(\cot^{-1}\tilde{k})$. As $\tilde{k} \to \infty$, $c(\tilde{k}) \to \tilde{k}$, and as $\tilde{k} \to 0$, $c(\tilde{k}) \to \coth(\pi/2)$.

$$\ell_{12}(\tilde{k}) = \cot^{-1}\left[\frac{\coth(\tilde{k}\ell_2)}{\tilde{k}}\right] = \cot^{-1}\left[\frac{c(\tilde{k})}{\tilde{k}}\right] \tag{11.84}$$

The functions $\ell_{12}(\tilde{k})$ and $\tilde{k}\ell_2(\tilde{k})$ are shown in Figure 11-42. As $\tilde{k} \to 0$, $\tilde{k}\ell_2 \to \pi/2$ and $\cot \ell_{12} \to \infty$, hence $\ell_{12} \to 0$. As $\tilde{k} \to \infty$, $\tilde{k}\ell_2 \to 1/\tilde{k}$.

The R_{12} and R_{34} Functions

The values of R_{12} and R_{34} at M and E can now be determined.

$$\begin{aligned} R_{34}^M(\tilde{k}) &= \sqrt{\ell^{*2}+1}\,\frac{c(\tilde{k})}{\sqrt{c(\tilde{k})^2+\tilde{k}^2}} & R_{34}^E(\tilde{k}) &= \sqrt{\ell^{*2}+1}\sqrt{\frac{c(\tilde{k})^2-1}{c(\tilde{k})^2+\tilde{k}^2}} \\ R_{12}^M(\tilde{k}) &\approx \tfrac{1}{2}(\ell^*+1)e^{\ell_{11}}e^{\ell_{12}(\tilde{k})} & R_{12}^E(\tilde{k}) &\approx \tfrac{1}{2}(\ell^*+1)e^{\ell_{11}}e^{\ell_{12}(\tilde{k})}\frac{\sqrt{\tilde{k}^2+1}}{\tilde{k}} \end{aligned} \tag{11.85}$$

These functions are shown in Figure 11-43a and 11-43b.

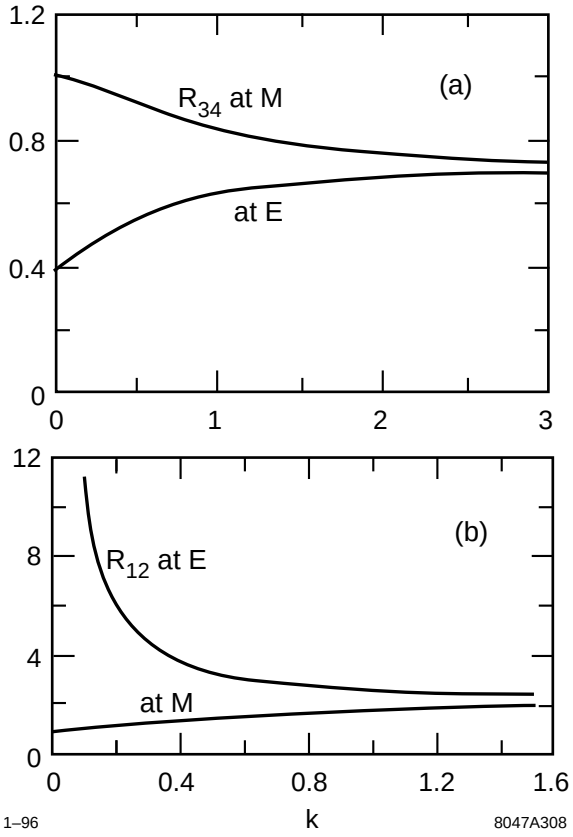


Figure 11-43. *(a) The ray $R_{34}(\tilde{k})$ at M and E. (b) The ray $R_{12}(\tilde{k})$ at M and E.*

Chromaticity

An important parameter for final-focus system design and operational tolerances is the chromaticity, which is given by

$$\int ds\, \kappa_2(s)\beta_z(s) = \frac{1}{\beta_2^*}\int ds\, \kappa_2(s) R_{12}^{z^2}(s) \equiv \frac{L_z^c}{\beta_z^*} \quad . \tag{11.86}$$

The subscript z here can be x or y. $\kappa_z = |\kappa|$ if the quadrupole is focusing in z plane, and $\kappa_z = -|\kappa|$ otherwise.

It is best to consider separately the contributions from the integration of regions specified by ℓ_{11}, ℓ_{12}, and ℓ_2. For the vertical plane, the integral for the regions ℓ_{11}, ℓ_{12} is given by

$$\frac{(\ell^{*2}+1)}{2}\left(\ell + \frac{\cot\ell}{\cot^2\ell + 1}\right) \tag{11.87}$$

where $\ell = \ell_{11}$ or $\ell = \ell_{12}$. The contribution of region ℓ_{11} is independent of $\tilde{k}$ and given by

$$\ell_{11}^{cy} = \frac{\ell^{*2}+1}{2}\cot^{-1}\ell^* + \frac{\ell^*}{2} \quad . \tag{11.88}$$

The contribution from regions ℓ_{12} and ℓ_{l2} is given by

$$\ell_{12}^{cy} + \ell_2^{cy} = \frac{(\ell^{*2}+1)}{2}\left(\ell_{12} + \tilde{k}^2\ell_2\frac{c^2(\tilde{k})-1}{c^2(\tilde{k})+\tilde{k}^2}\right) \quad . \tag{11.89}$$

$\ell_{12}^{cy} + \ell_2^{cy} = 0.13(\ell^{*2}+1)$ at $\tilde{k}=1$ and approaches 0 as $\tilde{k}\to 0$. These two contributions to the vertical chromatic length add up to a function of ℓ^* plus $(\ell^{*2}+1)/2$ times a function of $\tilde{k}$. At $\tilde{k}=1$

$$\ell^{cy} = \frac{\ell^{*2}+1}{2}\left(\cot^{-1}\ell^* + 0.27\right) + \frac{\ell^*}{2} \quad . \tag{11.90}$$

The chromatic length $L^{cy} = \ell^{cy}/\tilde{k}$. Multiplying and dividing by L^* gives

$$L^{cy} = \frac{L^*}{\ell^*}\left[\frac{(\ell^{*2}+1)}{2}\left(\cot^{-1}\ell^* + 0.27\right) + \frac{\ell^*}{2}\right] \tag{11.91}$$

This expression diverges to $+\infty$ for $\ell^*\to 0$ and as $\ell^*\to\infty$ it goes to $L^*(1.0+0.135\ell^*)$. There should be a minimum value somewhere. For $\tilde{k}=1$, the minimum occurs at $\ell^*=1.85$; the chromaticity at $\ell^*=1.0$ is 12% larger and at $\ell^*=0.8$ is 24% larger than this minimum. In Figure 11-44a, we show a plot of L^{cy}/L^* for $\tilde{k}=1$ and $\tilde{k}=0.5$. The large crosses marked 250, 500, and 750 show the parameters for doublets at these three beam energies with $d=0.3$ m (Section 11.6.5). These solutions were done by numerical interaction on the value of x/x' at M, and do not make the assumption that $x/x'=1$ as in the analytic case. Note how well they fall in the analytic approximative curves!

For the horizontal plane, the chromatic contribution from Q1 is given by

$$\ell_1^{cx} = -\frac{1}{8}(L^*+1)^2e^{2\ell_1} - \frac{1}{2}(L^{*2}-1) + \frac{1}{2}\ell^* \tag{11.92}$$

and the contribution from Q2 by

$$\ell_1^{cx} = \frac{1}{8}(L^*+1)^2e^{2\ell_1}\left[\frac{\tilde{k}^2+1}{\tilde{k}}\cot^{-1}\tilde{k} + \tilde{k}\right] \quad . \tag{11.93}$$

The horizontal chromatic length can also be written as the sum of two terms: one which is a function of ℓ^*, and another which is a function of ℓ^* times a function of $\tilde{k}$. We show L^{cx}/L^* for $\tilde{k}=1$ and $\tilde{k}=0.5$ in Figure 11-44b. It also has a minimum (at $\ell^*=6.5$) equal to 5.3. At $\ell^*=2$, it is 1.24 times this minimum, at $\ell^*=1.5$, 1.53 times, and at $\ell^*=1$, a factor of 2.3 times the minimum. Since the β_x^* is a 100 times larger than β_y^*, the horizontal chromaticity is less critical than the vertical.

At this point we can draw the following conclusions:

- The analytic solution based on the R_{12} value-to-slope ratio (x/x') being one between quadrupoles gives the same doublet parameters as the numerical solution (within a few percent).
- There is a minimum horizontal and vertical chromatic length that can be achieved by a final doublet.
- Since the chromaticity is the chromatic length divided by the appropriate β^*, and the desired β_x^*/β_y^* ratio is about 100, the minimum horizontal chromaticity is twenty times smaller than the minimum vertical chromaticity.
- For an L^* of 2 m, field strengths based on a 1.2-T pole-tip field strength, and a beam energy of 250 GeV, the minimum is achieved for the vertical chromatic length and the horizontal chromatic length is about 40% larger than the minimum.

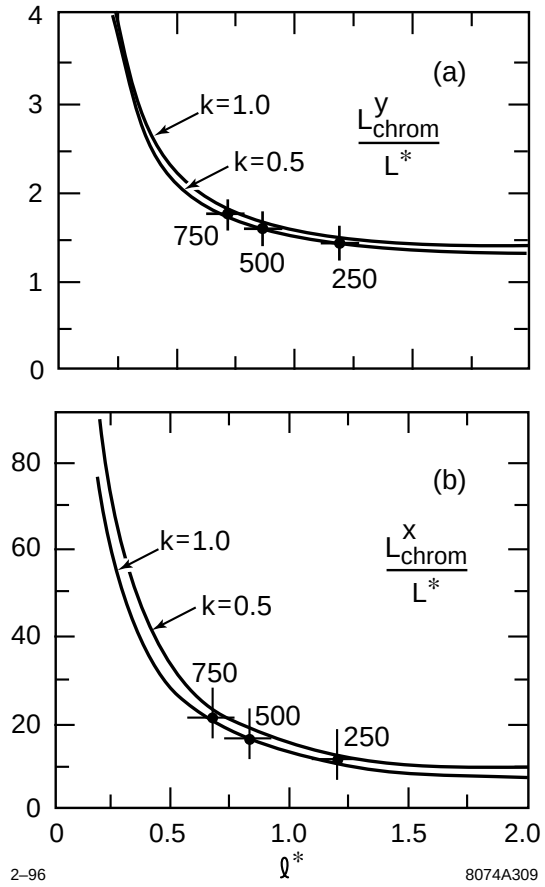


Figure 11-44. *(a) The vertical chromatic length for $\tilde{k} = 1$ and $\tilde{k} = 0.5$ as a function of $k_1 L^* = \ell^*$. (b) The horizontal chromatic length for $\tilde{k} = 1$ and $\tilde{k} = 0.5$ as a function of $k_1 L^* = \ell^*$.*

- The worst situation occurs at 750-GeV beam energy where the vertical chromatic length is 25% larger than the minimum and the horizontal chromatic length is three times the minimum.
- These solutions are surprisingly insensitive to changes in the strength of Q2 (Figure 11-44 shows chromatic lengths for equal field strengths and with the Q2 field strength reduced by a factor of 4).

We now turn our attention to the impact of increasing the separation between the magnetic elements Q1 and Q2. We will show that the chromatic lengths are also surprisingly insensitive to this parameter.

Doublets with Large Q1-Q2 Separations

In Table 11-32, we show the chromatic lengths for three cases with parameters appropriate for the 250-GeV beam energy. The pole-tip field was set at 1.35 T, the aperture of Q1 was chosen to be 5 mm, and L^* remains equal to 2 m. The separations were $d = 0.3$ m, 1.5 m, and 3.0 m.

Figure 11-45 shows the doublet solutions corresponding to the smallest and largest separation, $d = 0.3$ m and $d = 3.0$ m. The vertical chromatic length changes very little: in fact, it decreases slightly with increasing separation. The

Q1-Q2 separation (d)	Vertical chromatic length (L_y^c/L^*)	Horizontal chromatic length (L_x^c/L^*)
0.3 m	1.17	14
1.5 m	1.14	16
3.0 m	1.11	19

Table 11-32. *Chromatic lengths for several Q1-Q2 separations.*

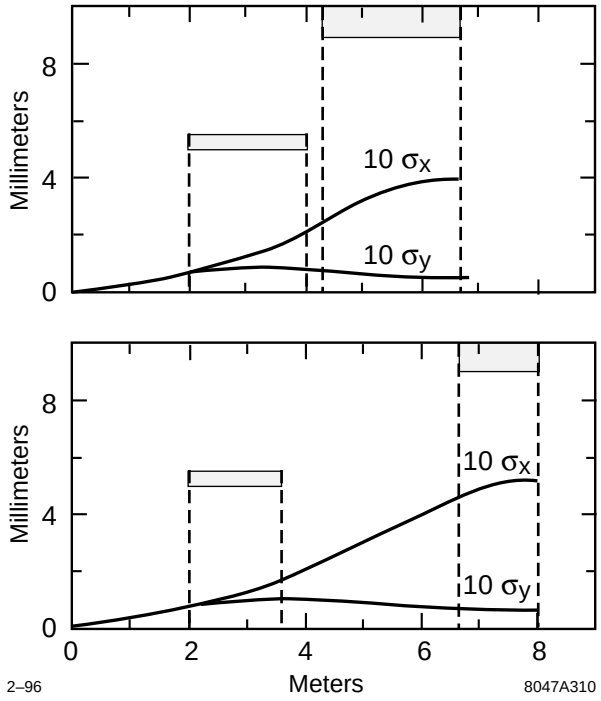


Figure 11-45. *Doublets for separations $d = 0.3\,m$ and $d = 3.0\,m$ with field strengths and L^* held constant.*

major effect is on the horizontal chromaticity which increases by about 33%, and on the R_{12}^E which increases in Q2 by about 25%. It is interesting to note that the length of both Q1 and Q2 decreases as d increases.

A large separation between Q1 and Q2 can be advantageous: i) the region of the detector fringe can be avoided, and ii) space is available for an adjustable magnet to allow for energy variability. For example, a superconducting magnet could be placed in the fringe region, since there the beam separations are sufficient that the entrance beam-line quadrupole not interfere with the exit beam line.

Skew Quadrupole from Tilt

The vertical displacement of an incoming ray due to the tilt of one or both of the quadrupoles is given by

$$\begin{aligned} \Delta y^* &= 2\int R_{34}(s)\,\theta(s)\,\kappa(s)\,x(s)\,ds \\ &= 2x'^*\int R_{34}(s)\,\theta(s)\,\kappa(s)\,R_{12}(s)\,ds \end{aligned} \tag{11.94}$$

where $\theta(s)$ is the quadrupole tilt specified as a function of s. If we suppose that $\theta(s)$ is constant over some range of s, we can substitute $\kappa(s)\,R_{12}(s) = -R_{12}(s)''$, integrate twice by parts and use $\kappa(s)\,R_{34}(s) = R_{34}(s)''$ to get

$$\Delta y^* = \theta x'^*\,[R_{34}(s)\,R'_{12}(s) - R'_{34}(s)\,R_{12}(s)]_{s_1}^{s_2} \quad . \tag{11.95}$$

We note that because of the boundary conditions $R_{12} = R_{34}$ and $R_{12'} = R_{34'}$ at $s = L^*$, and $R_{12'} = R_{34'} = 0$ at the doublet exit, that $\Delta y^* = 0$ if the whole doublet is rotated. If only parts of the doublet are rotated this term is quite large. For example, by requiring $\Delta y^*/y^* < 1/5$ (for 2% luminosity loss), we get the condition

$$\theta < \frac{1}{5}\sqrt{\frac{\varepsilon_y}{\varepsilon_x}}\,\frac{\sqrt{\beta_x^*\beta_y^*}}{[R_{34}(s)\,R'_{12} - R'_{34}(s)\,R_{12}(s)]_{s_1}^{s_2}} \quad . \tag{11.96}$$

Inserting a typical emittance ratio of 1/100 and a $\sqrt{(\beta_x^*\beta_y^*)} \approx 1$ mm, and assuming the bracket can be on the order of 1 m, we find the requirement that $\theta < 20\ \mu$r. Of course, the skew quadrupole aberration can be corrected globally, and this should not be interpreted as an absolute requirement. It does indicate that extreme care should be taken in controlling the tilt of the doublet elements, to the order of 0.1 mr if possible.

11.6.3 Final Doublet Wake Effects

Resistive-wall Wake

We use the standard resistive-wall wake formula (described in Chapter 9, Section 11.8.10).

$$\Delta y'^T_R = 2\,\frac{N r_e}{\gamma\sigma_z}\,\frac{L}{g^3} < f_R > \sqrt{\lambda\sigma_z}\,\Delta y \tag{11.97}$$

where $< f_R >= 0.82$ is the average of the longitudinal shape function of the wake distribution, r_e is the classical electron radius, $\lambda = \rho/(120\pi)$ is called the skin depth (ρ being the resistivity), N is the number of particles per bunch, σ_z is the bunch length, L is the length of the section, g is the beam-pipe radius, and Δy is the offset of the beam from the center. It is the jitter we are primarily concerned with, because a constant steering can be corrected, as long as it is not so large as to effect the beam emittance. In estimating the importance of this term, we use a beam jitter of 1 sigma and integrate the kick along the trajectory of the particle. Since the absolute alignment of the doublet can not be much worse than the maximum of this 1-sigma trajectory (from Oide and second-order dispersion effects), and this term must give a kick that displaces the beam at the IP by less than 0.2 σ_y^*, emittance growth will not be a concern (see Section 9.2.4 for a comparison of emittance growth with centroid kicks).

The contribution of the resistive-wall wake can be compensated quite easily by increasing the aperture of the quadrupole Q1. The $1/g^3$ is a strong fall-off and apertures of 4 mm are adequate for the 1-TeV-c.m. parameter set. In fact we may take an aperture somewhat larger than this to minimize background effects.

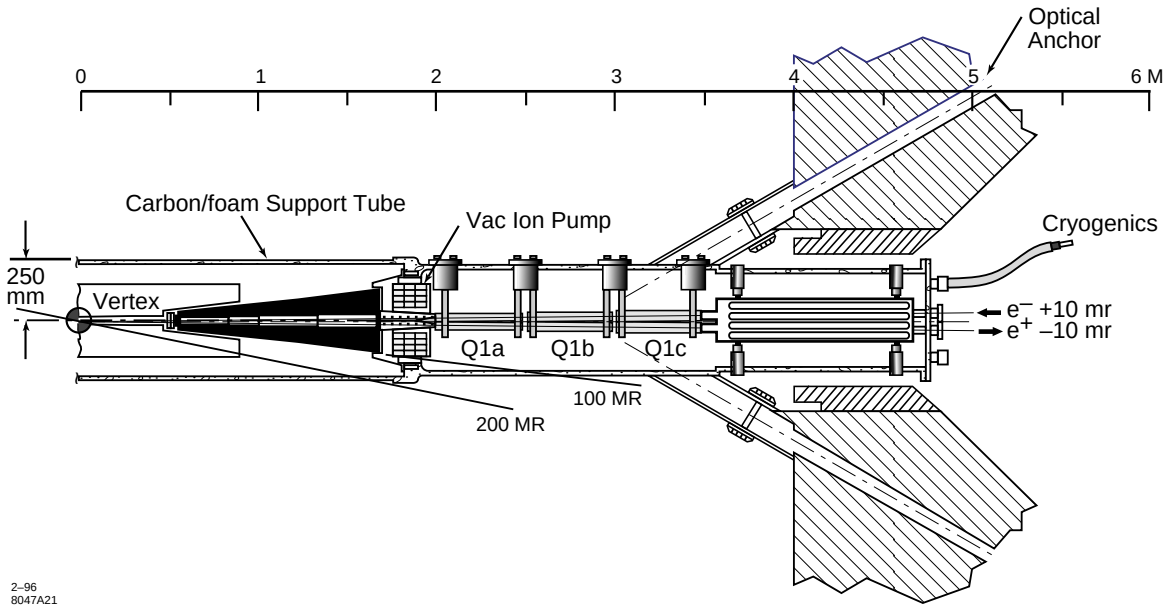


Figure 11-46. *Current final doublet layout with 20-mr crossing angle.*

Geometric Wakes

To minimize geometric wakes we suppose that the walls will be tapered anywhere the beam pipe diameter changes. This could presumably occur at the entrance to the final doublet, at the transition between Q1 and Q2, and at the IP end of Q1. In the collimation section, Section 9.2.4, it was shown that there is a broad optimum taper angle at

$$\Theta_T^{opt} \approx 1.1 \left(\frac{\lambda \sigma_z}{g^2} \right)^{\frac{1}{4}} \tag{11.98}$$

which equals about 10 mr for $g = 5$ mm. Hence the 0.3-m free space between Q1 and Q2 optimally accommodates a 3-mm change of beam-pipe radius. Since radii changes between quadrupoles are less than 4 mm, the total wake from the between-quadrupole region will not be much more than twice the resistive wall part, hence negligible.

At the IP end of Q1, the beam pipe will presumably taper up to the tungsten mask and proceed across the IP along the inner radius of the vertex chamber. This section of beam pipe is marked HOM shield in Figure 11-46.

The surface of the beam pipe at the end of Q1 will be complex since there are two holes here, one from the exit beam line and one from the entrance beam line. Though not explicitly considered in the calculations, we imagine that the beam-pipe surface is defined by two cones, each centered on one of the beam-line quadrupole bores, and that these cones terminate either at their intersection with one another or at the inner surface of the tungsten mask. (We have not considered any multibunch effects which might result from resonances of this IP region cavity.) The geometric wake is given by (see Eq. 9.27)

$$\Delta y'^T_G = 2 \frac{N r_e}{\gamma \sigma_x} \frac{(b-g)^2}{g b L_T} < f_G > \Delta y \quad . \tag{11.99}$$

Using $< f_G >= 0.28$, and $N r_e/(\gamma \sigma_z) \approx 2 \times 10^{-7}$ we have

$$\Delta y^* = L^* \Delta y'^T_G \approx 10^{-7} L^* \frac{b}{L_T} \frac{\Delta y}{g} \quad . \tag{11.100}$$

Δy at the face of the quadrupole (the smallest part of the cone) is given by $L^* y'_{\rm IP}$. Assuming we have about 1-sigma jitter, this becomes

$$\frac{\Delta y^*_{\rm rms}}{y^*_{\rm rms}} \approx 10^{-7} L^* \frac{b}{L_T} \frac{L^* y\prime^*_{\rm rms}}{g y^*_{\rm rms}} = 10^{-7} \frac{b}{L_T} \frac{L^{*2}}{g\beta^*_y} \quad . \tag{11.101}$$

Taking $\beta^*_y = 100\,\mu$m, $L^* = 2$ m and $g = 4$ mm, we obtain $\Delta y^*_{\rm rms}/y^*_{\rm rms} \approx b/L_T$. Parameters foreseen for b (2 cm) and L_T (35 cm) would be adequate.

11.6.4 Synchrotron Radiation Effects

The Oide Effect Including Horizontal Motion

The Oide effect [Oide 1988] begins with a calculation of the change in the IP position due to the change of energy of a particle at point s in the beam line. This is determined by the integral

$$R^{z6}_p(s) = \int_0^s ds' \, \frac{d\theta^z}{ds'}\,(s') R^z_{12}(s') \tag{11.102}$$

where the subscript p indicates the integral is to be taken along the path followed by the particle from the point s to the origin. For a particle following a betatron trajectory in an aligned quadrupole

$$\frac{d\theta^z}{ds} = -\kappa^z(s) z(s) = -\kappa^z(s) R^z_{12}(s) z'_{\rm IP} \tag{11.103}$$

hence

$$R^{z6}_p(s) = -z'_{\rm IP} \int_0^s ds' \kappa^z(s') R^2_{12}(s') = -z'_{\rm IP} L^{cz}(s) \quad . \tag{11.104}$$

If at the position s a photon of energy u is emitted, the change of the IP position for this particle will be

$$\Delta z_{\rm IP}(s) = R^{z6}_p(s) \frac{u}{E} \quad . \tag{11.105}$$

With a correct accounting of statistics [Sands 1985], the spread in the spot size for an ensemble of particles following this path will then be given by

$$(\Delta z_{\rm IP})^2 = \int ds\, \Delta z_{\rm IP}(s)^2 n(u,s) du = \int ds\, R^{z6}_p(s)^2 \int du \left(\frac{u}{E}\right)^2 n(u,s) \tag{11.106}$$

where $n(u,s)$ is the probability of emitting a photon with energy u (per unit length and energy.) The last integral can be performed and equals

$$\int du \left(\frac{u}{E}\right)^2 n(u,s) = c_s r_e \lambda\!\!\!^{-}_e \frac{\gamma^5}{|\rho(s)^3|} \quad . \tag{11.107}$$

Since $\frac{1}{|\rho|} = \frac{|B|}{p/q} = \frac{\sqrt{B_x^2+B_y^2}}{p/q} = \left[\frac{1}{\rho_x^2} + \frac{1}{\rho_y^2}\right]^{\frac{1}{2}}$, this integral can be written

$$(\Delta z_{\rm IP})^2 = c_u r_e \lambda\!\!\!^{-}_e \gamma^5 z'^2_{\rm IP} \int ds L^{cz}(s)^2 |\kappa(s)^3| \left[x'^2_{\rm IP} R_{12}(s)^2 + y'^2_{\rm IP} R_{34}(s)^2\right]^{\frac{3}{2}} \quad . \tag{11.108}$$

The integral in Eq. 11.101 becomes a triple integral where one integrates over a Gaussian distribution in $x'_{\rm IP}$ and $y'_{\rm IP}$. We change variables to $x'_{\rm IP} = t\,\sigma^*_{x'}$ and $y'_{\rm IP} = v\,\sigma^*_{y'}$, then in the $(t\,,v)$ space introduce polar coordinates r and ψ. The radial integral in this space is $\int\limits_0^\infty dr\, r^6\, e^{-\frac{r^2}{2}} = 15\sqrt{\frac{\pi}{2}}$. The angular integral for the vertical direction is

$$\begin{aligned} I(s) &= \frac{1}{2\pi}\int_0^{2\kappa} d\psi \sin^2\psi \left[\cos^2\psi\,\sigma^{*\,2}_{x'}\,R_{12}(s)^2 + \sin^2\psi\,\sigma^{*\,2}_{y'}\,R_{34}(s)^2\right]^{\frac{3}{2}} \\ &= \left[\frac{\sigma^{*\,2}_{x'}\,R_{12}(s)^2 + \sigma^{*\,2}_{y'}\,R_{34}(s)^2}{2}\right]^{\frac{3}{2}} \frac{1}{2\pi}\int_0^{2\kappa} d\psi \sin^2\psi\,[1+\lambda(s)\cos(2\psi)]^{\frac{3}{2}} \\ \text{where}\qquad & \lambda(s) = \frac{\sigma^{*\,2}_{x'}\,R_{12}(s)^2 - \sigma^{*\,2}_{y'}R_{34}(s)^2}{\sigma^{*\,2}_{x'}\,R_{12}(s)^2 + \sigma^{*\,2}_{y'}R_{34}(s)^2} \le 1. \end{aligned} \tag{11.109}$$

It is possible to find a good approximation to the final integral by expanding the bracket in a power series. When the integral is carried out on the series, the terms converge very rapidly. Keeping the leading terms, the resulting expression for $I(s)$ is

$$I(s) \approx \frac{1}{16\sqrt{2}}\left[\sigma^{*\,2}_{x'}\,R^2_{12}(s) + \sigma^{*\,2}_{y'}\,R^2_{34}(s)\right]^{\frac{1}{2}}\left[\sigma^{*\,2}_{x'}\,R^2_{12}(s) + 7\sigma^{*\,2}_{y'}\,R^2_{34}(s)\right] \quad . \tag{11.110}$$

The integral for the horizontal moves the coefficient 7 to the first term in the final bracket. Inserting the result of Eq. 11.110 into Eq. 11.108 we have, for the vertical direction

$$(\Delta y_{\rm IP})^2 \approx \frac{15\sqrt{\pi}\,c_u r_e \lambda_e \gamma^5 \sigma^{*\,2}_{y'}}{32}\int ds\, L^{cz}(s)^2 |\kappa(s)|^3 \left[\sigma^{*\,2}_{x'}\,R^2_{12}(s) + \sigma^{*\,2}_{y'}\,R^2_{34}(s)\right]^{\frac{1}{2}}\left[\sigma^{*\,2}_{x'}\,R^2_{12}(s) + 7\sigma^{*\,2}_{y'}\,R^2_{34}(s)\right] \quad . \tag{11.111}$$

Thus a good indicator for the importance of the Oide effect, taking into account both the horizontal and vertical motion in the final doublet, can be obtained from a single quadrature.

It is appropriate to first look at the situation described by Oide [Oide 1988] when $\sigma^*_{x'} = 0$. Then

$$I(s) = \sigma^{*\,3}_{y'}\,R_{34}(s)^3 \frac{1}{\pi}\int\limits_0^\kappa d\psi \sin^3\psi = \frac{8}{15\pi}\,\sigma^{*\,3}_{y'}\,R^3_{34}(s) \tag{11.112}$$

and

$$(\Delta y_{\rm IP})^2 = 4.2 r_e\,\lambda_e\,\gamma^5 \sigma^{*\,5}_{y'}\int ds\, L^{cy}(s)^2\,|\kappa(s)|^3\,R_{34}(s)^3 \quad . \tag{11.113}$$

For a vertical IP divergence angle of 22 mr and $E = 500$ GeV, the constant in front of the integral in Eq. (11.113) evaluates to $0.02\,\text{nm}^2$. Hence for a 2% contribution to spot size we need the (dimensionless) integral to remain below 8. The integral over Q2 can be made small by choosing a small value of $\tilde{k}$. The value of $\Delta y^2_{\rm IP}/\sigma^2_y$ is given for each of the doublets in the doublet designs of Section 11.6.5.

Oide Effect with Jitter

In the integral over the Gaussian distribution of $y'_{\rm IP}$, it is possible to offset the distribution assuming a nonzero centroid $y'_{0,\,IP}$. With no centroid offset, the average value of $< y'^5_{\rm IP} >= 6.8\sigma'^5_y$. With a centroid offset, the average values are given in Table 11-33.

Centroid offset $y'_{0\ IP}/\sigma_{y'}$	$\left\langle (y'_{\rm IP} + y'_{0\ IP})^5 \right\rangle / \left\langle y'^5_{\rm IP} \right\rangle$
0.00	1.00
0.25	1.16
0.50	1.70
0.75	2.60
1.00	4.00
1.50	10.00
2.00	22.00
3.00	87.00

Table 11-33. *Growth of synchrotron radiation with centroid offsets.*

Photons radiated per electron (in Q1)	σ/σ_0	L_0/L	Luminosity loss was overestimated by:
0.22	1.012	1.006	2.0
0.25	1.036	1.014	2.6
0.29	1.090	1.025	3.6
0.35	1.330	1.050	6.6
0.46	3.170	1.140	15.6
0.71	10.620	1.380	25.3
1.00	37.700	1.860	42.7

Table 11-34. *Beam-shape modifications due to synchrotron radiation in Q1 are very non-Gaussian. Here we compare luminosity loss that would be estimated from rms beam size increase with proper luminosity integrals (based on [Hirata 1989]).*

At an orbit displacement corresponding to 1 sigma, the rms of the vertical displacement due to synchrotron radiation effects in the final doublet are doubled. The conclusions that one would derive from looking at Table 11-33 are considerably tempered when one takes into account the non-Gaussian nature of the bunch distribution. Table 11-33 indicates that if we limit the rms beam size increase as an on-axis beam to 2%, then when missteered by 1 σ, the rms would be 4%. However, according to Table 11-34, an rms increase of 4% in reality implies a 2% luminosity loss. (Table 11-34 is based on information taken from Table 1 in [Hirata 1989]. In that study, β^* was varied, keeping the doublet configuration fixed. This is a bit unrealistic. The appropriate studies must be carried out to correctly evaluate the effects of large synchrotron radiation for our parameter configuration.)

11.6.5 Nominal Final Doublet Designs

A Doublet for 250 GeV

A possible doublet for the 250-GeV beam energy is shown in Figure 11-47. The IP parameters are: $E = 250\,\text{GeV}$, $\varepsilon_x = 10^{-11}$ m-rad, $\varepsilon_y = 10^{-13}$ m-rad, $\beta_x^* = 10^{-2}$ m, $\beta_y^* = 10^{-4}$ m. The pole-tip field of Q1 equaling 1.5 T was calculated for a SmCo magnet with an outer radius of 20 mm. The pole-tip field in Q2 is 1.2 T. The steering from the end of Q2 is given by $R_{12} = 10.6$ m and $R_{34} = 1.5$ m. The stay-clears are 27 σ_x and 60 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute about 0.1% to the vertical beam size. The synchrotron radiation, assuming a final

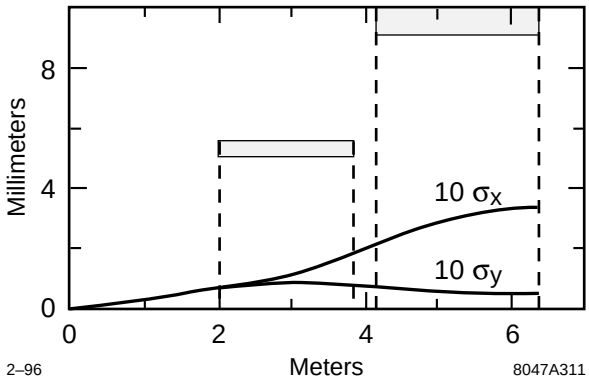


Figure 11-47. *A final doublet diagram for a 250-GeV beam energy showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2$ m, $d = 0.3$ m, $B_{T1} = 1.5$ T, $B_{T2} = 1.2$ T, $a_1 = 5$ mm, $a_2 = 9$ mm. The chromaticities are $\xi_x = 2200$ and $\xi_y = 29,000$.*

Gaussian shape from the photon emission, would contribute 1% to the beam size. In fact, the shape is not Gaussian, and the luminosity degradation will be less than this number for an aligned doublet. Two-thirds of the beam size growth from synchrotron radiation comes from the vertical motion alone. By choosing a weaker Q2, the horizontal motion contribution was held to 1/2 the vertical contribution.

A Doublet for 500 GeV

A final doublet for a 500-GeV beam energy is shown in Figure 11-48. The IP parameters are: $E = 500$ GeV, $\varepsilon_x = 510^{-12}$ m-rad, $\varepsilon_y = 5\,10^{-14}$ mr, $\beta_x^* = 2.5\,10^{-2}$ m, $\beta_y^* = 10^{-4}$ m. Note that the beta function ratio is now enlarged to 240 from 100 for the 250-GeV design. This was to keep the parameter n_γ, the number of beamstrahlung photons per electron, to about 1.0. It is also a help for us in reducing the horizontal chromaticity and the contribution of the horizontal motion to synchrotron radiation.

The steering from the end of Q2 is given by $R_{12} = 15$ m and $R_{34} = 1.8$ m. The stay-clears are 42 σ_x and 73 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute 0.2% to the vertical beam size. The synchrotron radiation, assuming a Gaussian shape for the final beam after photon emission, contributes 2% to the beam size. The horizontal contribution is now less than 1/4 of the vertical.

A Doublet for 750 GeV

Figure 11-49 shows a final doublet for a beam energy of 750 GeV. The assumed IP parameters are: $E = 750$ GeV, $\varepsilon_x = 3.410^{-12}$ m-rad, $\varepsilon_y = 3.4 \times 10^{-14}$ mr, $\beta_x^* = 3.8 \times 10^{-2}$ m, $\beta_y^* = 1.6 \times 10^{-4}$ m. The steering from the end of Q2 is given by $R_{12} = 19$ m and $R_{34} = 2.0$ m. The stay-clears are 51 σ_x and 96 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute 0.1% to the vertical beam size. The synchrotron radiation, assuming a Gaussian-shape beam after photon emission, contributes 2% to the beam size.

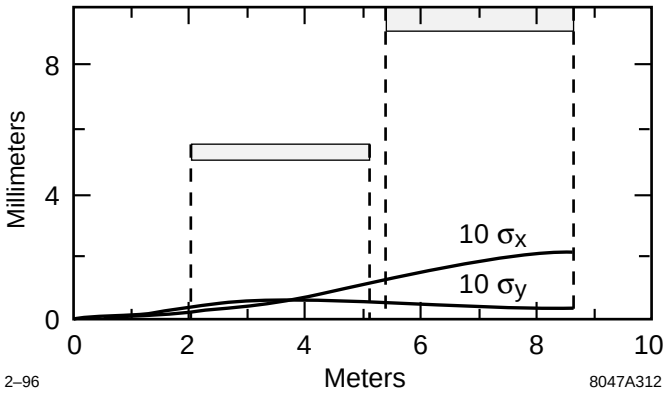


Figure 11-48. *A final doublet diagram for a 1-TeV-c.m. energy showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2$ m, $d = 0.3$ m, $B_{T1} = 1.5$ T, $B_{T2} = 1.2$ T, $a_1 = 5$ mm, $a_2 = 9$ mm. The chromaticities are $\xi_x = 1300$ and $\xi_y = 31,000$.*

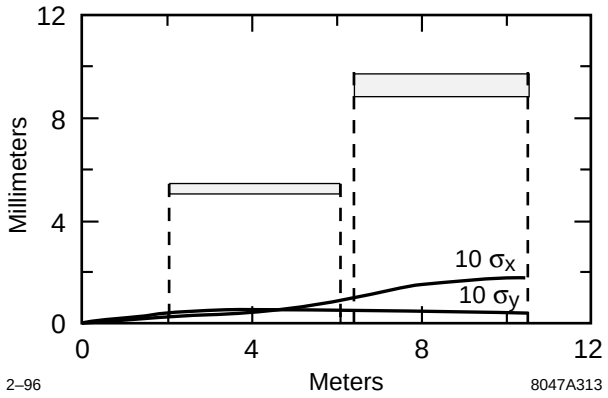


Figure 11-49. *A final doublet diagram for a beam energy of 750 GeV showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2$ m, $d = 0.3$ m, $B_{T1} = 1.5$ T, $B_{T2} = 1.2$ T, $a_1 = 5$ mm, $a_2 = 9$ mm. The chromaticities are $\xi_x = 1100$ and $\xi_y = 23,000$.*

Final Quartets

Since the doublet performance is not a sensitive function of the distance between Q1 and Q2 (in fact separations of a couple of meters appear satisfactory: the vertical chromaticity decreases slightly, while the horizontal chromaticity increases about 20%) we may place a variable element between Q1 and Q2 to obtain energy adjustability.

We propose to place a variable superconducting quadrupole (Q1.5) between Q1 and Q2 in the region of the detector fringe. This quadrupole provides adjustability, and avoids permanent magnet material in the fringe region where the perpendicular components of the solenoidal field become large, and can cause demagnetization.

The quadrupole Q2 is now outside the solenoid and easy to support and adjust. Because of its length, it may be practical to divide it into two pieces. Thus we arrive at the "final quartet."

Figure 11-50 shows two possible final quartets for 250 and 500 GeV beam energy, respectively. Q2 was not divided into two parts so only three elements appear in this figure.

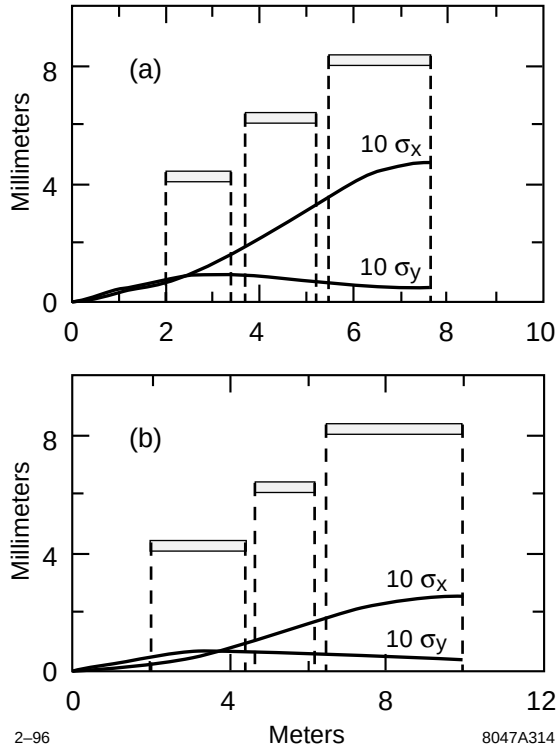


Figure 11-50. *Final quartets at 250- and 500-GeV beam energy.*

11.6.6 Tolerances

Steering Jitter Tolerances

Luminosity Loss

The luminosity for missteered beams, without disruption, goes as

$$L \propto \frac{1}{\sqrt{\sigma_R^2 + \sigma_L^2}} \exp\left(-\frac{\Delta y^2}{2\left(\sigma_R^2 + \sigma_L^2\right)}\right) \tag{11.114}$$

where Δy is the distance between the two beams. The beam separation Δy may arise from random uncorrelated missteering of the two beams, in which case the average displacements satisfy $\Delta y^2 = \Delta y_R^2 + \Delta y_L^2$. Assuming the right and left σ are equal, and the right and left error budget is the same, 2% luminosity loss occurs at rms values $\Delta y_L = \Delta y_R = (1/5)\sigma$.

If the offset is arising from correlated motion of elements to the right and left of the IP, then a 2% loss occurs at an rms $\Delta y = \sqrt{2}/5\sigma$. If the motion is correlated and sinusoidal in nature, then $< \Delta y^2 >= 0.5\ \Delta y_{\max}^2$, hence a 2% loss occurs at $\Delta y_{\max} = (2/5)\sigma$.

There is considerable vertical disruption for all NLC parameters, which means that the separated beams are steered toward one another during the collision. The luminosity formula follows the same relation with σ replaced by 3 σ in

Machine Section	Steering Budget	Accumulated Jitter	Luminosity Loss
To end of linac	$\sqrt{2}/5\sigma$	$\sqrt{2}/5\sigma = 0.28\sigma$	1%
Collimation system:	$1/5\sigma$	$\sqrt{3}/5\sigma = 0.35\sigma$	1.5%
w/ amplification:		$\sqrt{3}/4\sigma = 0.43\sigma$	2.3%
Big bend	$1/10\sigma$	0.44σ	2.4%
Final focus	$1/5\sigma$	0.48σ	2.9%
Final doublet	$1/5\sigma$	0.52σ	3.4%

Table 11-35. *Proposed NLC jitter budget.*

Aberration	Displacement					
	y_{Q1}	y_{Q2}	y_{Q1T}	y_{Q2T}	y_E	$R_{Ey'E}$
$y_{\rm IP}$	1.63	-0.63	0.04	0.04	0	-1.00
$\eta_{\rm IP}$	-1.33	-0.55	0.11	-0.23	1.88	-2.29
$\eta_{2,IP}$	1.35	0.47	-0.09	0.25	-1.07	2.20
$\eta_{{\rm IP}\vert S}$	2.40	-1.99	0.21	-0.15		
$\eta_{2,IP\vert S}$	-2.32	1.85	0.19	0.17		
$\eta_{{\rm IP}\vert SD}$	-0.85	0.72	-0.07	0.08		

Table 11-36. *The strength of aberrations for a variety of doublet displacements and incoming beam errors. This table was compiled for the 250-GeV final doublet of Figure 11-47.*

the exponent of Eq. 11.114 at design intensities. If the intensity is one-half the design, it is appropriate to replace σ by 2 σ in the exponent. To be conservative, we make the latter assumption.

11.6.7 Steering Jitter Budget

So that 2% luminosity losses do not proliferate to an unacceptable sum, we have assigned the budgets indicated in Table 11-35. A budget is also necessary because with a large beam jitter it becomes difficult to carry out beam-based alignment, wakefield effects are enhanced, and synchrotron radiation in the final doublet is enhanced (Section 11.6.4).

Steering Coefficients. With this section, we begin a discussion of aberrations which arise from the displacement of the doublet elements or from an incoming beam that has a nonzero position or slope. The aberration strength can be related to the displacement by a set of dimensionless coefficients, such as are given for the 250-GeV beam energy in Table 11-36. We will discuss the meaning of each of these coefficients, beginning with the steering coefficients.

There are potentially two coefficients relating IP motion to incoming beam centroid parameters. Since the final doublet focuses parallel rays to the IP, one of them is zero. The minus sign in Eq. 11.115 comes from the convention we use that the longitudinal distance is increasing, moving away from the IP.

$$\frac{\partial y_{\rm IP}}{\partial y_E} = 0 \qquad \text{and} \qquad \frac{\partial y_{\rm IP}}{R_{34}^{E\to{\rm IP}}\partial y'_E} = -1 \tag{11.115}$$

We include the R_{34}^E ($1.5\,\text{m} < R_{34}^E < 1.8\,\text{m}$) with the y'_E so that steering and dispersion coefficients are dimensionless.

There are two coefficients relating IP beam-position change to quadrupole displacements,

$$\frac{\partial y_{\rm IP}}{\partial y_{Q1}} \approx 1.6 \qquad \text{and} \qquad \frac{\partial y_{IP}}{\partial y_{Q2}} \approx -0.6 \quad , \tag{11.116}$$

which must add to one because a displacement of both elements shifts the focal point by the same distance. There are two coefficients relating IP motion to quadrupole pitch angle. The subscript T designates "tilt." $y_{Q1T} \equiv 0.5\, L_{Q1}\, y'_{Q1}$ is defined as the displacement of the end of the quadrupole farthest from the IP.

$$\frac{\partial y_{\rm IP}}{0.5 L_1 \partial y'_{Q1}} \approx 0.04 \qquad \text{and} \qquad \frac{\partial y_{\rm IP}}{0.5 L_2 \partial y'_{Q2}} \approx -0.04 \quad . \tag{11.117}$$

These are quite small because the R_{34} function is symmetric in Q1 and rather flat in Q2.

We can draw the following conclusions from these steering coefficients:

- If each quad is supported as a rigid unit, the tilt of this unit is less important by more than an order of magnitude than its position.
- Mounting both quadrupoles in a single long barrel should be avoided. In such a case, for the Π-mode vibration of the barrel, Q1 would be moving upward while Q2 was moving downward, and both would move the IP position in the same direction.
- Mounting the Q1 quadrupoles in a single barrel would be better. In such a case, the support of the barrel should be directly beneath Q1. Then the Π-mode would consist of a rotation of the Q1 quadrupole as a whole, and as noted above, this has a very small effect on the IP position.

Steering Jitter Tolerances. We assume, that at about 5 Hz, the steering feedback can reduce the impact of jitter. At frequencies above 5 Hz, element changes in the IP position could be measured and compensated by steering.

With the Q2 quadrupoles in separate barrels, it is appropriate to consider the four degrees (four is enough since we can neglect tilt) to be specified by the four linear combinations: i) $y_{Q1R} + y_{Q1L}$, ii) $y_{Q1R} - y_{Q1L}$, iii) $y_{Q2R} + y_{Q2L}$, and iv) $y_{Q2R} - y_{Q2L}$. The symmetric combinations move the IP symmetrically, so the beams remain in collision. Only the two antisymmetric motions are significant. The Q2 coordinates are less important than the Q1 by a factor of 2.7, so the jitter tolerance on the system is determined by the combination $y_{Q1R} - y_{Q1L}$. If it is 1 nm, the beams will miss by 1.6 nm. Assuming a doublet jitter budget of 1/2% luminosity loss, and assuming a factor of 2 from disruption (valid at 1/2 the design intensity), the tolerance on $y_{Q1R} - y_{Q1L}$ is $\sigma^*_y/5$.

We note that the motion of Q1R and Q1L should be highly correlated because they are separated by only 6 m. This is a fraction of a wavelength for motions up to 15 Hz. At these frequencies, seismic motion is very small (<0.1 nm) at most locations. Thus it is the cultural noise that must be carefully controlled at these frequencies.

Stability Tolerances from Dispersion

Dispersion Budget. It is assumed that dispersion is being tuned at regular intervals, every hour or less. Since the dispersion aberration is the most sensitive to beam position, it will likely require the most tuning. Ideally, the dispersion tuning would be automated. Dispersion at the IP implies a centroid off-set that is correlated with pulse-to-pulse energy. This correlation could be detected in the steering feedback data, and continuously minimized.

Dispersion at the IP can arise from several sources; the largest source is from the beam position in the final doublet and at the CCY sextupoles. However, this beam position in the final doublet is with respect to the beam position in the

Machine Section	Dispersion Budget	Luminosity Loss
To end of linac	1/5 σ	3% tent.
Collimation system	1/10 σ	0.5%
IP switch/big bend	1/10 σ	0.5%
Final focus/final doublet	1/5 σ	2%

Table 11-37. *Proposed NLC dispersion budget.*

final-focus system, because incoming jitter will not give rise to dispersion. It is the final-focus system, especially the CCY and the final telescope that must be stabilized. A proposed dispersion budget is shown in Table 11-37.

Allowing a 2% loss in luminosity from residual dispersion, the rms value of the dispersion aberration at the IP must be less than 1/5 the beam size, $< \eta_1\delta >_{\rm rms} \leq \sigma y^*/5$. Using a square distribution of width $\pm\Delta$, the rms of $\eta_1\delta$ is $< \eta_1\delta >_{\rm rms} = \eta_1\Delta/\sqrt{3}$. Thus, for $\Delta = 510^{-3}$ ($\delta_{\rm rms} \approx 3\times10^{-3}$) the first-order dispersion must satisfy $\eta_1 < 67\sigma_y^*$. At 1 TeV and $\sigma_y^* = 3$ nm, we must have $\eta_1 \leq 200$ nm. This sets the first-order dispersion scale.

11.6.8 Sources of First-order Dispersion in the Doublet

Dispersion arises from a change of beam position at the doublet entrance or a change of beam position at the IP. These are given in Table 11-36 as

$$\frac{\partial\eta_1}{\partial y_E} \approx 1.9 \qquad \text{and} \qquad \frac{\partial\eta_1}{R_{34}^{E\to IP}\partial y'_E} \approx -2.3 \quad . \tag{11.118}$$

The first term, corresponding to a displacement of the incoming beam or a vertical motion of the doublet system as a whole, indicates a tolerance of 200/1.9 = 105 nm for the 1-TeV machine. The vertical beam size at this point is about 33 μm, so this tolerance is about 1/300 of the beam size. This deserves some explanation.

The tolerance arises from the fact that passing off-axis through large chromaticity gives rise to dispersion. Since the chromaticity is compensated by the sextupoles in the CCY section of the final-focus system, an incoming jitter of the beam will pass off-axis through three large chromaticity points (the two sextupoles and the final doublet). Since the chromaticity has been adjusted to cancel, the dispersion will cancel as well. However, if the final doublet moves with respect to the final-focus system, or a quadrupole in the final telescope moves, or if the strength of one of the dipoles in the final telescope changes, then the dispersion will not cancel. So this problem can be approached in three ways:

- Control all systems to meet the stability requirements for time intervals equal to the time between dispersion tunings (this amounts to 20 nm for a few quadrupoles in the final telescope). For the FFTB, the observed motion (measured relative to a wire alignment system) is less than a micron per hour. This is a ground surface installation, and one might expect to do better underground.
- Have very sensitive BPMs ("rf BPMs") located at crucial positions, such as at the main CCY sextupoles and final doublet quadrupoles, and use all BPMs in the final-focus system to determine the incoming jitter of any pulse. Then with a minute's worth of pulses (10,800), select 1% (108 pulses) with the smallest jitter amplitude for further analysis. The amplitudes of these pulses are small enough so that the higher sensitivity rf BPMs can be used, since for these selected pulses the distribution width at the final doublet entrance will be 330 nm. Now, if the rf BPM had the required sensitivity (<100 nm), any one of these pulses would determine if something in

the beam line had moved, because it is a change in relative readings in the system that is of consequence. 100 pulses is an abundance.

- Use the fact that the presence of dispersion at the IP implies energy dependent steering. A correlation of IP position (as determined by beam-beam pulse-to-pulse deflections) with pulse-to-pulse energy from both right and left beams can be used to continuously correct the dispersion.

The second equation in Eq. 11.118 is of importance if the IP collision point wanders. A proper analysis requires consideration of the two doublets and the optimal method for their operation as a pair. In addition to holding the entry position (or the position in Q2) to each doublet constant, some other measurement must be used. One might imagine a good choice would be to minimize the change in the absolute sums of BPM readings in the two exit quadrupoles. This is not a good choice, because the ray that begins at the corrector in front of Q2 has a small value at the exit quadrupole in comparison to the size of this ray in Q1, where the dispersion is created. A better choice is to minimize the sum of the absolute changes of BPM readings in the two Q1s.

So let us imagine that this absolute sum is minimized and the beams are held in collision. To begin with, we could assume that the IP position does not change. The dispersion which arises from a small change in Q1, subject to this constraint, is:

$$\left.\frac{\partial \eta_1}{\partial y_{Q1}}\right|_s = \frac{\partial \eta_1}{\partial y_{Q1}} - \frac{\partial \eta_1}{R^{E\to IP}\partial y'_E}\,\frac{\partial y_{\rm IP}}{\partial y_{Q1}} \quad . \tag{11.119}$$

The values for these quantities are also shown in Table 11-36

$$\left.\frac{\partial \eta_1}{\partial y_{Q1}}\right|_s = 2.4 \qquad \text{and} \qquad \left.\frac{\partial \eta_1}{\partial y_{Q2}}\right|_s = -2.0 \tag{11.120}$$

The size of these coefficients emphasizes the importance of keeping the beam passing through the centers of these quadrupoles. In other words, misalignments of 200/2.4 = 83 nm is the maximum displacement of Q1 before the dispersion becomes unacceptable. As mentioned, this must be diagnosed and continually tuned or compensated. The beams must be held in collision, and the dispersion of each beam must remain unchanged. The latter can be arranged by changing the position at E (the IP position does not change, but the dispersion does.)

Stability Tolerances from Waist Motion

The strength aberration is also given by the chromaticity function. If the strength of both magnets in the doublet changed together, the tolerance on the strength would be:

$$\frac{\Delta k}{k} \le \frac{1}{10\xi} \approx 3.3 \times 10^{-6} \tag{11.121}$$

where we have budgeted a 1/2% luminosity loss for this aberration to the final doublet, and used 30,000 for the vertical chromaticity. If only the strength of Q1 changed, this tolerance is about 20% smaller.

Thermal Stability for Permanent Magnets. The material SmCo has a strength coefficient of $-4 \times 10^{-4}/^\circ$C, which would lead to the conclusion that the magnets should be stable to $10^{-2}\,^\circ$C in times that are on the order of waist tuning times (which we would like to have at one-hour intervals). This could conceivably be accomplished by providing a controlled thermal environment for the magnets. Temperature deviations of this order could be measured and small waist corrections introduced to compensate thermal drift.

Another possibility is to use a permanent magnet material that is more stable. There exists permanent magnet materials, $Sm_xEr_{l-x}Co$, containing erbium, that have very flat temperature coefficients, down to 10^{-6}/°C. In fact there are even materials which increase in strength with increasing temperature, at 10^{-5}/°C which could be chosen so as to compensate for the thermal expansion of the doublet. The remnant field of these materials is about 0.9 T, whereas the straight $SmCo$ remnant field can be as high as 1.06 T. However, 0.9 T seems adequate, and can give a pole-tip field of 1.35 T.

Temporal Stability of Permanent Magnets. The temporal stability of permanent magnet materials has been measured to be about $10^{-6}/hr$. Our experience with permanent magnets in accelerator environments also confirms this number [Spencer 1995]. Thus the temporal stability of permanent magnets is acceptable. The temporal stability of the $Sm_xEr_{l-x}Co$ needs to be checked.

Static Alignment Tolerances

Static Tolerances from Second-Order Dispersion. If we suppose that there is no knob in the final-focus system to adjust second-order dispersion, then this aberration will establish limits on static alignment tolerances. Synchrotron radiation also sets a limit on static alignment. The lowest of these two limits must be taken as the static alignment specification.

In calculating this tolerance, we assume that steering errors arising from static misalignments are being corrected by steering elements at E (the entrance end of Q2), and the dispersion is being corrected by adjusting the beam position at E. The appropriate coefficients are of the form:

$$\frac{\partial \eta_2}{\partial y_{Q1}}\Big|_{sd} = \frac{\partial \eta_2}{\partial y_{Q1}}\Big|_{s} - \frac{\partial \eta_2}{\partial y_E}\,\frac{\frac{\partial \eta_1}{\partial y_{Q1}}\Big|_{s}}{\frac{\partial \eta_1}{\partial y_E}\Big|_{s}} \quad . \tag{11.122}$$

The aberration term $\eta_2\delta^2$ contains a steering aberration, $\eta_2 < \delta^2 >$. After subtracting this steering, the rms spot size growth is $\eta_2\sqrt{\left(\left\langle\delta^4\right\rangle - \left\langle\delta^2\right\rangle^2\right)}$. For a square distribution of width $\pm\Delta$, this quantity equals $0.3\eta_2\Delta^2$. If the limit on luminosity loss from the second-order dispersion is taken to be 2% and $\Delta = 5\times10^{-3}$, then the second-order dispersion must satisfy $\eta_2 \le 3\times10^4\sigma_y^*$. At 1 TeV for $\sigma_y^* = 3$ nm, $\eta_2 \le 90\,\mu$m. This distance sets the scale of the static alignment tolerances arising from second-order dispersion effects.

Referring to Table 11-36, we see

$$\left.\frac{\partial \eta_2}{\partial y_{Q1}}\right|_{SD} = -0.85 \qquad \left.\frac{\partial \eta_2}{\partial y_{Q1}}\right|_{SD} = -0.72 \qquad \left.\frac{\partial \eta_2}{\partial y_E}\right|_{SD} = -1.0p7 \quad . \tag{11.123}$$

The fact that all of these coefficients are close to unity, indicates the absolute alignment tolerances arising from second-order dispersion are on the order of 90 μm. A smaller absolute alignment tolerance comes from synchrotron radiation considerations.

Static Tolerances from Synchrotron Radiation. Table 11-33 indicates that for a 0.5-sigma vertically-missteered beam, the synchrotron radiation integral is larger by a factor of 1.7. That means that a displacement of the incoming beam by about 16 μm results in an additional 1.4% spot size increase. (This estimate may be a factor of two too high. See Table 11-34). Though we have not computed the synchrotron radiation increase from element misalignments, this tolerance arises from the integral in Q1, hence it is the orbit in Q1 which sets this number. Since the missteered beam

has an orbit maximum which is somewhat larger in Q1, the misalignment tolerance will be somewhat larger than the 16-μm number deduced above. This number is a factor of five smaller than the static tolerance limits arising from second-order dispersion.

Meeting Static Alignment Tolerances. To meet absolute (static) alignment tolerances, it will be necessary to locate the beam with respect to the magnetic centers in the final doublet system. One must either provide for beam alignment based on varying quadrupole strengths, or in the case of permanent magnets, provide BPMs with appropriate absolute alignment accuracy. For permanent magnet systems, we will assume that BPMs are located at all element centers and determine the position of the magnetic centers with an absolute accuracy denoted by $\Delta y_{\rm BPM}$. For example, rf BPMs may be integrated into the permanent magnet quadrupole structure.

We begin by assuming the BPM centers are coincident with the magnetic centers. Thus they establish the two doublet center lines (say right and left) to which the incoming beam may be steered. Let us begin by neglecting the steering that may be required to arrange that each incident beam is traveling along this center line. The IP separation of these beams will be given by

$$\Delta y = \frac{s_2}{s_2 - s_1}\left(y_{1R} - y_{1L}\right) - \frac{s_1}{s_2 - s_1}\left(y_{2R} - y_{2L}\right) \quad . \tag{11.124}$$

As steering is introduced to collide the beams, it should be done with equal corrector strength on each side. Then magnets Q1 or dipole coils around Q1 should be equally adjusted so that the required steering from correctors is zero on each side. If the BPM centers were equal to the magnetic centers, the doublet would now be perfectly aligned. However, there will be BPM error. Synchrotron radiation, as noted in the previous paragraph, sets the limit on the absolute accuracy of these BPMs to be $\Delta y_{\rm BPM} \leq 16\ \mu$m.

11.6.9 Superconducting Quadrupole Q1.5

We propose to place a superconducting quadrupole, Q1.5, upstream of the permanent magnet Q1. This quadrupole serves three purposes: It, first, can provide a high gradient up to about 300 Tm^{-1}; it, second, allows operating the same doublet over a wide range of beam energies; and, third, it avoids permanent-magnet material in the solenoid-fringe region. The exit face of Q1.5 is located 3.85 m from the IP, where the incoming and outgoing beam lines are separated by about 7.7 cm. The inner bore is 11–14. mm, and the pole-tip field at 1-TeV-c.m. energy is 3–4 T.

Using a regular superconducting quadrupole would lead to a considerable field in the region where the second beam will pass, which is not desirable. We, therefore, intend to use a special magnet configuration: the current distribution will be arranged in an inner and outer shell so as to approximate two concentric $\cos 2\theta$ distributions with opposite polarity. While such a magnet can generate about the same field gradient as a conventional superconducting magnet, it has the additional advantage that the stray field experienced by the other beam is essentially zero (it is about 2–3 G).

Figure 11-51 illustrates the proposed current distribution and magnetic field pattern at the front face of the superconducting quadrupole pair viewed from the IP. Figure 11-52 shows the field-strength variation along the centerline of the two magnets on a linear scale, assuming a pole-tip field of 2 T and an inner-bore radius of 14 mm. The field falls off rapidly outside of and between the two quadrupoles, and it is very linear inside of each magnet. Finally, in Figure 11-53 the absolute value of the field is depicted on a logarithmic scale. Note that the field should go through zero at the center of the magnet, and that the finite value shown is an artifact of the drawing.

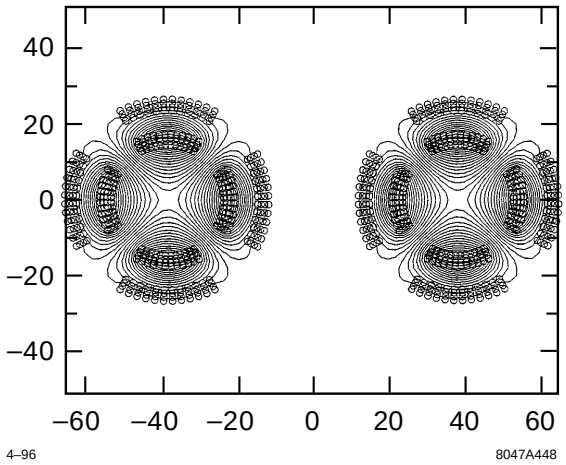


Figure 11-51. *Current and field distributions for the superconducting quadrupole pair Q1.5.*

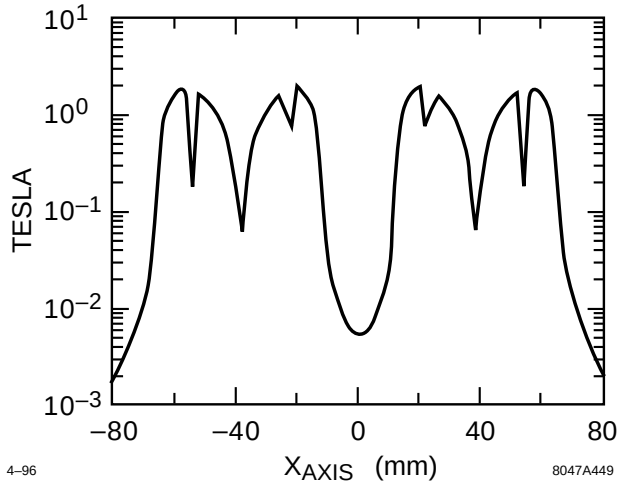


Figure 11-52. *Field strength along the centerline of the superconducting quadrupole pair Q1.5 on a linear scale.*

11.6.10 Summary

We have presented doublet designs for beam energies of 250, 500, and 750 GeV, whose chromaticities are close to the theoretical minimum and additionally

- have acceptable Oide effect synchrotron radiation,
- have acceptable resistive-wall and geometric wakes,
- have sufficient beam stay-clear,
- provide for a large range of adjustment to changing beam energy (about a factor of 2),
- are sufficiently stable thermally and temporally,

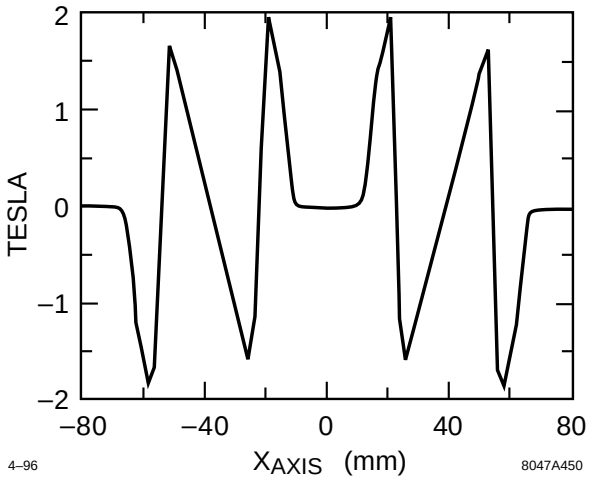


Figure 11-53. *Absolute field value along the centerline of the superconducting quadrupole pair Q1.5 on a logarithmic scale.*

- have no permanent-magnet material in solenoid fringe region,
- provide sufficient space for an exit beam path, exit beam-line elements and instrumentation at the 20mr crossing angle, and
- provide a mechanism (dipole coils) to achieve the required absolute alignment and to compensate solenoid field changes.

Final-doublet support is discussed in Section 12.4.2.

11.7 Crossing Angle, Crab Cavity, and Solenoid

Section 11.7 begins with a discussion of the factors involved in the choice of 20 mr as the IP crossing angle. Such a relatively large crossing angle implies the need for a crab cavity. The crab-cavity parameters and tolerances are discussed in Section 11.7.2. The crab cavity is placed either at the entrance of the final doublet, or if a quartet is used to obtain energy variability, then the crab cavity will be placed between the second and third element of the quartet so that the aperture of the cavity better conforms to the apertures of the quadrupoles. The crab cavity can be six $\lambda/3$ X-band cells with a total length of 5.2 cm.

This section concludes with a discussion of the effects of the detector solenoidal field including the skew aberrations, steering and dispersion effects, and synchrotron radiation.

11.7.1 Determination of Crossing Angle

The crossing angle must be large enough so that:

- There is no multibunch instability from parasitic beam-beam interactions.
- There is physical room for the final doublet and the exit beam port.
- The big bend angle provides sufficient muon protection for the detectors.
- The IPs are sufficiently separated to provide radiation and ground noise isolation.

The last two items are necessary only for beam-line geometries that interrelate the big bend angle with the crossing angle, as in the geometry we have chosen.

The crossing angle should be chosen as small as possible to minimize:

- The impact of the solenoidal field on steering, dispersion, and synchrotron radiation.
- The power requirements of the crab cavity.
- The voltage and phase stability tolerances of the crab cavity.
- The length of the big bend (especially at 750-GeV beam energy).

The NLC crossing angle parameter has been chosen to be $\theta_c = 20$ mr and the corresponding big bend angle is 10 mr (Figure 11-46).

11.7.2 Crab Cavity

The purpose of the crab cavity is to provide a transverse displacement within the bunch that varies with longitudinal distance according to

$$\frac{\partial x^*}{\partial z} = \frac{\theta_c}{2} \quad . \tag{11.125}$$

Using $\Delta x^* = R_{12}^E \Delta x'_E$ and $\Delta x'_E = \text{ eV}/E_0$, we find

$$\frac{\theta_c}{2} = \frac{\partial x^*}{\partial z} = R_{12}^E \frac{e\,\partial V(z)}{E\;\partial z} = R_{12}^E \frac{\text{eV}_{\max}}{E} k\cos(kz) \approx R_{12}^E \frac{\text{eV}_{\max}}{E}\frac{2\pi}{\lambda} \tag{11.126}$$

where in the last step we assumed the bunch was traveling through the cavity at the zero of the rf wave. R_{12}^E, which is the R_{12} matrix element from the entrance of the doublet to the IP, can vary somewhat with doublet design and machine energy, but it is in the range of 10 m to 20 m.

Crab Cavity Parameters

We have chosen the rf wavelength at X-band to keep the power requirements low and the structure length short. Table 11-38 compares X-band with S-band [Wilson 1995]. The parameters of the S-band cavity are "LOLA-III" taken from SLAC Report 17 (1963) and SLAC-PUB-135 (1965). The X-band parameters we will refer to as LOLA-X. Lolita-X is six cells of LOLA- X.

The required values for $V_{\max}$ and corresponding power are given in Table 11-39. We see that for the LOLA-X structure, the required input power is close to 0.25 MW for all energies. Since this is a very small power, the structure

Parameter	Symbol	Unit	Scaling	S-Band LOLA-III	X-band LOLA-X	X-band Lolita-X
Transverse shunt impedance	$r_\perp$	MΩ/m	$\omega^{1/2}$	11.7	23.4	23.4
Structure length	L	m	$\omega^{-3/2}$	3	0.375	0.0523
Attenuation	τ	none	1	1.04	1.04	0.15
Iris diameter	$2\,a$	mm	ω^{-1}	47	11.8	11.8
Q value	Q	1	$\omega^{-1/2}$	11,000	5,500	5,500
Group velocity	v_g/c	1	1	-0.0078	-0.0078	-0.0078
Fill time	TF	ns	$\omega^{-3/2}$	1,280	160	22
Number of $\lambda/3$ cells	N_c	1	$\omega^{-1/2}$	86	43	6
Deflection volts/$\sqrt{\text{MW}}$	$p_\perp c/\sqrt{P_0}$	MeV/$\sqrt{\text{MW}}$	$\omega^{-1/2}$	5.3	2.65	0.55

Table 11-38. *Scaling of crab-cavity parameters with wavelength.*

Beam Energy (GeV)	R_{12}^E (m)	θ_c (mr)	V (MV) S-band	V (MV) X-band	Power (MW) LOLA-III	Power (MW) LOLA-X	Power (MW) Lolita-X
250	8.9	20	4.8	1.2	0.8	0.20	4.8
500	17	20	5.0	1.3	0.9	0.24	5.6
750	21	20	6.0	1.5	1.3	0.32	7.4

Table 11-39. *Required values of maximum crab-cavity voltage for three energies.*

can be substantially decreased in length. The X-band structure with six cells (Lolita-X, 5.23-cm long) would suffice if driven with 7.4 MW of power. Shortening the structure also greatly reduces wakefield effects (almost as the square of the length).

A potential problem with the X-band structures is the iris aperture of radius 5.9 mm. We have chosen a somewhat larger Q2 aperture, to minimize resistive-wall wakes as well as provide a large beam stay-clear. At 250-GeV beam energy, the 23-σ_x aperture at the end of Q2 (see Figure 11-14) is about 11 mm for the parameter set with an 8-mm β_x^*. However, it is also possible to place the crab cavity between Q1 and Q2. For this parameter set R_{12} is still about 8 m, due to the space occupied by the superconducting Q1.

Crab Cavity Tolerances

We discuss five tolerance issues:

- Crab cavity voltage stability.
- Crab cavity pulse-to-pulse phase difference jitter.
- Crab cavity phase difference short-term stability (1 hour).

- Beam phase jitter.
- Crab cavity alignment.

Crab Cavity Voltage Stability. If the crab-cavity voltage changes, the "crabbed" angle changes. A 2% luminosity loss would occur for a crab angle error of $\theta_d/5 \approx 0.2\sigma_x/\sigma_z \approx 0.6\,\text{mr}$. For $\theta_c/2 = 10\,\text{mr}$, the permissible voltage error would be $\Delta\,\text{V}/\,\text{V} = (\theta_d/5)/(\theta_c/2) \approx 0.06$. This estimate yields a permissible 6% voltage error.

Crab Cavity Phase Difference Jitter. If the phase of the crab cavity on the left jitters with respect to the crab cavity on the right, the beams will not collide head-on. A 2% luminosity loss occurs at $\Delta x = (2/5)\,\sigma_x$ (right beam minus left beam). Since $\Delta x = (\theta_c/2)\Delta z$, the allowed $\Delta z = (4/5)\,(\sigma_x/\theta_c)$, which for $\theta_c = 20\,\text{mr}$ has the value $\Delta z = 12.8\,\mu$m, corresponding at X-band to about 0.2°. At S-band, the number of degrees would be an intimidating factor of four smaller.

It should be noted that this tolerance is a phase difference tolerance between the two crab cavities. It is a jitter tolerance, since steering correctors would compensate for any error in crab cavity phase difference which occurs for times greater than 1/5 s. To minimize phase difference jitter, both cavities should be driven by the same klystron. The phase jitter for an X-band system in such a circumstance has yet to be experimentally determined, but a simple experiment which uses a phase mixer to measure the relative phase in two rf transport lines coming from a splitter, driven by one klystron, can provide the required information. (Note: It is routine in the SLC klystron gallery to compare the S-band phase of the output from a SLED-I cavity with the rf reference phase driving that klystron. There the phase jitter is known to be less than or the order of 0.1°.)

It is interesting to estimate the thermal change on the transmission line on one side that would be result in a 0.2° phase change. The phase change is given by $\Delta\phi = \Delta(k_g L) = k_g L(\Delta k_g/k_g + \Delta L/L)$. Using $k_g = 2\pi/\lambda_g = 2\pi/\lambda(1-y^2)^{1/2}$ where $y = 0.61\lambda/a$ we have $\Delta k_g/k_g = (y^2/(1-y^2))\Delta a/a$. And $\Delta a/a = \Delta L/L = \alpha\Delta T$ with $\alpha = 1.8\,10^{-5}$, leads to $\Delta\phi = 2\pi L/\lambda\alpha\Delta T(1-y^2)^{-1/2} = 2.7°$ per °C (taking $L = 10\,\text{m}$, $\lambda = 0.026$). In other words, the temperature must be stable to 1/15°C in times of 1/5 s, or drifts should be smaller than 20°C per minute. It is expected that the temperature will be more stable than this.

A concern is that the phase difference may creep, so that the steering corrector kicks become large. Phases should be monitored for this, and a system installed to maintain the phase of the rf reaching each cavity to be the same, and at a null at the beam phase. 0.5 degrees corresponds to a horizontal steering at the IP of 1 σ_x. Probably 40-σ_x steering, corresponding to a crab-cavity-to-beam phase difference of 20°, is possible before unacceptable dispersion is created. The phase of the cavity with respect to the beam can be monitored by observing steering of the beam at a post-IP BPM as the power of the cavity is ramped. There will be no steering if the beam-pulse arrives at the rf null.

Incoming Beam Phase Jitter. Although the crab-cavity jitter tolerance is only 0.2°, this does not imply a need for a corresponding phase stability of bunches in the incoming beam. The function of the crab cavity is to align the bunch for a complete collision no matter how long. Hence the limits on beam phase stability arise from considerations of luminosity loss due to hourglass effects at the IP. If we can have $\Delta z_{\text{collision}} = \frac{1}{5}\beta_y^*$, and if the arrival of each beam is taken to be statistically unrelated, then we can have $\Delta z_{\text{beam}} = \sqrt{2}(100)\,\mu$m, which is about 1.9 degrees of X-band.

Crab Cavity Alignment for Multi-bunch Effects. If the beam passes through the cavity on-axis, no dipole mode is induced, even if the bunches are not at the zero phase. This is because the cavity does not put energy into the beam when steering it. So beam loading is entirely associated with misaligned cavities. We will assume a misalignment of 100 μm for numerical estimates.

Even if the cavity is misaligned, the induced fundamental field has a null at the longitudinal phase where the charge passes. As a result, if all particles are passing through the cavity at the same phase (modulo 2π), the induced field will change the slope of the field, but not its magnitude at the location of the bunches. We need an estimate for the change of the fundamental (dipole) mode and for the higher-order dipole modes which could steer the beam.

An estimate for the induced transverse field of 90 bunches traveling through the Lolita-X cavity off-axis by 100 μm can be found from the formula (P. Wilson)

$$V_{\perp}^{\rm ind}(n) = \frac{1}{2} I\, r_{\perp}\, L_s\, kx \left(\frac{1 - e^{-2\tau'} - 2\tau' e^{-2\tau}}{1 - e^{-2\tau}} \right) \quad \text{where} \quad \tau' = \tau\; Min\left[1,\; \frac{n\,\tau_{\rm sep}}{T_F}\right] \quad , \tag{11.127}$$

where n equals the bunch number and $\tau_{\rm sep}$ is the separation time between bunches. The expression in brackets reaches its maximum value of 0.15 at $n = 16$. The corresponding maximum induced voltage is 2.5 kV. Dividing by 1.2 MV, the crab-cavity voltage at 250 GeV, gives $V_{\rm ind}/V_{\rm crab} = 0.002$, or 0.2% change. This is comfortably smaller than the allowed 6% calculated in Section 11.7.2.

A rough estimate for higher-order dipole modes is that they might be about 10% of the induced fundamental, or about 0.25 kV for Lolita-X. This voltage could cause a missteering at the IP of $R_{12}(\text{ eV})/E = 17(9\ 10^3)/(500\ 10^9) = 8.3$ nm which is just about equal to 1/40 σ_x. This is also negligible.

Stability Tolerance from Energy Change. The Hamiltonian that gives the desired crab-cavity kick is of the form

$$H = \frac{\theta_c}{2R_{12}}\, x\, z \quad , \tag{11.128}$$

where the kick $\Delta x'$ is given by the derivative with respect to x, and has the linear z-dependence desired. However, there is also a change in energy implied by this Hamiltonian given by the derivative with respect to z. It is

$$\Delta\delta = \frac{\theta_c}{2R_{12}}\, x \quad . \tag{11.129}$$

The change in energy could be $\delta \approx 6 \times 10^{-6}$, which is also the stability tolerance of the final doublet. Inserting values for the crossing angle and R_{12}, we find the condition $x < 6$ mm. This is of no concern as an alignment stability requirement. This number also places a limit on the beam size at the crab cavity. This limit is met since the beam size is on the order of $R_{12}\sigma_{x'} < 0.3$ mm.

Feedback Systems

The physical quantities to be monitored and stabilized are i) the crab-cavity phases with respect to beam arrival times, ii) the crab-cavity voltage, iii) the crab-cavity phase difference between the two crab cavities, and iv) the crab-cavity alignment. The phase with respect to the beam, as pointed out above, can be checked by changing the crab voltage and looking for steering in the outgoing beam. The voltage itself must be set to minimize the horizontal beam size, which one should be able to monitor by vertical and horizontal beam deflection scans. It is equivalent to a beam size measurement.

By installing an insulated return line from each crab cavity to the tee (following the klystron) and comparing the phase of the two return lines, a phase correction can be introduced to keep the return phases stable. A schematic of this stabilizing system is depicted in Figure 11-54. The system is designed to keep large drifts from occurring in the phase difference. In addition, phase jitter must be small, and short-term phase stability will be taken care of by the IP steering feedback system.

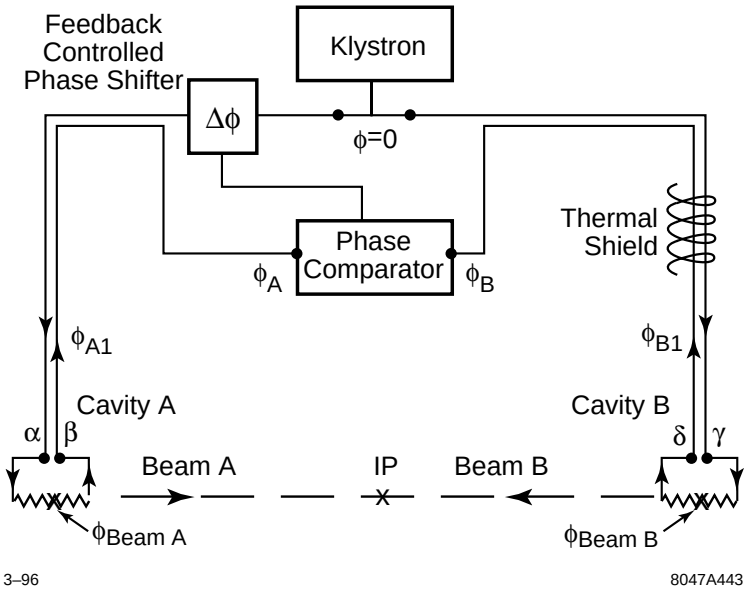


Figure 11-54. *Schematic of a system which stabilizes the phase difference between the two crab cavities.*

11.7.3 Solenoid Effects

We will assume that the maximum solenoidal field strength will be 3 T, and that the maximum crossing angle to be considered is 30 mr. Under these assumptions, the maximum field perpendicular to the beam trajectory is $B_\perp = 450$ Gauss. For this field, $B_\perp/B\rho = \{0.5\times10^{-4}, 0.25\times10^{-4}, 0.2\times10^{-4}\}m^{-1}$ at $E = \{250, 500, 750\}$ GeV respectively.

Skew Quadrupole Aberrations

The skew quadrupole term in the solenoid Hamiltonian is $1/2\, k_s(xp_y - yp_x)$ with $k = B_s/B\rho$. If there were no quadrupole within the solenoid, then this term is a constant (by rotational symmetry) and can be evaluated at the IP, to give an integrated beam-line generator of

$$G = LH = \frac{1}{2}\,\frac{B_s L}{B\rho}\left(x^*\, p_y^* - y^*\, p_y^*\right) \tag{11.130}$$

giving $\Delta y^* = (B_s L_s/B\rho)\, x^*$. Taking $B_x = 3$ T and $L_s = 4$ m leads to $B_s L_s/(B\rho) = \{0.014, 0.007, \text{and } 0.005\}$ for $E = \{250, 500, \text{and } 750\}$ GeV. Since $\sigma_x^* = \{100, 160, 160\}\ \sigma_y^*$, $\Delta y^*_{\rm rms}/\sigma_y^* = \{0.7, 0.6, 0.4\}$. So this term of the generator is small, but still too large by a factor of about five, and must be compensated with a skew quadrupole. We note that it is not the generator $p_x p_y$, which is the main term generated and corrected in the final-focus system. This skew quadrupole may have to precede the final-focus system, or be obtained from two skew quadrupoles on the opposite sides of an IP image in the final-focus system.

The interaction of Q1 and the solenoid will also produce the skew quadrupole generator $p_x p_y$, which may readily be compensated by the final-focus system skew quadrupole.

Steering and Dispersion

The solenoidal field in the region between Q1 and the IP will have a component $B_s\theta_c/2$, and will steer the beam by $\Delta y^* = B_s L^*/B\rho\ \theta_c/2$. For $\theta_c = 20\,\text{mr}$, this has the value $\Delta y^* = \{70,\ 35,\ \text{and}\ 25\}\ \mu\text{m}$ at $E = \{\ 250,\ 500,\ \text{and}\ 750\}$ GeV. This will also be the magnitude of the dispersion generated.

The two beams are deflected apart so steering must be introduced to correct this displacement. If this is done with a corrector at the Q2 end of the doublet, then because (see Table 11-36)

$$\frac{\partial \eta_1}{R_{34}^E \partial y\prime_E} = -2.3 \quad , \tag{11.131}$$

the dispersion generated by the doublet adds to the dispersion generated by the solenoidal field, giving a net dispersion 3.3 times the amount from the solenoidal field alone. This is a huge dispersion, and attempts to correct it will create unacceptable second-order dispersion terms and synchrotron radiation.

The only available alternative is to displace Q1 by $\Delta y_{\text{Q1}} = -\Delta y^*/1.6$ or to introduce an equivalent dipole field in the region of Q1. This will create dispersion which has the opposite sign but does not totally cancel the original. Theoretically, a small section at the IP end of Q1 could be displaced to re-steer the beam, in which case the dispersion would almost totally cancel. This does not work because the synchrotron radiation in such a corrector would be unacceptable. The remaining dispersion is $(1.63 - 1.33)/1.63 \approx$ 1/5 of the original (from Table 11-36: $\Delta\eta = -1.33\Delta y_{\text{Q1}}$).

The remaining residual dispersion of $\{21,\ 11,\ \text{and}\ 9\}\ \mu$m at $E = \{250,\ 500,\ \text{and}\ 750\}$ GeV can be corrected by displacing the incoming beam by $\{11,\ 5,\ \text{and}\ 4\}\ \mu$m. This is at the edge of acceptability vis-a-vis the generation of synchrotron radiation. For 20-mr crossing angles, these numbers are reduced by a factor of 1.5, and are acceptable.

Synchrotron Radiation

We first calculate the additional synchrotron radiation created by the solenoid for the on-axis particles. Then we calculate the effects of the increased dispersion function on the final doublet synchrotron radiation for focused particles (Oide effect + solenoid).

Solenoid Field Alone. The $R_{36}(s)$ function generated by the solenoid in the L^* region is given by

$$R_{36}(s) = \int R_{34}(s) d\theta = \int R_{34}(s) \frac{B_s \frac{\theta_c}{2}}{B\rho}\, ds = \frac{B_s \theta_c}{4B\rho}\, s^2 \quad . \tag{11.132}$$

The synchrotron radiation from this region results in

$$\frac{\Delta y_{\text{IP}}^2}{\sigma_{\text{IP}}^{y2}} = \frac{c_u r_e \lambda_e \gamma^5}{\sigma_{\text{IP}}^{y2}} \int ds\, R_{36}(s)^2 \left|\frac{1}{\rho(s)^3}\right| = \frac{c_u r_e \lambda_e \gamma^5}{\sigma_{\text{IP}}^{y2}} \int ds \left(\frac{B_s\theta_c}{4B\rho}\, s^2\right)^2 \left|\frac{B_s\theta_c}{2B\rho}\right|^3 = \frac{1}{20}\, \frac{c_u r_e \lambda_e}{\sigma_{\text{IP}}^{y2}} \left(\frac{B_s\theta_c L^* \gamma}{2B\rho}\right)^5 \tag{11.133}$$

Assuming $L^* = 2\,\text{m}$, $B_s = 3\,\text{T}$, and $\theta_c = 40\,\text{mr}$, $[\Delta y_{\text{IP}}/\sigma_{y^*}]^2 = \{0.04,\ 0.06,\ 0.06\}$. Fortunately, this term goes like $(B_s\theta_c)^5$, and can be reduced by reducing the crossing angle or the solenoidal field. For $\theta_c = 30\,\text{mr}$, the coefficients are reduced by a factor of 7.6, and $[\Delta y_{\text{IP}}/\sigma_{y^*}]^2 = \{0.005,\ 0.008,\ 0.008\}$. These are already acceptable values.

The contribution from the compensation of the solenoidal steering that lies within Q1 must be added. Here the value of $1/\rho$ is smaller and the R_{36} function from the solenoid plus quadrupole is becoming smaller. This term adds an additional 15% to the above coefficients (at $E = 250\,\text{GeV}$).

Solenoid Field Plus Final Doublet. The effect of the synchrotron radiation (Oide effect) in the quadrupole is enhanced by the presence of the solenoidal field because of the increased dispersion function to the IP. This will be negligible at a 20-mr crossing angle.

Tolerance on Solenoidal Field Strength Stability

If the solenoid field were to change, the orbit to the IP would change and the beams would not collide. With the beam separation proportional to the solenoidal field and between 200 and 80 μm, and a tolerance on the beam separation for a 1/2% luminosity loss of about 2.4 nm, the field must be stable to one part in 10^5. Because the effect is a steering effect, this is a jitter tolerance and should be met quite easily.

11.7.4 Summary

From the point of view of the requirements on the crab cavity and the solenoidal field, the considerations of this section suggest that the 20-mr crossing angle is acceptable. This allows us to couple the big bend angle (used to obtain IP separation and reduce muon backgrounds) with the crossing angle, and to have the space to use permanent magnets for the first entrance and exit quadrupoles. The permanent magnet exit quadrupole is helpful by providing focusing close to the IP (which is necessary because of the large exit beam divergence) securing the high-quality beam removal optical path crucial for post-IP diagnostics. Permanent magnet fields can be quite stable under thermal change.

It is an important result that the first quadrupole must be displaced by up to 45 μm when the solenoidal field is turned on to correct the dispersion. This is a large movement compared to the nanometer stability required for this quadrupole under colliding conditions.

An experiment with an X-band phase mixer should be undertaken to confirm the possibility of maintaining a 0.2° relative phase tolerance between the two crab cavities.

11.8 The Beam Extraction and Diagnostic System (The Dump Line)

The main functions of this system are to optimize luminosity, to characterize beam properties at the interaction point, and to transport beams from that point to a dump with minimal backgrounds in the detector from these or any secondary functions that may occur. Beam characterization includes the measurement of current, position, profile, polarization, energy, and low-order correlations on a bunch-to-bunch basis for feedback and stabilization. A magneto-optical and diagnostic layout is described that provides such functions. We also consider the possibilities for e, μ, n, and γ secondary lines and dump experiments as well as energy recovery and local reuse of an assumed 10 MW of power in each 500-GeV beam. An earlier condensed version of this section was given in [Spencer 1995b].

To accomplish all of this, we need to know the detailed composition and characteristics of the outgoing beam(s) under different, possible circumstances. Clearly, these characteristics depend on those assumed for the incoming beam. Based on an assumed set of incoming beam parameters and transport optics we then make predictions for the outgoing beams that are used to guide the design of the outgoing, dump line optics. Typically, these beams have a significant number of pairs and more photons than leptons. Thus, if background simulations for the detector then imply an unacceptable situation, this procedure has to be iterated until a consistent solution is achieved.

Our approach was to develop a "straw man" design, *i.e.*, to assume a set of upstream characteristics that are reasonable but conservative to provide a worst-case scenario that emphasizes the major problems in constructing the transport system. The current design is a reasonably versatile one that takes the primary, outgoing disrupted beam (including photons) from the IP to the dump with minimal loss of particles. We have also included the low-energy pairs that are produced during the strong beam-beam interaction. We have found no basic limitations on the achievable luminosity imposed by the dump line. Furthermore, it provides a range of configurations that allow a variety of options.

While our primary goal was to optimize the luminosity while disposing of the various secondary beam components, it is also important to provide any monitoring and feedback that can optimize the average, usable collision rate at the IP. Thus, beyond simply dumping the beam, there are other functions that run from the necessary to the desirable that will be considered a part of the dump line:

- Beam control and stabilization,
- Diagnostics and monitoring - including luminosity and polarization,
- e, μ, n, γ secondary beams and parasitic experiments,
- Polarized sources for γ, μ and e^+ beams, and
- Energy recovery.

Due to the high power in the outgoing photon beam and the cost of beam dumps, we decided that the photons and primary leptons should share a common beam dump. This implies an available distance for beam studies of $\approx$150 m before the outgoing, high-energy photons begin to cause problems for the present dump design.

Because the SLC was a prototype for the NLC, we begin by reviewing the SLC and the FFTB experience relative to the NLC design. Next, we summarize the more important working assumptions in terms of a table of parameters characterizing the input and the output optics and beam characteristics at the IP. Next we describe the design procedure and give some intermediate results that were used to arrive at the final design. A horizontal chicane is used to separate the outgoing photon and electron beams, which are comparable in number and density, to allow separate experiments, before both beams are recombined into a common dump. Some rather sophisticated measurements and capabilities are possible over the available distance that are compatible with a low-background dump. We have also tried to accommodate the possibility of beam dump experiments. However, various additional simulations need to be done to confirm this and certify specific designs. Certain hardware is discussed where relevant. Finally, we summarize our conclusions and discuss the work that still needs to be done.

11.8.1 Comparison with the SLC

One advantage for the NLC dump design relative to that for the SLC is a horizontal crossing angle ($\theta_{c,x}$=2$\times$10 mr) at the IP that allows us to avoid kickers and septa for separating the counterpropagating beams. A related but more subtle advantage of this is that it allows us to reverse the polarities between the ingoing and outgoing IP doublets for better control of the larger outgoing horizontal disruption angle and emittance.

The much smaller vertical spot sizes assumed for the NLC overcome other differences to make its disruption larger. Although this complicates the disposal problem, it still allows all of the SLC measurements to be done as well as others that may not be practical there. For example, we can measure the energy loss distribution and use precise rf BPMs in the beginning of the line to measure position, angle, and timing of individual bunches where this may be impractical for the SLC due to its larger peak deflection angles that produce rms outgoing angular spreads that are four times larger than our worst case.

However, the outgoing, disrupted energy spreads and energy offsets calculated for the NLC are more than a factor of twenty larger. The outgoing beam also includes a significant number of synchrotron radiation photons that can produce low-energy pairs quite apart from the beamstrahlung. Because the emittance and disruption angle is largest in the horizontal, this implies that we want our analyzing magnets to bend horizontally which should also simplify any beam dump experiments.

Rather than 30 kW in each SLC beam, one has to deal with nearly 10 MW in each NLC beam. There is a factor of ten in energy for 500-GeV beams in the NLC, the same rf pulse repetition rate of 120/s, and a factor of twenty or so in beam current per rf pulse from accelerating a multibunch train in each pulse. This increased beam power poses certain problems for intercepting detectors and implies significantly higher operating costs. For 10¢/kWh, this represents a potential refund of as much as $45K/day if energy is restored to the grid or otherwise recycled before it is dissipated in the dump.

11.8.2 Basic Design Procedures, Constraints, and Assumptions

The optics can only be realized after we know the characteristics of the outgoing beam from the IP. The procedure was to take the upstream final-focus design in the form of TRANSPORT [Brown 1977] and convert this to DIMAD [Servranckx 1990] for predicting the spot characteristics at the IP due to emittance growth from synchrotron radiation in the dipoles and quadrupoles. One assumption was that the effect of the energy loss on the optics, especially in the quadrupoles, can be neglected. The predicted beam parameters resulting from DIMAD were then used in ABEL91 [Yokoya 1986] to predict the composition and characteristics of the outgoing beam as well as to produce ray sets for all particle types for tracking. The ABEL results were checked against analytic calculations where possible and used to simulate the dump line with TRANSPORT (for design) and TURTLE (for tracking) [Carey 1982]. Table 11-40 gives several configurations that were used with those labelled `A' the current ones for 250- and 500-GeV incident beams.

In the Table, $\mathcal{L}_{e^\pm}$ and $\mathcal{L}_{e^-e^-}$ are the predicted incident channel luminosities for these parameters whereas $\mathcal{L}_{\rm Compton}$ and $\mathcal{L}_{\gamma\gamma}$ are secondary to the e^+e^- channel. $\mathcal{L}_{\gamma\gamma}$ is rather large because it includes contributions from real and virtual photons and the beamstrahlung parameter Υ is fairly large. Υ is an important measure of the QED and QCD backgrounds expressed in terms of $\gamma\gamma \rightarrow e^+e^-$ pairs and minijets ($\gamma\gamma \rightarrow$ X + anything). Such parameters are described in more detail below. Because there is still debate on how to calculate the hadronic backgrounds in terms of the various possible quark and gluon contributions at these energies, they are not listed. There is interesting physics available here if it could be measured but an upgraded SLC would be a better place to find new hadrons, quark molecules or glueballs.

Because electrons and positrons are generally interchangeable, we will often talk about electrons but all parameters such as N_B, the number of electrons (or positrons) per bunch apply equally to both beams unless otherwise stated. In fact, N_B can also be the number of photons per bunch [Spencer 1995c] in a general linear collider (GLC). The last column for 500-GeV beams was our primary case for the optics design. In the absence of other input, we used only the **disrupted** beam from this worst case having the lowest $\beta^*_{x,y}$ and highest Υ. The numbers in the table and the ray sets that were used were predicted using a modified version of ABEL that was made operational on the SLAC VM system. Calculations were done for all of the columns shown and the output was archived for general use.

The length of the extraction line was set by the outgoing beamstrahlung photons and their predicted size at the dump:

$$12\sigma^\gamma_{x'} \times L_D \leq 2R_D \ , \tag{11.134}$$

where L_D is the distance from the IP to the dump face and R_D is its radius. For a dump window of 20 cm in diameter and outgoing angular spreads of $\sigma_{x',y'}$=92,43 μr, taken from Figure 11-58, we have ≈150 m of usable length available. The apertures of the outgoing magnets were sized to pass these photons (and pairs) outside the detector so that the outgoing photons pass through the system into the dump without serious loss.

$\sqrt{s}_{\rm NLC}$	(GeV)	500	500A	1000	1000A	`Worst'
$f_{\rm rep}$	(Hz)	180	180	120	120	120
n_B		90	90	75	90	75
$\gamma\epsilon_y$	(10^{-8} m)	5	8	5	10	5
N_B	(10^{10})	0.65	0.65	1.1	0.95	1.1
σ_x^*/σ_y^*	(nm)	319.7/3.2	285.9/4.52	360/2.3	226.1/3.57	245/2.5
σ_z	(μm)	100	100	100	125	125
β_x^*/β_y^*	(mm)	10/0.1	8/0.125	25/0.1	10/0.125	10/0.1
$\mathcal{L}_G$	($10^{34}{\rm cm}^{-2}{\rm sec}^{-1}$)	0.53	0.42	1.05	0.96	1.42
D_x/D_y		0.073/7.3	0.090/5.70	0.049/7.60	0.132/8.33	0.104/10.2
$\theta_D = \theta_{x,\max}$	(μrad)	233	257	175	238	256
Υ		0.09	0.10	0.27	0.30	0.32
$\mathrm{H}_D \equiv \mathcal{L}/\mathcal{L}_G$		1.37	1.42	1.30	1.36	1.56
$\mathcal{L}_{e^\pm}$	($10^{34}{\rm cm}^{-2}{\rm sec}^{-1}$)	0.73	0.60	1.37	1.30	2.20
$\langle s\rangle/s_{\rm NLC}$			0.972		0.913	
$\langle \mathrm{E}_e - \mathrm{E}_\diamond\rangle/\mathrm{E}_\diamond$		0.019	0.032	0.0711	0.103	0.127
$\delta_B \equiv \delta\mathrm{E}_e^{\rm rms}/\mathrm{E}_\diamond$		0.042	0.065	0.120	0.143	0.167
$\mathcal{L}_{100}/\mathcal{L}_{e^\pm}$		0.441	0.376(0.405)	0.305(0.323)	0.195(0.218)	0.182(0.206)
$\mathcal{L}_{90}/\mathcal{L}_{e^\pm}$		(0.83)	(0.80)	(0.84)	(0.75)	(0.69)
N_γ/N_B		0.88	0.98	1.21	1.67	1.77
$\langle \mathrm{E}_\gamma\rangle/\mathrm{E}_\diamond$		0.027	0.033	0.059	0.062	0.072
N_p/N_B	(10^{-7})[a]		3.95		8.69	
$\langle \mathrm{E}_p\rangle$	(MeV)[a]		23.4		25.8	
$\mathcal{L}_{\rm Compton}$			0.23		0.80	
$\mathcal{L}_{\gamma\gamma}$	($10^{34}{\rm cm}^{-2}{\rm sec}^{-1}$)		0.10		0.56	
$\mathcal{L}_{e^-e^-}$			0.23		0.45	

[a] N_p/N_B is the ratio of outgoing positrons per incident electron for energies $\mathrm{E}_p > 2$ MeV and angles $\theta \geq 10$ mr.

Table 11-40. *Beam-beam effects for e^+e^- at $E_{\rm cm} = 0.5$, 1.0 TeV and $\gamma\epsilon_x = 5 \times 10^{-6}$ m. Quantities enclosed in (...) are calculated analytically rather than from the simulations.*

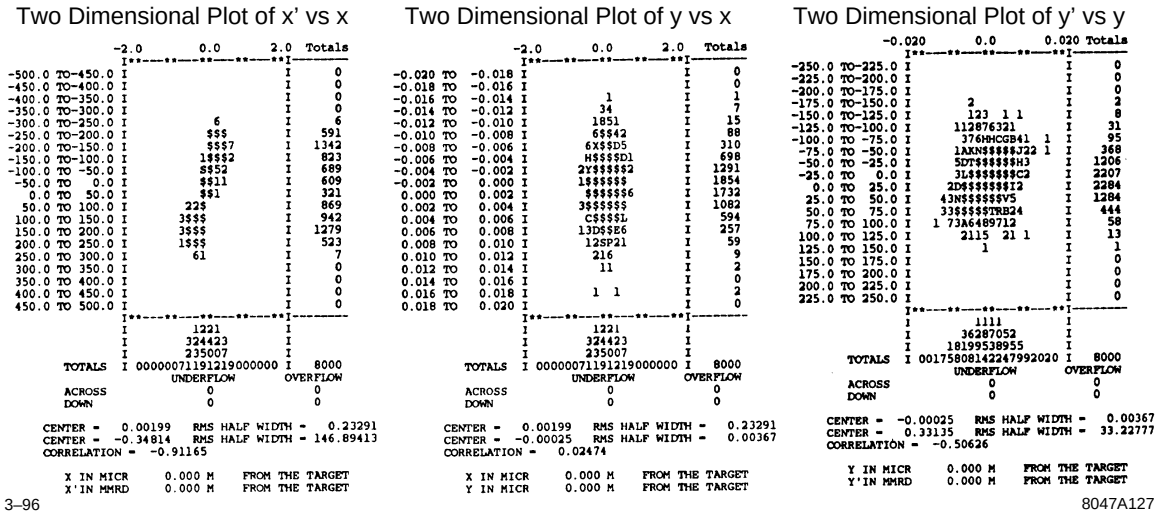


Figure 11-55. *Disrupted beam distribution at the IP for the `Worst' case in Table 11-40*

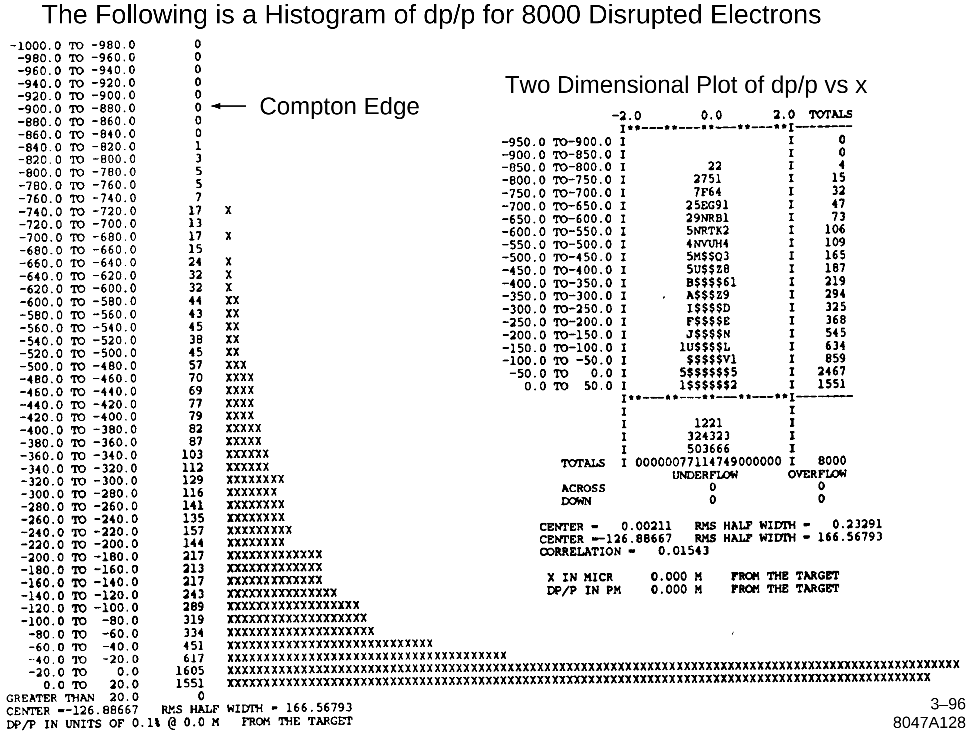
11.8.3 The Beam-Beam Calculations

Histograms of our initial working set of 8000 disrupted particles are shown in Figures 11-55 and 11-56. From these figures we see the shearing of the upright ellipses from the focusing by the other beam just after the interaction as well as emittance growth in both transverse and longitudinal directions. However, the energy loss and spreading with a very long tail extending nearly down to the Compton edge (50 GeV for 500-GeV beams and a Nd:YAG laser) is probably the most significant effect. These and the subsequent histograms based on them were taken directly from third-order TURTLE. The beamstrahlung distributions that are shown next were also done with TURTLE for uniformity.

Figures 11-57 and 11-58 show the outgoing photon ray set for the same input beam. From the Table there are actually more outgoing photons than electrons but a thinning factor was used. When the configurations are frozen we will enlarge the ray sets and also characterize them analytically. The average photon energy and distribution looks very much like a bremsstrahlung spectrum and relates directly to that for the outgoing disrupted electrons.

The effective, transverse photon emittances are typically smaller than the disrupted, charged particle beam because of the differing effects of disruption. Although this is discussed later, we note that the photon spot size provides a better, realtime measure of the incoming, unperturbed beam profile than the disrupted electrons. The outgoing photons were important in determining the length and acceptance of the extraction line because they can't be controlled in the same way as charged beams and because they have enough high energy photons to allow a good, nonintercepting profile monitor for bunch-to-bunch measurements outside of the solenoid. Figures 11-59 and 11-60 show the TURTLE output at various distances from the IP after tracking the beams through the quadrupoles inside the detector. From these figures and Figure 11-65 below, we see from the rms beam sizes and the numbers of particles outside the picture frame, that the disrupted primary beams and their associated beamstrahlung should clear the quadrupole apertures and solenoid with negligible loss.

From these figures, the primary disrupted beam is still the determinant of the magnet apertures but neither beam imposes any severe restrictions—at least while they are in the solenoid. The most damaging effect is then expected to come from the low-energy pairs. We have calculated their spectrum and detector simulations are now being done. This is complicated by the low energies, the crossing angles, and the strong solenoidal fields at the IP that necessarily couple the transverse coordinates.



The Following is a Histogram of dp/p for 8000 Disrupted Electrons

```
-1000.0 TO -980.0      0
 -980.0 TO -960.0      0
 -960.0 TO -940.0      0
 -940.0 TO -920.0      0
 -920.0 TO -900.0      0
 -900.0 TO -880.0      0
 -880.0 TO -860.0      0
 -860.0 TO -840.0      0
 -840.0 TO -820.0      1
 -820.0 TO -800.0      3
 -800.0 TO -780.0      5
 -780.0 TO -760.0      5
 -760.0 TO -740.0      7
 -740.0 TO -720.0     17   X
 -720.0 TO -700.0     13
 -700.0 TO -680.0     17   X
 -680.0 TO -660.0     15
 -660.0 TO -640.0     24   X
 -640.0 TO -620.0     32   X
 -620.0 TO -600.0     32   X
 -600.0 TO -580.0     44   XX
 -580.0 TO -560.0     43   XX
 -560.0 TO -540.0     45   XX
 -540.0 TO -520.0     38   XX
 -520.0 TO -500.0     45   XX
 -500.0 TO -480.0     57   XXX
 -480.0 TO -460.0     70   XXXX
 -460.0 TO -440.0     69   XXXX
 -440.0 TO -420.0     77   XXXX
 -420.0 TO -400.0     79   XXXX
 -400.0 TO -380.0     82   XXXXX
 -380.0 TO -360.0     87   XXXXX
 -360.0 TO -340.0    103   XXXXXX
 -340.0 TO -320.0    112   XXXXXX
 -320.0 TO -300.0    129   XXXXXXXX
 -300.0 TO -280.0    116   XXXXXXX
 -280.0 TO -260.0    141   XXXXXXXX
 -260.0 TO -240.0    135   XXXXXXXX
 -240.0 TO -220.0    157   XXXXXXXXX
 -220.0 TO -200.0    144   XXXXXXXX
 -200.0 TO -180.0    217   XXXXXXXXXXXXX
 -180.0 TO -160.0    213   XXXXXXXXXXXXX
 -160.0 TO -140.0    217   XXXXXXXXXXXXX
 -140.0 TO -120.0    243   XXXXXXXXXXXXXXX
 -120.0 TO -100.0    289   XXXXXXXXXXXXXXXXXX
 -100.0 TO  -80.0    319   XXXXXXXXXXXXXXXXXXX
  -80.0 TO  -60.0    334   XXXXXXXXXXXXXXXXXXXX
  -60.0 TO  -40.0    451   XXXXXXXXXXXXXXXXXXXXXXXXXXXX
  -40.0 TO  -20.0    617   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  -20.0 TO    0.0   1605   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
    0.0 TO   20.0   1551   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
GREATER THAN  20.0     0
CENTER =-126.88667   RMS HALF WIDTH = 166.56793
DP/P IN UNITS OF 0.1% @ 0.0 M   FROM THE TARGET
```

Two Dimensional Plot of dp/p vs x

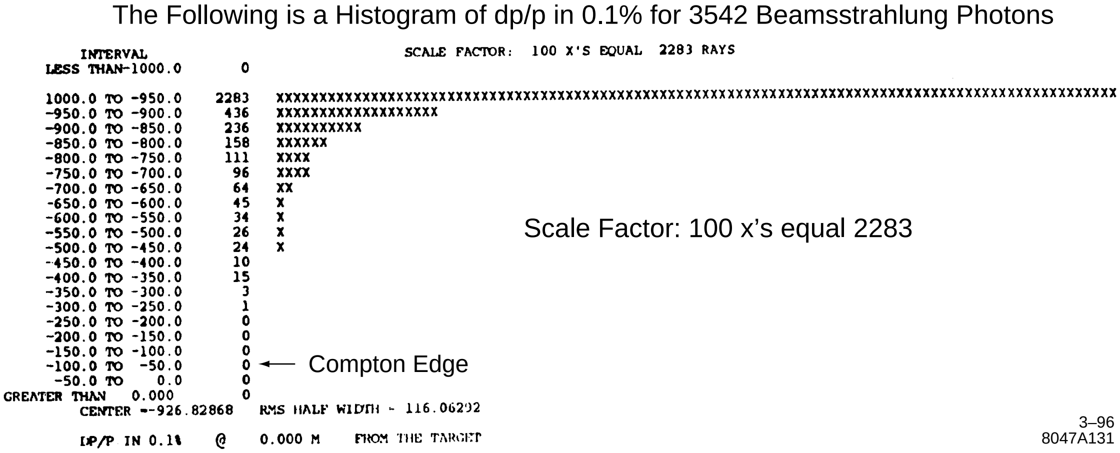
```
                       -2.0         0.0         2.0  TOTALS
                       I**----**----**----**----**I--------
-950.0 TO-900.0 I                          I       0
-900.0 TO-850.0 I                          I       0
-850.0 TO-800.0 I            22            I       4
-800.0 TO-750.0 I           2751           I      15
-750.0 TO-700.0 I           7F64           I      32
-700.0 TO-650.0 I          25EG91          I      47
-650.0 TO-600.0 I          29NRB1          I      73
-600.0 TO-550.0 I          5NRTK2          I     106
-550.0 TO-500.0 I          4NVUH4          I     109
-500.0 TO-450.0 I          5M$$Q3          I     165
-450.0 TO-400.0 I          5U$$Z8          I     187
-400.0 TO-350.0 I          B$$$$61         I     219
-350.0 TO-300.0 I      .   A$$$Z9          I     294
-300.0 TO-250.0 I          I$$$$D          I     325
-250.0 TO-200.0 I          F$$$$E          I     368
-200.0 TO-150.0 I          J$$$$N          I     545
-150.0 TO-100.0 I         1U$$$$L          I     634
-100.0 TO -50.0 I          $$$$$V1         I     859
 -50.0 TO   0.0 I         5$$$$$$$5        I    2467
   0.0 TO  50.0 I         1$$$$$$$2        I    1551
                I**----**----**----**----**I--------
                I                          I
                I            1221          I
                I           324323         I
                I           503666         I
        TOTALS  I 0000007711474900000 0    I    8000
                          UNDERFLOW          OVERFLOW
        ACROSS                0                 0
        DOWN                  0                 0

    CENTER =    0.00211   RMS HALF WIDTH =    0.23291
    CENTER =-126.88667   RMS HALF WIDTH = 166.56793
    CORRELATION =    0.01543

     X IN MICR      0.000 M    FROM THE TARGET
     DP/P IN PM     0.000 M    FROM THE TARGET
```

3–96
8047A128

Figure 11-56. *Disrupted energy distribution at the IP in units of 0.1% for the `Worst' case in Table 11-40. The insert shows the correlation with the horizontal position* x*.*



The Following is a Histogram of dp/p in 0.1% for 3542 Beamsstrahlung Photons

```
      INTERVAL                              SCALE FACTOR:  100 X'S EQUAL  2283 RAYS
   LESS THAN-1000.0        0

   1000.0 TO -950.0     2283   XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   -950.0 TO -900.0      436   XXXXXXXXXXXXXXXXXXX
   -900.0 TO -850.0      236   XXXXXXXXXX
   -850.0 TO -800.0      158   XXXXXX
   -800.0 TO -750.0      111   XXXX
   -750.0 TO -700.0       96   XXXX
   -700.0 TO -650.0       64   XX
   -650.0 TO -600.0       45   X
   -600.0 TO -550.0       34   X
   -550.0 TO -500.0       26   X
   -500.0 TO -450.0       24   X
   -450.0 TO -400.0       10
   -400.0 TO -350.0       15
   -350.0 TO -300.0        3
   -300.0 TO -250.0        1
   -250.0 TO -200.0        0
   -200.0 TO -150.0        0
   -150.0 TO -100.0        0
   -100.0 TO  -50.0        0
    -50.0 TO    0.0        0
GREATER THAN   0.000       0
      CENTER =-926.82868   RMS HALF WIDTH = 116.06292

      DP/P IN 0.1%    @    0.000 M    FROM THE TARGET
```

3–96
8047A131

Figure 11-57. *"Worst" case outgoing photon energy distribution from Table 11-40.*

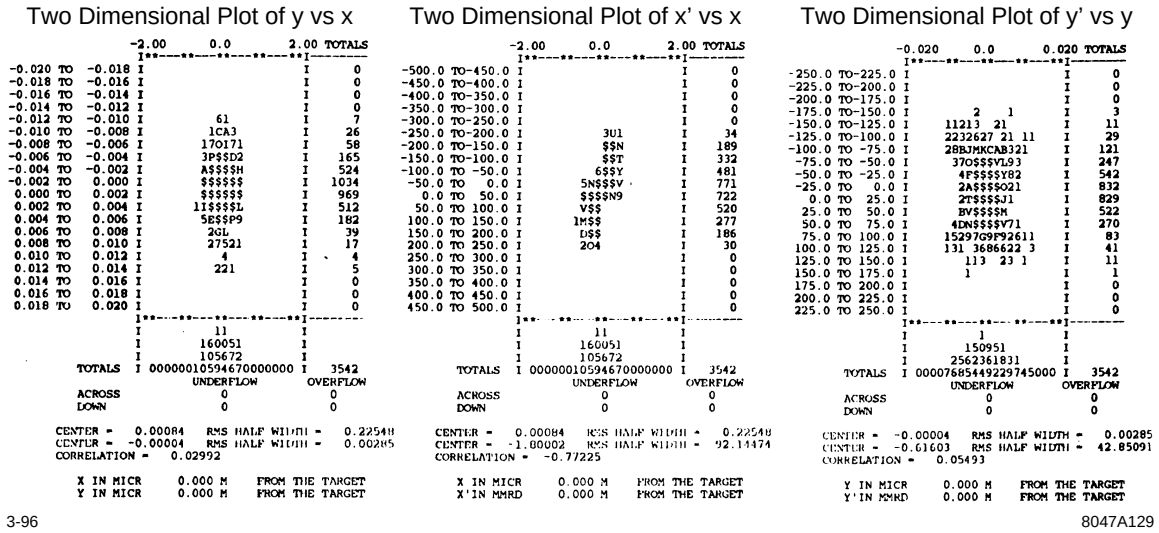


Figure 11-58. *`Worst' case transverse, outgoing photon distribution from Table 11-40.*

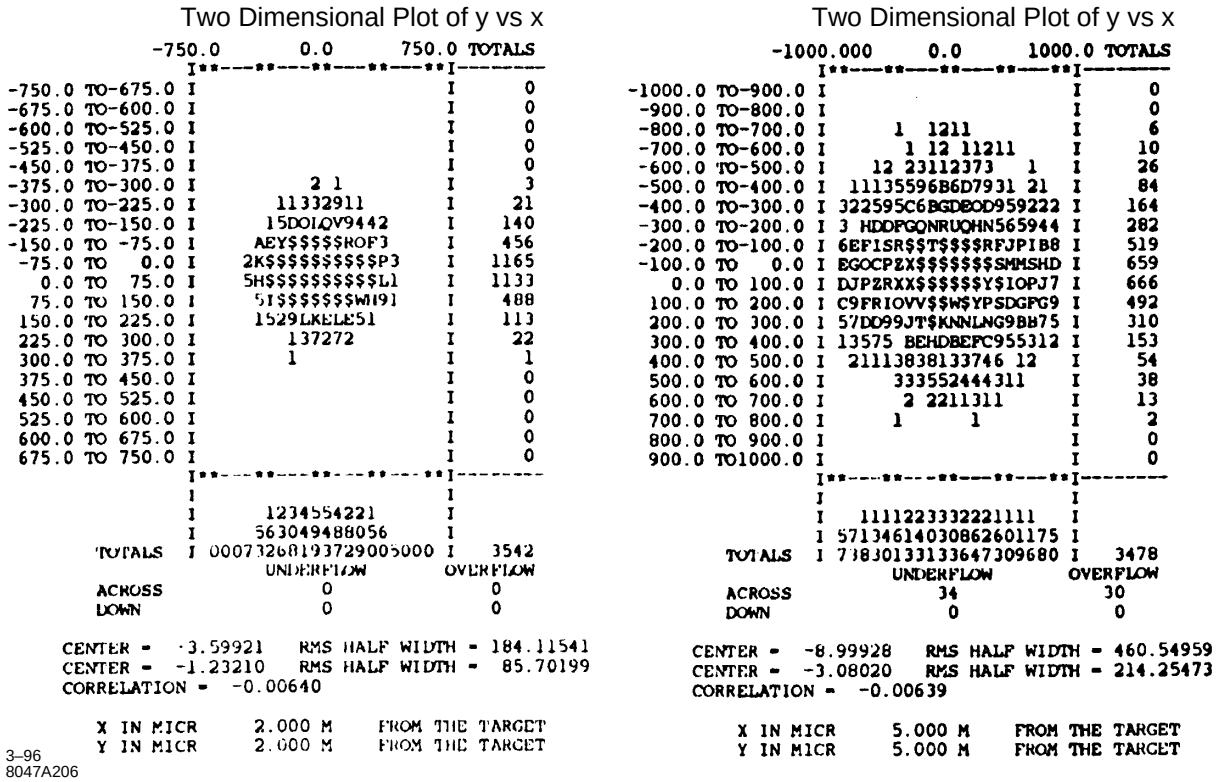


Figure 11-59. *Transverse, outgoing photon distribution at various distances from the IP in the solenoid (L_{sol}=5 m). The picture frame is ≥1/8 of the relevant quadrupole aperture.*

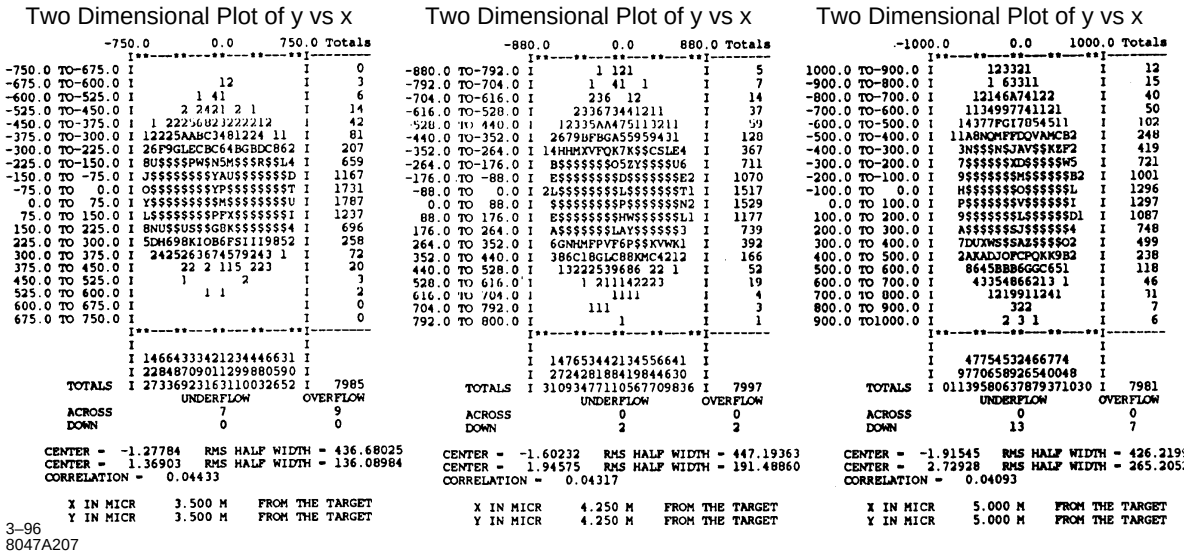


Figure 11-60. *Transverse, disrupted beam distribution at various distances from the IP in the solenoid. The picture frames are always ≥ 1/8 of the nearest quadrupole aperture.*

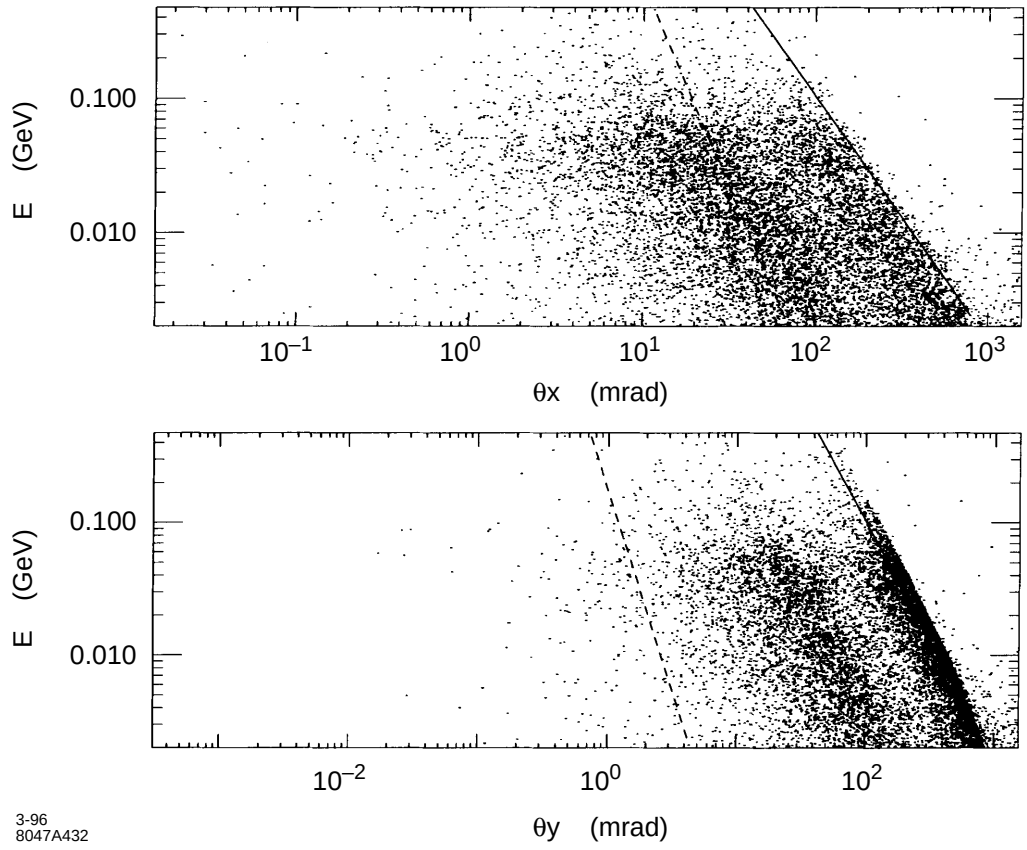


Figure 11-61. *Outgoing charged pairs at 1 Tev for Case 1000A in Table 11-40.*

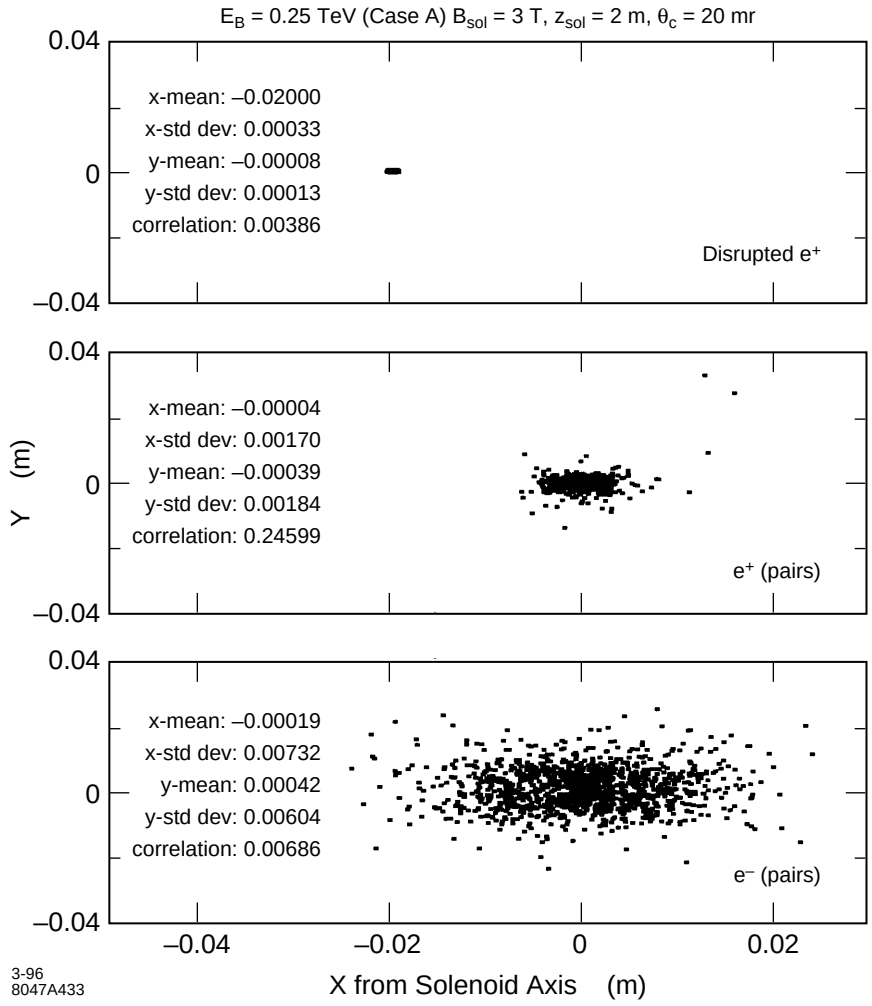


Figure 11-62. *Outgoing charged beams at 2 m from the IP in a strong solenoidal field.*

Such beams necessarily experience and are influenced by nonlinear fields in both the transverse and longitudinal directions before exiting the solenoid. Of equal importance, the response of an outgoing pair to the collective beam field after they are produced depends drastically on the charge and energy of each component.

Figures 11-61–11-64 show some characteristic predictions for the Case A columns in Table 11-40 corresponding to cuts $10 \leq E \leq 500$ MeV and $\theta \geq 10$ mr. The figures illustrate how the different species of outgoing particles complicate the design of the line—especially when it is desired to pass them without intercepting any (except via the required magneto-optical fields). While there are comparatively large angular and energy spreads between the incoming and outgoing electron and photon beams as shown above, the pairs are much worse and also include two opposite charge states. Figure 11-61 and 11-64 show this when we allow the pairs to drift 10 m after leaving the solenoid. In these figures, we assumed two different solenoid lengths for comparison. The results can be compared to the previous estimate of the length of the extraction line based on the outgoing photon beam characteristics. Fortunately, the pairs do not change our previous design in any essential way as we will discuss.

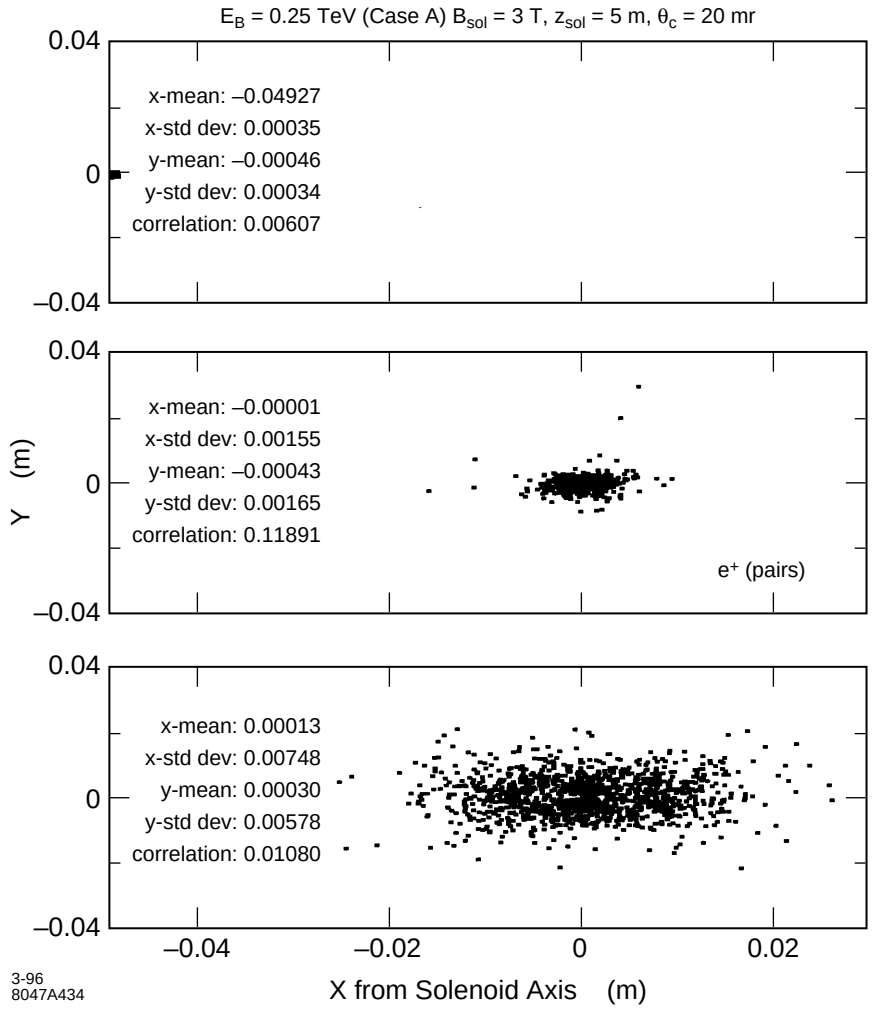


Figure 11-63. *Outgoing charged beams at 5 m from the IP in a strong solenoidal field.*

Besides constraining our optical and hardware design options, this forced us to continually constrain the apertures as well as the field integrals during the design optimization to ensure viable magnets that allow all of these beams to pass. Other constraints were related to the requirements imposed by the different possible measurements and experiments that might be required. For example, the electron spin rotation and depolarization from the bends and the beam-beam interaction as well as the energy resolution necessary to resolve the low-energy electrons near the Compton edge (the most sensitive ones for monitoring beam polarization) constrained the strength and disposition of the dipoles. A related constraint was the need to capture off-energy bunches from the linac or ones that did not collide and therefore lose any energy from the strong beam-beam interaction. Lastly, because the pairs also carry information about the luminosity in terms of the beam sizes at the IP, we wanted to pass them outside the solenoid to make measurements on them before dumping them in separate, reentrant dumps.

The simplest procedure is to sample the higher energy part of the pair spectrum because this is the most accessible and easiest to measure with the necessary resolution. The pair dumps are modest in both size and cost as indicated in

Table 11-40. The relative proportion of Breit-Wheeler, Bethe-Heitler and Landau-Lifshitz pairs depends on the cuts, but Breit-Wheeler production from two real photons is generally negligible at those energies.

We have laid out the essential assumptions and presented the results of calculations that provided the various input beams required for the detector simulations and the optics design. On the assumption that the detector studies show no irreconcilable conflicts with the ingoing or outgoing IR optics, the problems to be solved become reasonably well defined.

For example, given that l^*=2 m, we showed what the distributions of the outgoing beam components look like along the solenoid axis, *i.e.*, what the masking, the magnets, and the diagnostics within the detector must look like to accommodate such beams and provide the required beam stay-clear. For these outgoing beams, we anticipated the optics design and showed that we can get all the beam components out of the solenoid with virtually no loss if we can design rather strong quadrupoles with an open structure that can achieve their assumed characteristics in a strong solenoidal field ($\approx$3 T). A stronger solenoid for containment of the lower energy pairs appears desirable but this depends on the topology of the aperture, the crossing angles, etc. Ultimately, it is a question of the cost of the ideal structure *vs.* the practical question of what the detector design can accommodate. The next section gives the complete optics design from the IP to the dump that allows a variety of configurations and hardware for the options discussed in the introduction. Subsequent sections discuss these options and their hardware.

11.8.4 Beam Line Optics

In the SLC, the angular distribution predicted for the outgoing disrupted beam due to the strong focusing from the counterpropagating beam at the IP was a major factor in sizing the apertures of the strong IR quadrupoles. This was $\approx 10\ \sigma^*_{x'}$ for round beams which translates to $\approx 10\ \sigma_x$ at the exit of the last quadrupole adjacent to the IP. For the NLC, a closely analogous situation exists concerning the outgoing angles but perhaps the most important characteristic that dictates the optics is the energy spread in the outgoing beam. This also complicates the energy recovery for various methods such as rf deceleration where we might want the beam to pass through an accelerating structure where the energy-phase structure of the bunch at the cavities becomes important as does the transverse size. In the NLC, there will also be a significant number of lower-energy particles from beamstrahlung as well as pair production and the effective disruption parameter for such particles, that scales inversely with their energy, can become very large.

We note that the primary problem of transporting the outgoing beam to the dump is much simpler after it exits the detector. Figures 11-59 to 11-64 show the outgoing electrons, positrons and photons after various quadrupoles for various assumptions about the detector—typically assumed to end at 5.0 m from the IP. These results were computed through third-order with TURTLE and are dominated by the low order, chromo-geometric terms. In each Figure, the picture frame is $\leq$1/8 of the magnet aperture and the beam σ's are given in μm and μr. In all cases, the pole-tip fields and apertures of the magnets (including the beam dump) were sized to give $R_M \geq 12\sigma_{x,y}$ for the outgoing photons and disrupted electrons. In the initial doublet, that extends beyond 10 m from the IP, the apertures are $\geq 15\sigma_{x,y}$.

A number of configurations were tested and shown to be adequate under the stated assumption that we must take **all** photons and electrons to the beam dump with the stated size limit. Figure 11-65 shows one possible configuration that has space available for experiments, monitoring, and higher-order multipoles. It satisfies our criteria using only dipoles and quadrupoles. Further, some of the quadrupoles that are shown here were not excited for this particular optics. They act as potential tuning elements, *e.g.*, if QDD3 is turned off, the full energy, undisrupted beam will propagate to the dump in a point-to-parallel configuration. Lower energy particles will be overfocused so that the envelope of electrons tracked with TURTLE intersects that for the photons near 19 m from the IP. Thus, our first space available for beam size and polarization measurements comes between 10.5–16.2 m.

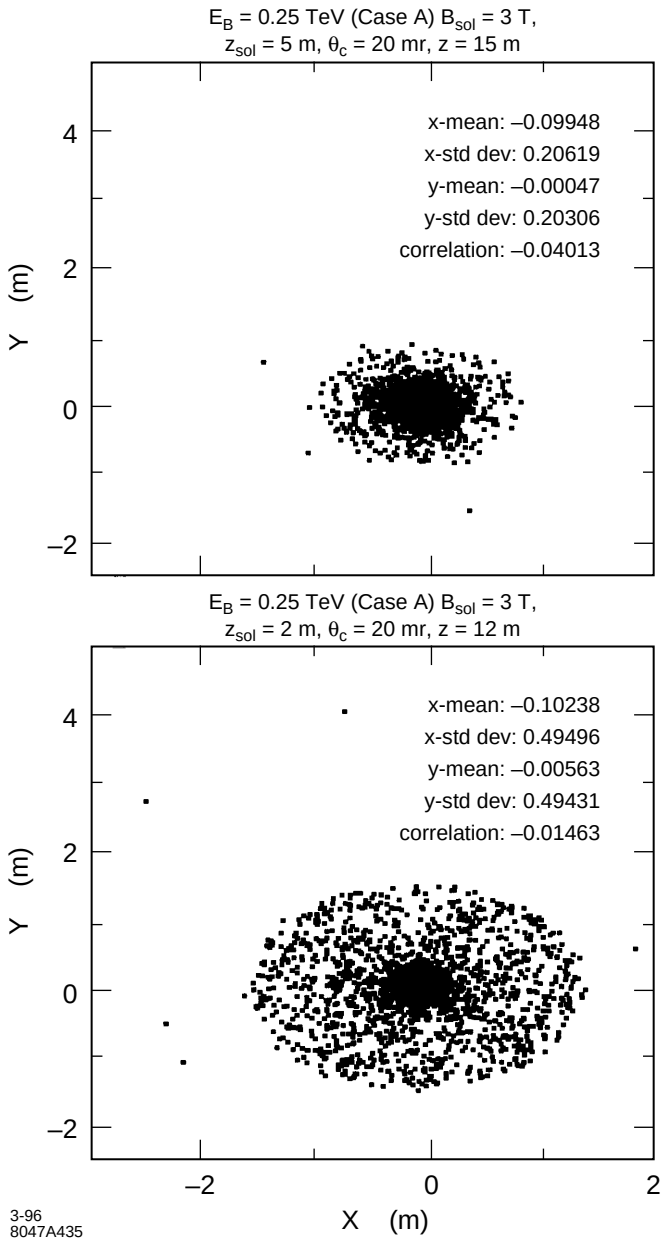


Figure 11-64. *Outgoing beams 10 m after a strong solenoidal field with $L_{\rm sol} = 2$ and 5 m.*

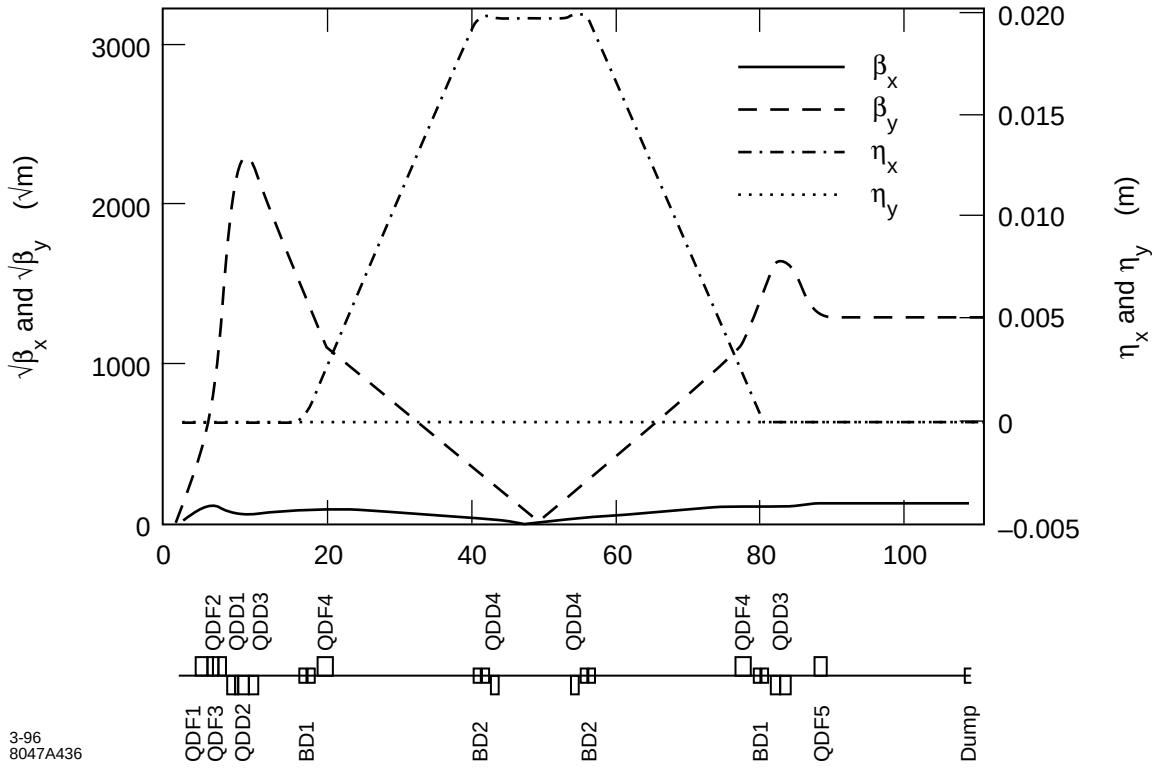


Figure 11-65. *Optical layout for the NLC dump line for the `worst' case in Table 11-40.*

The first dipole of the horizontal chicane, used to separate the outgoing photon and electron beams, begins at 16 m. It allows separate e and γ measurements or experiments before recombining both beams into a common dump at 110 m. The various bends in the chicane can serve as spectrometers as well as the chicane itself. Figure 11-65 shows the Twiss functions when the four bends are sized to separate the two beams by 12 sigma. Their maximum separation is

$$\Delta x = 2\rho_B(1-\cos\theta_B) + L_1\tan\theta_B \ , \tag{11.135}$$

where θ_B is the bend angle for the full energy of any one rectangular dipole of length L_B=$\rho_B sin\theta_B$ and L_1 is the separation between bends BD1 and BD2. Notice that this is just the dispersion η_x in the center of the chicane. This separation requires a minimum distance of

$$L_{\min} = 2\rho_B\sin\theta_B + L_1 \ . \tag{11.136}$$

The change in the bunch separation, due to the chicane, after this point in the central region is

$$R_{56} = \frac{\delta l_z}{\delta p/p} = 2\rho_B(\tan\theta_B - \theta_B) + L_1\tan^2\theta_B \ . \tag{11.137}$$

This is proportional to the rf phase shift [Spencer 1995a]. Thus we have a common beam pipe and sufficient dispersion to measure the energy and spread of the undisrupted beam.

For example, if we want to use the first bend for analyzing low-energy particles from the IP or from a laser interaction before this bend, then the first-order resolving power for some downstream location L is

$$R_1(\rho,\theta,L) = \frac{\rho(1-\cos\theta) + L\cdot\tan\theta}{[x_i\cos\theta + x_i'(\rho\sin\theta + L/\cos\theta)]} \ . \tag{11.138}$$

In the middle of the chicane R depends on the optics we impose. $R_{12}\rightarrow$0 for point-to-point so R=$\Delta x/10\sigma_x^*$=8000 for a magnification of 10. This region of the chicane can then resolve a single-beam, undisrupted energy spread of δp/p=0.0125% while the region directly in front of BD2 gives $R\approx$800 or a 0.13% capability.

Notice that there were several factors that constrained the bends: electron spin rotation as well as the energy resolution necessary to resolve low-energy electrons near the Compton edge (required for monitoring beam polarization). Further, dipoles drive many higher-order aberrations that act to blow the beam up that require higher multipoles to correct. These were not needed to get the beams into the dump with the 12-sigma constraint through the line.

Characteristic Angles and Some Related Disruption Effects

There are several processes that influence the beam and various characteristic angles that deserve discussion. The typical emission angle for most high-energy radiative processes such as incoherent bremsstrahlung is $\theta_r \approx 1/\gamma \approx 1\,\mu$r at 500 GeV. As the photon energy decreases, the angles begin to grow compared to θ_r as does the interaction volume of the electron. When the electron interacts with the collective field of the other bunch, it sees transverse electric and magnetic fields that are nearly equal. It is easy to show that these fields can bend an electron by angles significantly greater than θ_r so this angle is physically interesting. The resulting radiation, similar to synchrotron radiation, is called beamstrahlung.

A characteristic angle for the "full energy" primary, disrupted electrons is [Hollebeek 1981]:

$$\theta_d \equiv \frac{2Nr_e}{\gamma(\sigma_x+\sigma_y)} = \frac{D_{x,y}\sigma_{x,y}}{\sigma_z} \quad , \tag{11.139}$$

where $\sigma_{x,y}/\sigma_z$ is sometimes called the diagonal angle and equals θ_d for $D_{x,y}$=1. Typically, in the e^+e^- channel, the maximum disruption angle is $\theta_{x,max}$=θ_d because the disruption parameter in the vertical is so large that one gets overfocusing or a thick lens effect whereas the focusing over the length of the beam in x is weaker but cumulative or more like a thin lens. The rms angles are $<\sigma_{x,y}>= 0.550\theta_d$. For , the situation reverses and the vertical disruption angle dominates as expected.

Beyond θ^*=$\sigma_{x',y'}\ll\theta_d$, another angle of relevance here is the spin precession angle θ_s. This can be expressed in terms of the spin tune

$$\nu_s = \frac{E[GeV]}{0.44065}\cdot\frac{\Theta}{2\pi} \quad , \tag{11.140}$$

where Θ is some deflection angle in radians. For the bends used here, this is typically twice the maximum disruption angle θ_d=256 μr according to Table 11-40. Further, the bends rotate the polarization in a correlated way in the bend plane between the input and output rather than depolarize the beam assuming that radiative effects (spin flip) are small and that the emittance is small and that there are no significant orbit distortions. The effective polarization after such a bend is $P_{\rm eff} = P_{\rm inc}\cos(2\pi\nu_s)$=0.42–0.84.

There are still several depolarizing effects to be considered. The rms disruption angles cause a net depolarization. A previous section gives another due to an rms gaussian energy spread δ. Clearly, an energy spread will give rise to a spread in precession frequencies that can cause some depolarization but this appears small for the disruption angles and bend angles used here which is about 0.1 % for θ_d=256 μr and <2 % for a 1 mr of bend even at 500 GeV. This classical effect can be avoided insofar as the beam polarization measurement is concerned because the beam is dispersed with sufficient resolution to emphasize the undisrupted portion of the beam. A worse effect, pointed out by Yokoya and Chen, relates to spin-flip during the beamstrahlung process. This can be related to n_γ in Table 11-40. and was one constraint on our bend angles of 1 mr or so but this is not an intrinsic limitation and could prove useful to monitor beam overlap.

The final angle of relevance is the crossing angle θ_c at the IR. For multibunch trains, we may need to introduce a crossing angle and design the FF quadrupoles accordingly. This can decrease the collision efficiency ζ so that we have to introduce variable, crab-crossing cavities [Palmer 1990] that rotate the beams to the appropriate orientation at the IP or CP (the $e\gamma$ conversion point) to restore ζ. We have shown that such cavities will be required for all incident channels [Spencer 1995c] so we will assume that they exist.

Optical Considerations on Energy Recovery

For energy recovery, we want the purely transverse dispersive as well as the **mixed** transverse-longitudinal terms in the optics to be zero to all orders in $\delta \equiv (E-E_o)/E_o$. However, for energy recovery, we want to add another constraint that would not be important otherwise—the synchronicity or time (phase) variation with energy that makes different energy electrons arrive with the correct but differing phases required to lose (or gain) the correct energy. Until otherwise noted, this implies that a high-order achromat is desired (that is not otherwise required), with the corollary condition that we want a first-order or linear R_{56} term that we would like to make variable while otherwise maintaining the high-order achromatic behavior.

In practice, it is the energy-angle correlations as well as the purely angular and chromatic terms that are of most importance because of the small spots and the strong disruption at the IP. The most relevant terms without bending magnets, in standard notation, are:

Second Order:	$(1\|26), (3\|46)$	$(1\|16), (3\|36)$
Third Order:	$(1\|266), (3\|466)$	$(1\|166), (3\|366)$

where the second set of terms in each row is less important. The $(1|26)$ and $(3|46)$ terms are dominant but would not require correction if we only wanted to take the beam to the dump. In fact, they would help the loading on the dump window. The bunch length terms $(5|22)$ and $(5|44)$ enter at the level of 0.1 μm or 0.1 % of σ_z. If we add bending magnets, the only additional terms of importance are $(1|66)$ and $(1|666)$.

Although not essential for most of the other applications, reducing such terms improves things—especially the spectrometer resolving power that is important for the Compton measurements as well as measuring the energy distribution of the outgoing disrupted beam which in turn measures the overlap and distribution of the incoming beams. We should note that it appears more practical to use the 10 MW of power than to recover it to the grid.

Configurations for Beam Monitoring and Experiments

We have already discussed the point-to-parallel configuration for the primary beam. Of course, the undisrupted portion of the beam goes to the dump anyway but this has several different single-beam applications with differing focal points along the way to the dump. Some of these possibilities are discussed below and shown schematically in Figure 11-66.

11.8.5 Beam Control Hardware

The initial quadrupoles are assumed to be permanent magnets (PMs) made from SmCo and NdFe. We have actually made and used such multipoles at SLAC since 1982. Prototypes for the SLC final-focus quadrupoles were made and tested under a variety of conditions such as very strong transverse fields [delCorral 1982]. Figure 11-67 shows several PM systems including the SLC prototype, an SLC split-ring sextupole and an NLC prototype quadrupole with

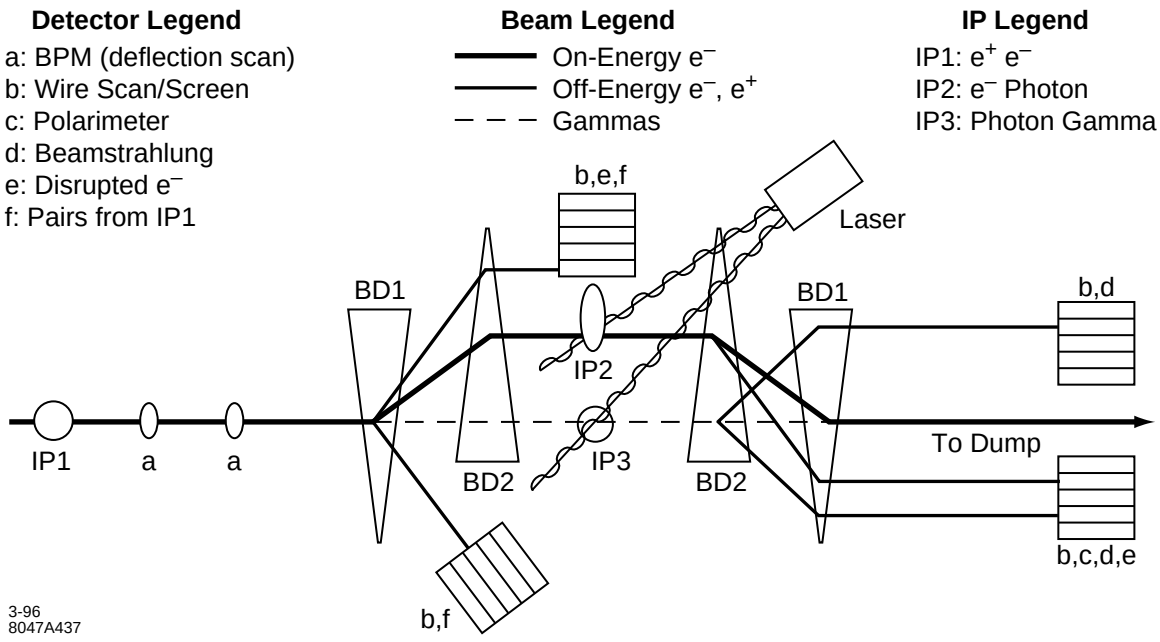


Figure 11-66. *NLC extraction line and diagnostic layout.*

a gradient G>500 T/m that was made in 1989. More recently our experience with the permanent magnet sextupoles used in the SLC damping rings since 1985 was reported [Gross 1994].

An important characteristic demonstrated by these sextupoles, applicable to any PM multipole that has symmetry about the median plane, is that there does not have to be any material in the median plane of the magnet. Thus, the large number of low-energy pairs that spread out from the solenoid axis to the position of the primary outgoing beam can be passed without necessarily intercepting any material. While it appears that conventional PMs can be used in the IR, some sample calculations using a new technique by which magnets with open structures can be designed and still achieve very strong fields also appear to be possible. Additionally, some of the dipoles are combined function magnets (dipoles/quadrupoles) to simplify the transmission of the larger energy spreads and emittances in the outgoing beam from the IP.

Finally, to avoid the nonlinearities in the end-field of the detector solenoid, we propose using a special wiggler in that region that is driven by the strong solenoid field itself. This element serves as a different kind of mirror plate for the detector field that shields the beam from the nonlinear transverse and longitudinal fields that occur in the end region by rerouting and reshaping the flux there. Ultimately that flux must be returned to the detector's conventional endcap by topologically distorting the wiggler's exterior mirror plate. Currently, a superconducting quadrupole (300 T/m) occupies this region and accomplishes a similar result.

11.8.6 Beam Monitors and Diagnostics

One of our primary goals is to provide diagnostics to help put the beams into collision, to maintain this state, and to measure and optimize luminosity. The latter involves diagnostics to characterize the beams: polarization, energy,

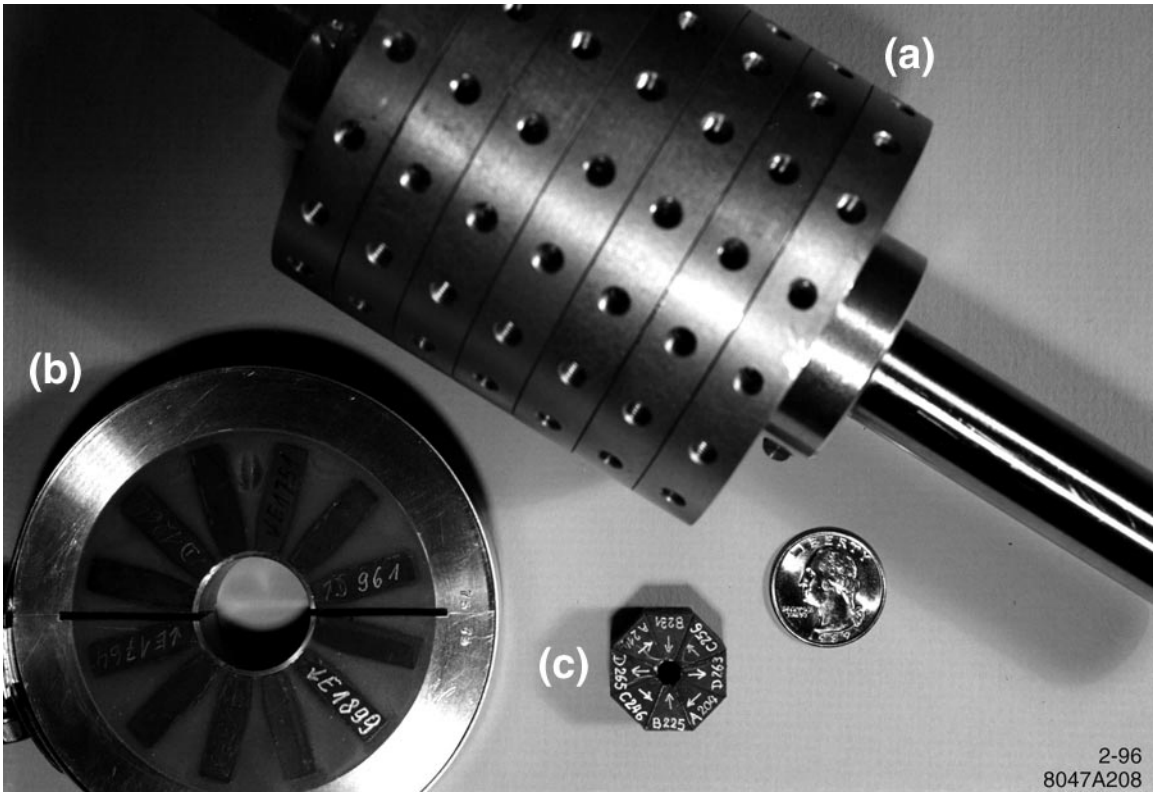


Figure 11-67. *Representative permanent magnet multipoles: (a) a 7-layer quadrupole for the SLC, (b) an SLC damping ring sextupole, and (c) a 500 T/m NLC quadrupole prototype.*

intensity, profile, positions (both absolute and relative), and the disruption of such quantities from the beam-beam and environmental factors.

We also need to dump the beams cleanly and safely. This involves the possibility of other functions such as secondary beams: e, μ, n and γ as well as other forms of energy recovery before final disposal.

The guiding principle in the instrument layout shown in Figure 11-66 was to minimize the material in the high-power beams. Thus, there is a significant use of lasers to control the production of additional particles. Rf BPMs for both position, angle, and timing information are assumed to begin near 5 m where the outgoing beams are still small and C-band cavities [Hartman priv] can have apertures comparable to the quadrupoles ($\approx \lambda/4$). Nevertheless, since beamstrahlung is unavoidable, there are possibilities [Norem 1995] to use either Compton or beamstrahlung photons that could prove useful for monitoring the position, size, and correlations of the bunches at the IP on a bunch-to-bunch basis. Clearly, beamstrahlung is quite sensitive to any changes in these parameters at the IP. In fact, the photon distribution is a better measure of the bunch profile at the IP than the outgoing, disrupted electrons as shown by Figures 11-55 to 11-60. We also assume BPMs, wire scanners, and screens similar to those used for the SLC [Field NIM].

The usual monitors such as BPMs and wire scanners are deployed at key locations for stabilization and feedback for position and energy control. It is understood that the BPMs include both amplitude and phase sensing. There are also more arcane monitors that look at the beamstrahlung from the IP as well as Compton backscattering that have

nanometer resolutions [Norem 1995] and can function on a bunch-to-bunch basis during colliding beam operation or with single beams.

Luminosity Measurements

To estimate the difficulty of this measurement, one can begin by looking at what the Bhabha rate might be for an "annular" detector subtending some angular range $\delta\theta$ near the front face of the first quadrupole in the dump line after the IP. Because this quadrupole (QDF1) is located at l^*=2 m with a radius of 7.5 mm, a reasonable minimum angle θ_1 occurs in the range $\theta_1 \approx$1–3 mr. If the predicted rate is reasonable for a bunch or even a train, we can then consider the backgrounds predicted in the beam-beam simulations to get an estimate of the signal-to-noise ratio. One can then consider the necessary characteristics required for the detector based on the characteristics of the backgrounds.

For unpolarized Bhabha scattering, a good approximation for the angles of interest here (a few hundred mr or less) is

$$\sigma(\delta\theta) = \int_0^{2\pi}\int_{\theta_1}^{\theta_2} \frac{d\sigma}{d\Omega}(e^+e^- \to e^+e^-) \approx \frac{4\pi\alpha^2}{s}\left[\frac{1}{\sin^2(\frac{\theta}{2})} + 4\ln\left(\sin\frac{\theta}{2}\right)\right]_{\theta_2}^{\theta_1} \quad . \tag{11.141}$$

Using $4\pi\alpha^2 = 261\,\mathrm{GeV}^2\,\mathrm{nb}$ with 250-GeV beams at very forward angles we have

$$\sigma(\delta\theta)|_{\theta_1}^{\theta_2} \approx 1.05\left(\frac{2}{\theta_1\theta_2}\right)^2(\theta_2-\theta_1)(\theta_2+\theta_1)\ \mathrm{pb} = 4.2\frac{1}{\theta_1(mr)}\frac{1}{\theta_2(mr)}\left[\frac{(\theta_2-\theta_1)(\theta_2+\theta_1)}{\theta_1\theta_2}\right]\mu\mathrm{b} \quad . \tag{11.142}$$

Assuming $\mathcal{L} = 6\cdot 10^{33}$ at 250 GeV for the 500-A column in Table 11-40 implies a Bhabha rate

$$R_B^{500A} = \mathcal{L}\sigma(\delta\theta)\Big|_{\theta_1=1mr}^{\theta_2=3mr} \approx 22,400/\mathrm{s} = 125/\mathrm{train} = 1.4/bunch \tag{11.143}$$

for a rep-rate of 180 pps in the linac with 90 bunches/train. Increasing θ_2 to 100 mr increases this rate by less than 15%. Again, using the column for 1000 A in the Table for 500-GeV beams, the corresponding numbers are

$$R_B^{1000A} = \mathcal{L}\sigma(\delta\theta)\Big|_{\theta_1=1mr}^{\theta_2=3mr} \approx 12,100/\mathrm{s} = 101/\mathrm{train} = 1.1/\mathrm{bunch} \quad . \tag{11.144}$$

Increasing θ_2 to 100 mr increases the rate to 114/train or 1.3/bunch. θ_2 was made small to be compatible with an rf shield that would extend from the last quadrupole toward the IP with a radius comparable to that of the quadrupole. While such a shield is not as important for the outgoing beam lines, there are no serious penalties imposed on the luminosity monitor for having one.

The value of θ_1=1 mr is reasonable from Figures 11-55, 11-58, and 11-59 for the typical beams of outgoing photons and electrons that we expect. While these two beams have comparably high intensities and are not expected to limit the measurement, the pairs are clearly a problem even though we expect on the order of 10^4/bunch crossing or 10^6 less than for the incident beams (or photons) per crossing. This is demonstrated in Figure 11-62 for the 500-A configuration. However, because these particles are predominantly below 1 GeV, some form of energy discrimination should be sufficient to deal with them.

However, because the pairs that are relevant here are in the higher-energy part of the pair spectrum, two important points can now be made. First, a sampling measurement of these particular pairs is itself a measure of the luminosity and is therefore a useful adjunct with a distinctly different signature. Further, this pair rate, in this location, is a problem that has to be dealt with anyway. Second, we can avoid this problem, at least for the Bhabha luminosity measurement, by going further downstream. This is clearly demonstrated in Figure 11-63 that shows the different beams at 5-m from the IP. However, from Figures 11-59 and 11-60 showing the x and y profiles for both the photons and electrons, this

is not the preferred location. Clearly, the optimal location is somewhere in between but a location near 2 m appears acceptable for making both measurements—especially because this allows the possibility of varying the solid angle subtended by the detector via tromboning or moving it longitudinally toward the IP. We expect almost 100 pairs within the quad aperture there.

Because halving the angle θ_1 quadruples the rate, the main questions are then how small we can make this angle, and whether it is practical to move the detector longitudinally to vary the subtended solid angle for different operating conditions such as beam scans. There are a number of interesting questions concerning the detector and its energy and timing discrimination for individual bunches.

Polarization Measurements

We propose to measure the longitudinal electron beam polarization (P_e) with a Compton polarimeter as shown in Figure 11-56 similar to how it is done for the SLC [Woods 1994]. The polarimeter will detect Compton-scattered electrons from the collision of the longitudinally polarized electron beam with a circularly polarized photon beam. The detector will be a segmented threshold Cherenkov counter similar to that in use at the SLC.

The counting rates in each Cherenkov channel are measured for parallel and anti-parallel combinations of the photon and electron beam helicities. The asymmetry formed from these rates is given by

$$A(E) = \frac{R(\rightarrow\rightarrow) - R(\rightarrow\leftarrow)}{R(\rightarrow\rightarrow) + R(\rightarrow\leftarrow)} = P_e P_\gamma A_C(E) \tag{11.145}$$

where P_γ is the circular polarization of the laser beam at the CIP at IP2 in Figure 11-66 and $A_C(E)$ is the Compton asymmetry function. $A_C(E)$ and the unpolarized Compton cross section are shown in Figure 11-68 for an electron beam energy of 250 GeV and a photon energy of 1.17 eV. The linear Compton spectrum is characterized by a kinematic edge at 46 GeV (180°) backscatter in the center of mass frame, and the zero-asymmetry point at 77 GeV (90° scattering in the center of mass frame).

The photon energy of 1.165 eV corresponds to using an Nd:YAG laser operating at 1064 nm. The long wavelength is chosen to avoid background from the reaction $e^-\gamma \rightarrow e^+e^-e^-$ that opens for

$$E_\gamma(\mathrm{eV}) > \frac{522}{E_e(\mathrm{GeV})} \quad . \tag{11.146}$$

However, this channel is mediated by a second photon and lower thresholds are possible when the second photon is real as from a secondary process such as Compton backscattering. Another laser photon can then produce a Breit-Wheeler pair. This cross channel process opens, in lowest order ($\eta \rightarrow 0$), for:

$$E_\gamma(\mathrm{eV}) > \frac{315}{E_e(\mathrm{GeV})} \tag{11.147}$$

and clearly depends on laser intensity and wavelength. Similarly, for beamstrahlung photons, the threshold is related to the Schwinger critical field by the laser's intensity and wavelength for a given incident electron energy.

The kinematics for Compton scattering gives the location of the Compton edge, $E_C(\mathrm{edge})$, and the edge asymmetry, $A_C(\mathrm{edge})$, as

$$\begin{aligned} E_C(\mathrm{edge}) &= E_e \cdot y \\ A_C(\mathrm{edge}) &= \frac{y^2 - 1}{y^2 + 1} \end{aligned} \tag{11.148}$$

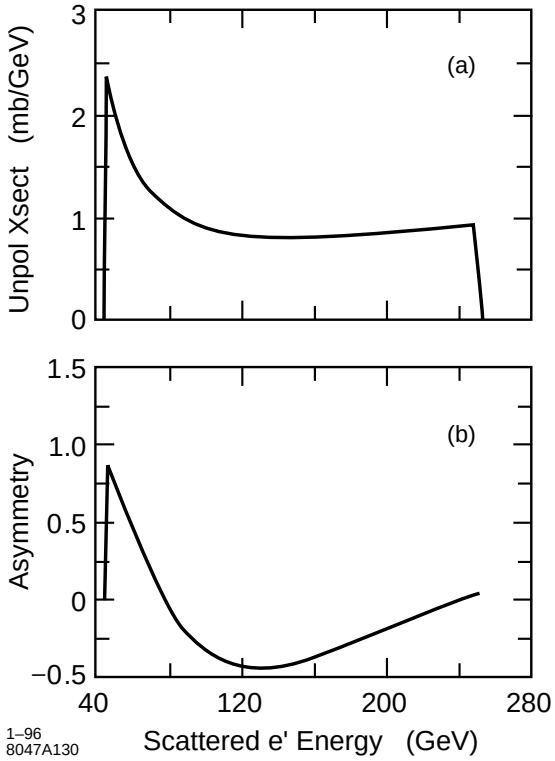


Figure 11-68. *(a) Linear Compton scattering and (b) asymmetry A_c for 250-GeV electrons.*

$$y = \frac{1}{1+\frac{4E_e E_\gamma}{m_e^2}} \tag{11.149}$$

where this is again in the limit that the classical, strong field parameter $\eta \ll 1$. The operant dimensionless, classical, strong field parameter is:

$$\eta = \frac{e\mathcal{E}_{\rm rms}\lambda\!\!\!^{-}_{\gamma}}{mc^2} = \frac{e\sqrt{\langle A_\mu A^\mu\rangle}}{mc^2} \quad . \tag{11.150}$$

This is described elsewhere [Bula 1995] where one sees the smearing effect on the edge from the nonlinear and multiple scattering effects due to high laser intensity. We have calculated this for an NLC example at 250 GeV using YACC, a variant of CAIN. A sample result is shown in Figure 11-69 and discussed further in the section on secondary beams.

Disruption of the electron beam during the collision process can have a significant effect on both the effective polarization of the electron beam during the collision process [Yokoya 1988] and on the measurement of beam polarization. The polarization of the undisrupted incoming beam can easily be determined by measuring the electron polarization when the positrons are absent. The effective polarization of the disrupted electron beam can be determined from Compton asymmetry measurements during collisions with studies of this asymmetry for different targeting of the the laser beam on the electrons at IP2. Corrections for the effects of jitter in the offsets between the beams at the IP can be made using multiple BPM measurements near the IP which is one objective of the spectrometer. Good knowledge of the disrupted electron beam energy spectrum will be required and this can be measured with a wire scanner near IP2.

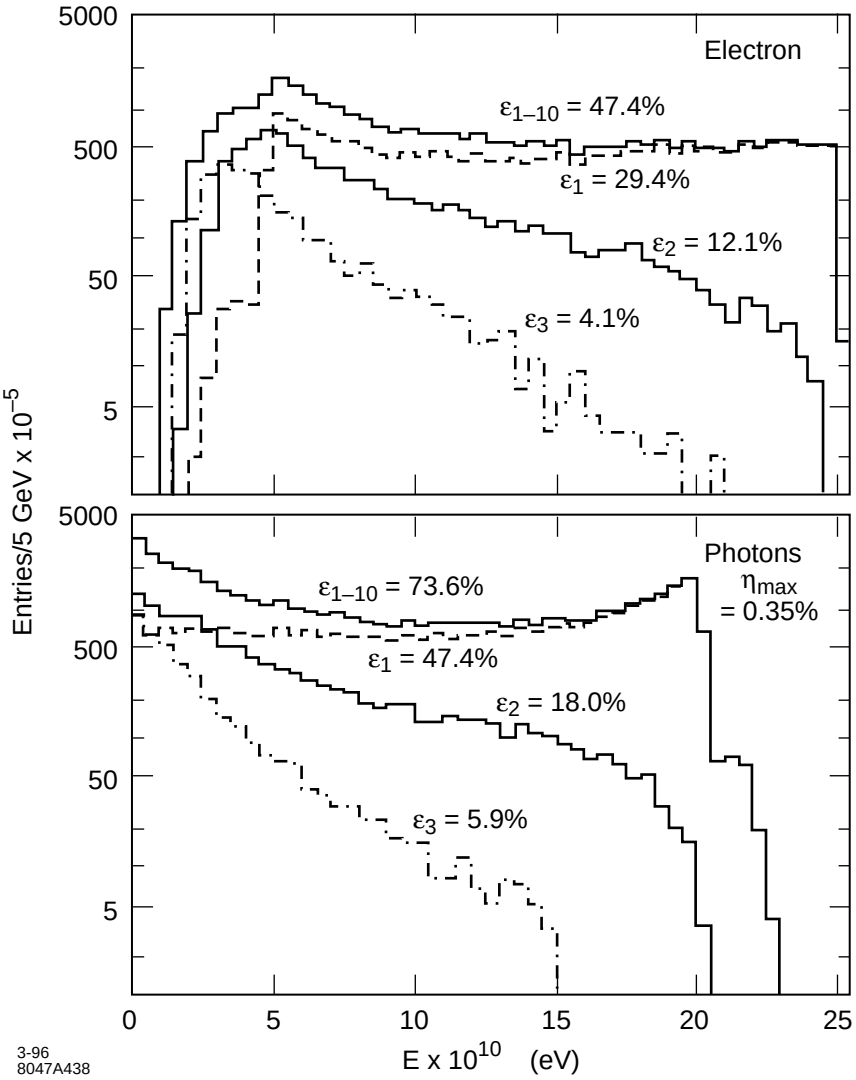


Figure 11-69. *Compton spectra for undisrupted beams at IP2 for high laser intensities.*

Because these electrons are quite low in energy and fall on the far out tail of the disrupted beam distribution, the only requirement is the ability to resolve their energy to within a few GeV, *i.e.,* we would like a resolving power at 500 GeV of:

$$R = \frac{P}{\Delta P} \geq 500.$$

We have provided $R > 2000$ for the full energy beam so there is no problem. We have pointed out before that the nonlinear effects are quite interesting in this regime but this will be discussed elsewhere.

Beam-Beam Deflection Scans

At the SLC, measurement of the deflection of one beam by the opposing beam at the IP, when they are offset, provides an important diagnostic which allows measuring, optimizing, and maintaining luminosity [Bambade 1989]. At the NLC, deflection angle measurements could prove just as valuable. For undisrupted beams, the luminosity varies exponentially with the square of the separation or offset between beams:

$$\frac{\mathcal{L}(\Delta_{x,y})}{\mathcal{L}(0)} = e^{-\Delta^2_{x,y}/2\Sigma^2_{x,y}} \tag{11.151}$$

where $\Delta_{x,y}$ is the transverse offset of the two beams and Σ_y (Σ_x) is the quadrature sum of the rms vertical (horizontal) beam size for the two colliding beams. A good test of our ABEL code is to compare to this prediction for a relevant set of NLC design parameters (Case A) while calculating a deflection scan. The results are shown in Figure 11-70 and are labeled by H_D but are more appropriately termed a collision efficiency factor ζ [Spencer 1995c] especially because the analytic calculations do not include disruption effects.

The deflection angles can be determined from BPM measurements before and after the IP. For flat, upright Gaussian beams which are centered on the x axis, the mean y deflection angle is given by:

$$< \theta_y >= \frac{2r_e N}{\gamma}\Delta_y \int_0^\infty dt \frac{e^{-\Delta^2_y/(t+2\Sigma^2_y)}}{(t+2\Sigma^2_y)^{3/2}(t+2\Sigma^2_x)^{1/2}} \tag{11.152}$$

where r_e is the classical electron radius, N is the number of target particles, and γ is the beam energy divided by the electron mass. Sample deflection scans for the NLC design parameters of Case A corresponding to Figure 11-70 are given in Figure 11-71.

According to Ref. [Raimondi 1995], the effect of beam-beam disruption will cause larger deflection angles than given in the above formula for y deflection angles, but will have a negligible effect on horizontal deflection angles. Our results indicate observable differences in both transverse directions, *e.g.,* 6–7% in the peak deflection angle in x and 9–10% in y. In both cases, the location of the peak deflection is pushed outwards but one sees that the predominant effect of the disruption is to push the location of the y peak outwards beyond 10σ or so for the chosen parameters. In fact, we can summarize the results of a number of calculations as follows: 1) small disruptions ($D_{x,y} \ll 1$) can use the analytic results while large Ds need simulations, 2) for a given set of parameters typical of Case A, the Y deflection angle converges to the same value ($\approx$220 μr) independent of σ_y, 3) the location at which this value is achieved comes at about $10\sigma_y$ in Figure 11-71, 4) this is true for a large range in σ_y, but 5) this value depends on the the value of σ_x assumed.

We can understand these results as due to the significant disruption/focusing of the beam in the collision region. A 200-μr angle over a 200-μm bunch length can give an offset of 40 nm or $\geq$ 10 σ. It seems advisable to test the characterizations made above with real experiments on the SLC where practical.

Because the maximum deflection angle will be comparable to or less than θ_d in Table 11-40, we may conclude that a radially uniform aperture in the outgoing quadrupoles should be acceptable for doing this diagnostic test even though

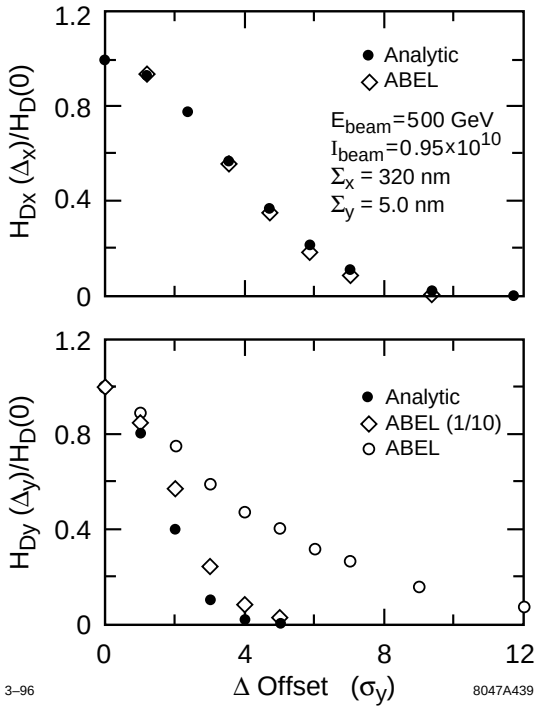


Figure 11-70. *Comparison of beam-beam enhancement factors H_D (or efficiencies ζ).*

the outgoing beamstrahlung now has a pointing angle that can sweep out angular apertures of nearly half a mr. This is not a problem because the apertures were sized to allow such angles. Rather, it provides another reason to place the primary beam dump closer to 100 m rather than the 150–200-m distance that would otherwise be allowed. Pair production is a worse problem for the quadrupole acceptances.

Another signature for disruption could come from measuring the photon moments as well as those of the electrons according to Refs. [Norem 1995, Field NIM]. The drawbacks of the deflection scan are obvious so that other alternatives, involving less beam time overhead and aperture, are being considered. This is one reason for the good energy resolution at IP1 in Figure 11-66. However, regardless of the method of choice, the deflection scan can be expected to provide a well-understood comparison that can be used for the NLC.

11.8.7 Beam Dump

The primary beam dumps have to dispose of essentially all of the power at the highest beam energy (750 GeV). Appendix 11.A argues that water should be the primary absorber in a cylindrical vessel housing a vortex-like flow of water with vortex velocity ≈1–1.5 m/s normal to the beam momentum. The vessel is 1.5-m, diameter and has a 6.5-m-long water section, followed by ≈1 m of water-cooled solids to attenuate a 750-GeV EM cascade shower. The beam enters through a thin window ≈1-mm thick and 20-cm diameter. Production of ≈5 l H_2/16 MW beam power from radiolysis [Walz, 1967] can be mitigated with a catalytic H_2/O_2 recombiner that has a closed-loop system that contains all radioisotopes.

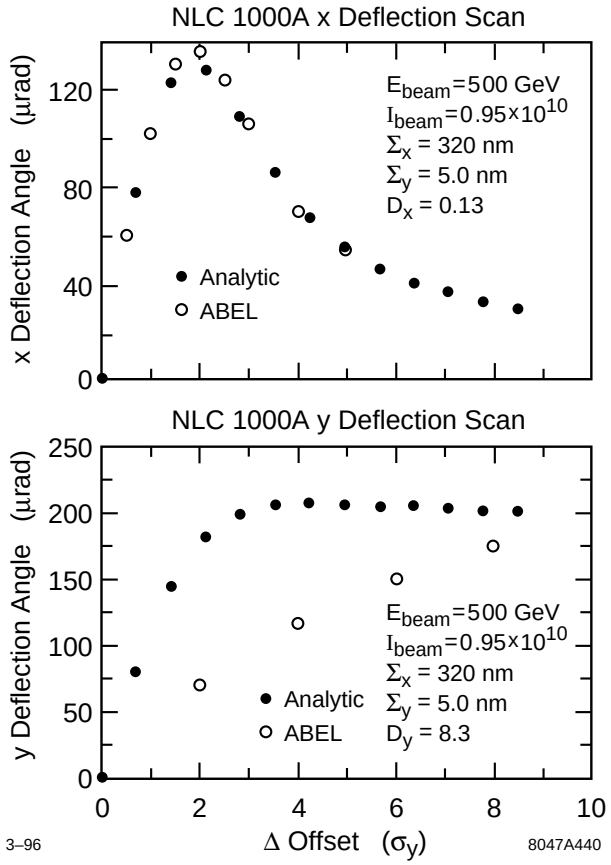


Figure 11-71. *Comparison of beam deflection scans corresponding to Figure 11-70*

The reader is referred to Appendix 11.A for the complete details. Here we only summarize the main characteristics and point out that a window size of 20 cm appears practical from all standpoints. We also note that there is the possibility to do some classical beam dump experiments after the dump which may be facilitated by dispersing the beams horizontally.

11.8.8 Secondary Beams

Beyond the usual possibilities, there are interesting opportunities to produce intense beams of higher mass leptons and hadrons as well as neutrons that haven't really been explored yet but could provide both a unique and comparatively inexpensive alternative source to nuclear reactors for some applications.

Muon Possibilities

Discussions on muon colliders have been resurrected recently so that the associated possibility of muon beams and colliders based on very high-energy, high-quality electron beams naturally arises. In fact, this is almost unavoidable

when one considers the backgrounds from muons that are also unavoidable and other questions of energy recycling and beam disposal.

Most discussions about muons use proton beams to produce the muons but Barletta and Sessler have looked at electroproduction recently. They concluded it was not viable but also did not consider the kinds of beams we will assume here. An important aspect is that we can use all the cooling tricks that have been discussed to cool the very large emittances from proton production although our assumption is that we can make them in a way that avoids many of the steps (and problems) normally required because we start at higher energies where there are several intrinsic advantages. The procedure is to produce high-energy photons as in the preceding section which are then used to produce the muon beams. Photoproduction was proposed previously for producing high-energy polarized positrons and there are a number of variants including the use of crystals and high-power lasers or FELs.

Photon Possibilities

This possibility relates to the second interaction region (IP2) and to the nonlinear QED experiment on the FFTB line at SLAC (E144). Clearly, similar problems are encountered with the Compton polarimeter and profile monitors based on photons. Related questions concern beam stability and synchronization so that much of the hardware has a number of interrelated uses. In Figure 11-69, we gave a worst-case scenario for IP2 assuming we were limited in both laser wavelength and power to the specified quantities. We point out that ignoring nonlinearities in the laser-electron interaction so that we can use unlimited laser power if we could get it, we could conceive of higher luminosity in this channel than currently appears slated for IP1. Table 11-41 gives some characteristic numbers. Additional results can be found in the section on the second IP and elsewhere [Spencer 1995c].

11.8.9 Energy Recovery and Its Applications

Following the introductory discussion, we note that it costs a significant amount to build and continuously dispose of such high-power beams in an environmentally acceptable way. Depending on the options that are pursued, there are at least four beam dumps required—not counting possible secondary lines. If we suppose that it takes another 4% of the total beam power just to dispose of the beams, this is another $1,000/day in power costs per dump. Clearly, this is worth pursuing if, in fact, one can find some nondisruptive means that are themselves environmentally acceptable. This question has been raised in the past at SLAC but the power involved and its subsequent uses implied it was not economically feasible.

Some ways to look at rf recovery methods are to remember how klystrons produce rf power from an electron beam, how recirculating systems work, and how high-frequency colliders such as CLIC propose to produce its rf power. If we want to store it or use it in some associated storage rings then we should consider superconducting systems. Any subsequent physics application should justify the cost. Ideally, whatever the sequence, one would end up with a final, dumped beam that has been reduced to an energy below the neutron production threshold to minimize activation of long-lived isotopes.

One obvious synergistic application is to try to drive an FEL with this beam in such a way that the same FEL can be used to produce variable, high-energy photon beams. It is interesting to point out that the Stanford recirculating SCA proposed and tested energy recovery to improve FEL efficiency [Rohatgi 1987]. One can expect such methods to provide up to 25 MV/m.

c.m. Energy (TeV)	0.5	1.0	1.5	TeV
Bunch Current, N_{max}	0.70	1.45	1.45	$\times 10^{10}$
Bunches/Train	100	75	75	
Bunch Separation	42	42	42	cm
Trains/s	180	120	120	s^{-1}
Coupling, $K \equiv \epsilon_y/\epsilon_x$	1.0	1.0	1.0	%
IR Beta, β_x^*/β_y^*	10/0.1	10/0.1	30/0.3	mm
IR Size, σ_x^*/σ_y^*	316/3.16	224/2.24	316/3.16	nm
Sizes at δz=4 cm	1.31/1.27	0.93/0.89	0.53/0.42	μm
Sizes at δz=8 cm	2.56/2.53	1.81/1.79	0.90/0.84	μm
Sizes at δz=12 cm	3.82/3.80	2.70/2.68	1.30/1.26	μm
Sizes at δz=20 cm	6.53/6.33	4.48/4.47	2.13/2.10	μm
Divergences, $\sigma_{x'}^*/\sigma_{y'}^*$	31.6/31.6	22.4/22.4	10.5/10.5	μrad
Laser $\lambda(\mu)$	1.05	2.1	3.2	μm
$< \delta t >_L$	0.1	0.1	0.1	ps
$< \delta l >_L$	0.3	0.3	0.3	ps
σ_C[b]	0.197	0.197	0.197	$\times 10^{-24}\mathrm{cm}^2$
Photons/Pulse, N_λ	4.2	4.7	2.9	10^{18}

Table 11-41. *Parameters for 0.5, 1 and 1.5 TeV. Electron emittances are $\gamma\epsilon_x$=5 $\times$ 10^{-6} m and $\gamma\epsilon_y$=5 $\times$ 10^{-8} m. The energy spread is σ_E/E=0.20 % and the bunch length is σ_z=100μ.*

11.8.10 Other Questions and Problems

There is a question concerning the sufficient separation between the various beams. This is especially relevant for the primary e and γ beams that are to be monitored and that have comparable densities and power. Another question is the e^- depolarization that we have not thought much about except for measuring it. Finally, there are all the alternatives for monitoring the luminosity, beam sizes, and offsets at the IP under normal operating conditions. We have proposed several methods in each case, and that work is still in progress. We have shown that there are significant differences between the SLC and NLC and that the NLC has more possibilities than the SLC because of the greater disruption and its secondary effects. We have also verified that virtually all of the corresponding diagnostic tools that are used for the SLC such as conventional deflection scans can be accommodated by the extraction optics and diagnostic devices even in the presence of pairs. However, this subject is a fertile area for exploration and it is being pursued together with that of strong PMs having open structures that are matched to the outgoing beam distributions and characteristics under differing diagnostic conditions. Finally, we strongly emphasize the importance of testing as many of the various calculations that have been presented here as practically possible. This could benefit the SLC as well as any future generation machine.

11.9 Conclusions and Comments

The NLC final-focus design fulfills all specifications, and achieves the desired IP spot sizes of about 300 nm$\times$4 nm, in the entire c.m. energy range between 350 GeV and 1.5 TeV, while maintaining a reasonable momentum bandwidth of $\pm 0.60\%$ or larger. The proposed design lends itself to an easy and straightforward upgrade to 1.5-TeV-c.m. energy, requiring only minor changes in geometry. The price to be paid for this flexibility is a final-focus length of about 2 km.

All possible sources of spot-size dilution, which are relevant on different timescales, have been budgeted, and are controlled by stabilization, tuning and maintenance systems. These systems also ease the tolerances on element vibrations, position drifts, and field stability. Most of the tolerances are not particularly tight, and have already been achieved at the FFTB or at other places. It is interesting that the luminosity loss caused by ground motion appears to be insignificant and that, therefore, the ground (bedrock) provides an ideal reference for magnet alignment and stabilization. In order not to destroy the coherence of the ground, care will be needed in the mechanical design of components and magnet supports. The primary engineering task will be to contain the effect of cultural noise.

The design of the NLC final focus profited greatly from experience with SLC and FFTB operation, whose influence is clearly exhibited by optical layout, aberration and tolerance budgets, tuning schemes, sensitivity studies, diagnostics, and operational procedures.

In conclusion, the NLC beam delivery system will not only produce small spot sizes at the IP, but it also promises redundant tunability, adjustability over a wide energy range, and, last not least, tolerable detector background.

11.A The Beam Dumps

If the efforts to economically recover a significant fraction of the spent electron beam energy are not successful, we will be faced with the tasks of safely disposing and dissipating these 10-MW beams. The working beam parameters are either:

90 Bunches/Train($\equiv n_B$) at $N_B \leq 0.85 \cdot 10^{10} \equiv 76.5 \cdot 10^{10}$/Train and $f_{\rm rep}$=180 Hz

or

90 Bunches/Train($\equiv n_B$) at $N_B \leq 1.25 \cdot 10^{10} \equiv 112.5 \cdot 10^{10}$/Train and $f_{\rm rep}$=120 Hz.

Then the beam energies and average power per beam are:

$E_\circ$ (GeV)	40	80	150	250	500	750
$P_{av,90/180}$ (MW)	0.88	1.76	3.30	5.51	11.02	16.52
$P_{av,90/120}$ (MW)	0.86	1.73	3.24	5.40	10.80	16.20

To assess the magnitude of the task, let us examine some engineering materials commonly used at accelerators. Specifically, we will look at the energy deposited in such materials at the beginning of an electromagnetic cascade and also in the region of peak energy deposition at shower maximum. At the beginning of the cascade (before any shower multiplicity), the average energy loss is approximately given by the ionization loss and can be written as

$$P'_{\rm av} \sim 1.6 \cdot 10^{-19} (-\rho \frac{dE}{dx}) f_{\rm rep} N_B n_B \tag{11.153}$$

and at shower maximum

$$P'_{\rm av,max} \sim \Pi_{\rm max} P'_{\rm av} \tag{11.154}$$

where $P'_{\rm av}$ and $P'_{{\rm av},max}$ are the average power depositions per unit length at the beginning and at shower maximum. ρ is the specific gravity of the material, dE/dx is the minimum ionization loss, and the other quantities are defined in Table 11-40.

For electrons, the location of the shower maximum $T^{e^-}_{\rm max}$, in units of radiation lengths $X_\circ$, and the maximum shower multiplicity $\Pi_{\rm max}$ are given by Rossi [Rossi 1952] as

$$T^{e^-}_{\rm max} = 1.01[\ln \frac{E}{\epsilon_\circ} - 1] \tag{11.155}$$

and

$$\Pi^{e^-}_{\rm max} = 0.31(\frac{E}{\epsilon_\circ})[\ln \frac{E}{\epsilon_\circ} - 0.37]^{-\frac{1}{2}} \tag{11.156}$$

where the critical energy $\epsilon_\circ$ of an absorber material is the unit of measure that makes these expressions independent of material. Typical values for carbon and lead are 76 and 7.6 MeV when one includes the effect of density variation, otherwise $\epsilon_\circ$ goes inversely with atomic number. The following tables give various quantities derived from the equations given above to judge the suitability of various materials for this application.

Table 11-42 gives the location of $T_{\rm max}$ as a function of beam energy. Table 11-43 gives the shower multiplicity $\Pi_{\rm max}$ at $T_{\rm max}$ as a function of beam energy. Table 11-44 gives the average power deposition per unit length at the beginning of the cascade (Π=1) for the two bunch train configurations (to first approximation energy independent). Table 11-45

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_0 =40	4.9	5.3	5.4	6.0	6.5	6.6	6.7	7.6	7.7
80	5.6	6.0	6.1	6.7	7.2	7.3	7.4	8.3	8.4
150	6.3	6.6	6.7	7.3	7.8	8.0	8.1	8.9	9.0
250	6.8	7.1	7.2	7.8	8.3	8.5	8.6	9.4	9.5
500	7.5	7.8	7.9	8.5	9.0	9.2	9.3	10.1	10.2
750	7.9	8.2	8.3	8.9	9.4	9.6	9.7	10.5	10.6

Table 11-42. *Location of $T^{e^-}_{max}$ for various materials (in units of X_0).*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_0 =40	48	65	74	121	195	224	244	537	584
80	90	123	132	231	371	428	467	1030	1122
150	162	220	237	415	670	773	843	1867	2034
250	260	350	382	670	1084	1252	1365	3031	3302
500	497	678	732	1287	2087	2413	2631	5793	6390
750	727	993	1072	1889	3067	3546	3868	8630	9407

Table 11-43. *The shower multiplicity $\Pi^{e^-}_{max}$ at $T^{e^-}_{max}$ for various materials.*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
$P_{av-90/180}$	68	80	44	98	158	255	282	500	284
$P_{av-90/120}$	67	79	43	96	154	250	277	490	279

Table 11-44. *Power deposition per unit length P'_{av} (W/cm) for the two bunch train configurations and various materials before shower development.*

give the same values for the two bunch train configurations at T_{max} as a function of beam energy. Note that the values in the two tables are very similar to each other.

It should be pointed out that calculations using the Monte Carlo code EGS gave power deposition values P'_{max} which are about 25–50% below those given in the tables using the equation for the shower maximum, whereas at the beginning of the shower, they are in very close agreement with the results obtained using P'_{av}.

As is readily apparent from the values presented in Table 11-45, removal of the vast amounts of thermal energy generated in solid materials in the region of shower maximum looks nearly impossible for the higher energy cases and the higher Z materials. Therefore, we will concentrate on examining the pros and cons of using water as the principal power absorption and dissipation medium in the beam dump.

11.A.1 The Beam Dump Vessel

For pulse- or bunch-train repetition rates of 180 Hz and 120 Hz, the train spacings are 5.6 ms and 8.3 ms, respectively. For a bunch spacing of 1.4 ns, the length of one train is 125 ns for 90 bunches and 105 ns for 75 bunches. Since energy

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	3.3	5.2	3.3	11.9	30.8	57.1	68.8	268	166
80	6.1	9.8	5.8	22.6	58.6	109.0	132.0	515	319
150	11.0	17.6	10.4	40.7	106	197.0	238.0	933	578
250	17.7	28.0	16.8	65.7	171	319.0	385.0	1516	938
500	33.8	54.2	32.2	126	330	615.0	742.0	2896	1815
750	49.4	79.4	47.2	185	485	904.0	1090	4315	2672

Table 11-45. *A. Maximum power deposition per unit length $P'_{\mathrm{av,max}}$ (kW/cm) for various materials as a function of beam energy with 90 bunches/180 Hz/0.85·$10^{10}e$.*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	3.2	5.1	3.2	11.6	30.0	57.1	67.6	263	163
80	6.0	9.7	5.7	22.1	57.1	107.0	129.0	505	313
150	10.9	17.4	10.2	39.8	103.0	193.0	234.0	915	567
250	17.4	27.7	16.4	64.3	167.0	313.0	378.0	1485	921
500	33.3	53.6	31.5	124.0	321.0	603.0	729.0	2839	1783
750	48.7	78.4	46.1	181.0	472.0	886.0	1071.0	4229	2625

Table 11-45. *B. Maximum power deposition per unit length $P'_{\mathrm{av,max}}$ (kW/cm) for various materials as a function of beam energy with 90 bunches/120 Hz/1.25·$10^{10}e$.*

is principally lost by ionization from the primary electrons and secondary electrons or delta rays which are highly relativistic, the local microvolume exposed to these rays will have reached final temperature in the time frame of about 10^{-15} s after an individual bunch has passed. This is some three orders of magnitude less than the bunch width, some six orders of magnitude less than the bunch spacing, and some eight orders less than the length of one bunch train. Therefore, in the time frames of the bunch and the bunch train final temperature is reached instantaneously. On the other hand, significant thermal relaxation takes place on a timescale which is long compared to the approximately 100-ns length of a bunch train (and, incidentally, the pressure or shock waves generated by the sudden thermal expansion of the beam-heated microvolume will propagate outward at the velocity of sound in the local medium [∼1450 m/s], which is also on a timescale very long compared to the bunch spacing and long compared to the length of the bunch train). So, the pertinent question to ask at the onset of this examination is "What is the expected temperature rise due to one bunch train at $\mathrm{T_{max}}$?".

From experience and from simulations with the Monte Carlo code EGS, we can expect a beam size at $\mathrm{T_{max}}$ of $2\sigma_b$ ∼0.8 cm for an input size of $2\sigma_b$ ∼1 mm and the higher beam energies under consideration. Let us define a heat source term for the region of 0≤r≤σ at $\mathrm{T_{max}}$ and E_o=500 GeV

$$\mathrm{S} \stackrel{\mathrm{def}}{=} \frac{P'_{\mathrm{av,max}}\mathrm{C}}{(2\sigma_b)^2\pi/4} \tag{11.157}$$

where $C = 0.393$ is a constant to reflect the assumed transverse Gaussian beam intensity distribution. Numerically, we find

$$\mathrm{S} = \frac{32.2 \cdot 0.393}{0.8\pi/4} \sim 25.2 \ \frac{\mathrm{kW}}{\mathrm{cm}^3} \tag{11.158}$$

for the 90-bunch/180-Hz case. Then the temperature rise per bunch train is

$$\Delta T_T = \frac{\mathrm{S}}{\rho c_P f_{\mathrm{rep}}} \tag{11.159}$$

where ρc_P=4.22 J/cm^3°C is the specific heat capacity for water. This is approximately constant over the temperature interval of interest (it only decreases ∼1.5 %). We then find $\Delta T_T \approx$33°C for the 90-bunch train/180-Hz case at 500 GeV and ≈50° C for 750 GeV.

For comparison, the maximum water inlet temperature in mid-summer at the SLC is ≈30° C (≡ 90° F). So, peak bunch train temperatures at T_{max} would be ≈65 and 80°C. We should recall that the heat source term was defined as an average for the body of rotation bounded by $2\sigma_b$, and therefore somewhat higher values can be expected for a Gaussian distribution near r=0. On the other hand, EGS simulations with a Gaussian input often show a double-humped distribution at a depth of several radiation lengths with a small depression at r=0 resulting from the transverse spread of the electromagnetic cascade. Averaging over 2σ is a reasonable approximation. As a matter of fact, peak temperatures which are somewhat higher than the ones calculated at T_{max} are consistently found at a depth short of T_{max}. The reason for this is that in the early part of the electromagnetic cascade, the increase in shower multiplicity dominates over the increase in the transverse size of the beam on account of the higher average energy of the shower particles and consequential smaller scattering angles. At a depth short of T_{max}, this role is reversed and it is at that depth where peak temperatures are found. The discrepancy appears to be larger in low-Z materials.

The next question which needs to be asked is, "What is the significance of a temperature of 80° C, for example, in the beam-heated microvolume of water?" We can make a categorical statement that volume boiling of water needs to be avoided at all cost. The consequences would be a great reduction in the density of the power absorption medium and therefore a downbeam shift of T_{max}, perhaps no longer fully containing the longitudinal cascade shower within the dump itself. The boiling point of water at, for example, 7 atmospheres (≡100 psig) is ≈160° C. So, assuming for the moment that there is no thermal relaxation during the interbunch train period of 6–8 ms, one could accept a second bunch train and not reach the boiling point at the local pressure; with thermal relaxation, it would be at least three trains. But for continuous operation at 120–180 Hz, we must move the water transverse to the beam momentum vector, fast enough, with some element of conservatism, that no two successive bunch trains target on the same microvolume of water. Suppose we use a water velocity of ∼1 m/s, then the distance traveled by the lump of water during the interbunch train period is $V_w \geq 10^3$ mm/s$\times 5.6 \cdot 10^{-3}$s≈ 5.5 mm. This is almost what we need to avoid significant overlap of the microvolume element bounded by 2σ from two successive bunch trains.

The next question which we need to answer is, "What are the longitudinal and transverse dimensions to guarantee complete containment of the electro-magnetic cascade of a 750 GeV/16-MW electron beam in water?" From EGS and analytical work, we find ∼40 $X_\circ$ are needed to sufficiently attenuate the beam longitudinally. This does not address containment of 750-GeV muons. Of this total, ∼15–18 $X_\circ$ should be water. The remainder could be water-cooled solid materials. For water, 1$X_\circ$≡35.7 cm. Thus, the beam would first take a cold bath in ∼5.5–6.5 m of water. An additional 1 m of solids would complete the attenuation of the longitudinal cascade. The total dump length would then be ∼7.5 m.

The dump diameter needs to be large enough to attenuate the radial shower. Based on EGS and experience, $D_\circ$ ∼1.5 m of water should be sufficient. The next question to be answered is, "How much flow rate is required to keep the bulk water temperature rise from exceeding ∼30–35° C?" From simple calorimetry, we find ∼6800 l/min (≡1750 gpm). This should not present any great logistics problems.

Two beam dumps were originally designed and built at SLAC which use a vortex-like flow arrangement to economically dissipate the power from 2.2-MW, 25-GeV electron beams [Walz 1967a]. A schematic of such a dump is shown in Figure 11-72. Water is injected approximately tangential to the dump wall through a series of evenly-spaced holes from an inlet flow header located at the periphery of the cylindrical shell of the dump vessel. The water then flows spirally toward the center of the dump vessel where the exit manifold is located. In the region of interest, the velocity follows the laws of potential flow theory, namely V×r=const. Introducing the electron beam at a radius of ∼0.3 m guarantees a flow velocity transverse to the beam of ∼1 m/s for the flow rate of 2000 l/min (≡550 gpm). After traveling one complete revolution in the vortex, the heated microvolume of water has undergone complete mixing and is returning at a lower orbit. No danger exists for volume boiling.

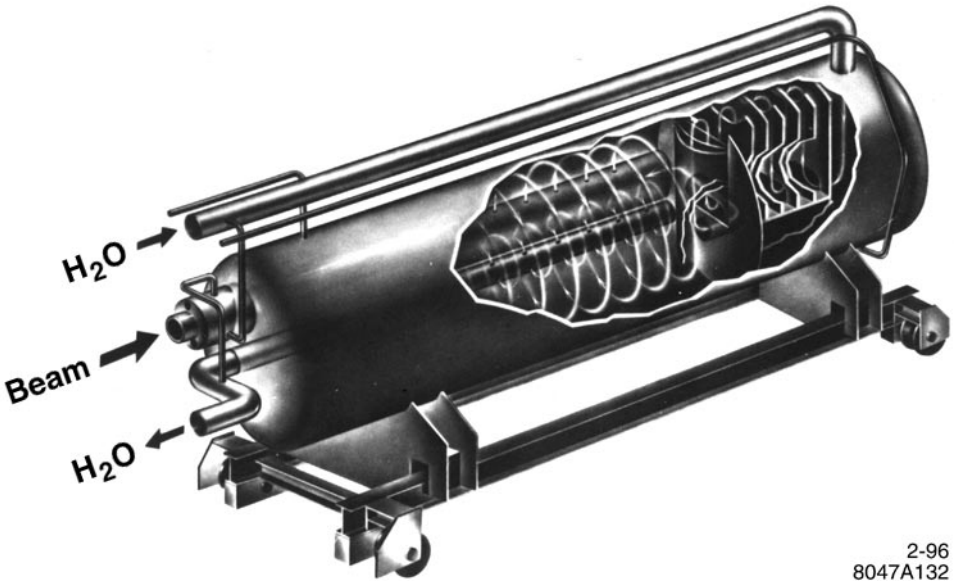


Figure 11-72. *Schematic representation of the primary beam dumps.*

11.A.2 The Window

The electron beam would enter the dump through a thin window separating the water from either vacuum or atmosphere. A number of materials such as aluminum, titanium, or copper appear suitable for this application, the latter being the most attractive. Since beam size and beam excursion from a nominal trajectory are expected to be small for an NLC, the window diameter can be small, say 10–15 cm. A reasonable wall thickness might then be δ=1 mm for a hemispherically-shaped window. The power deposited in the window can then readily be computed by multiplying $P'_{\rm av}$ by the window thickness and using the minimum ionization loss in copper $(-\rho\frac{dE}{dx})$ ~12.8 MeV/cm:

$$\mathrm{P}_w = \mathrm{P}'_{\rm av}\delta = 12.8\cdot 10^6(112.5\cdot 10^{10})120(1.6\cdot 10^{19})0.1 \sim 28\,\mathrm{W} \tag{11.160}$$

for the 90-bunch/120-Hz case. An EGS simulation at 50 GeV gave a value ~3% higher. To first approximation, we can neglect heat losses from the window by thermal radiation (and natural convection if there is atmosphere) to the upbeam world. We assume then that all heat is conducted through the window and removed by forced convection by water on the inside. Let us correct for a double convoluted Gaussian and obtain $\mathrm{P}_{\rm w,corr}$=28·0.393=10.9 W for the region 0≤r≤σ_b. After averaging over $2\sigma_b$ the effective heat flux into the water, neglecting lateral conduction for the moment, becomes q″=$\mathrm{P}_{w,corr}/\mathrm{A}_b$ = 10.86/0.01π/4≡1.38 kW/cm^2. This is in the nucleate boiling heat transfer range and, depending on flow velocity and subcooling, could be close to the transition from nucleate to film boiling. The latter is defined as a burnout condition.

Still neglecting lateral conduction, we find for the maximum temperature difference from the outside to the inside of the window (δ=1-mm thickness and k=3.9 Ws/mm^2°C/cm for copper).

$$\Delta T = \frac{q''\delta}{2}\mathrm{k} = \frac{1.38\cdot 10^3\cdot 0.1}{2\cdot 3.9} \sim 18°\mathrm{C} \quad . \tag{11.161}$$

To get an idea by how much lateral conduction reduces this temperature difference and the heat flux into the water, we use Fourier's Law of conduction

$$k\nabla^2 T + S = \rho c\frac{\partial T}{\partial \tau} \quad . \tag{11.162}$$

For steady state and internal heat generation, this reduces to the familiar Poisson equation

$$\nabla^2 T + \frac{S}{k} = 0 \quad . \tag{11.163}$$

For the axisymmetric case, the Laplacian ∇^2 can be written as

$$\frac{\partial^2 T}{\partial r^2} + \frac{1\partial T}{r\partial r} + \frac{\partial^2 T}{\partial z^2} + \frac{S}{k} = 0 \quad . \tag{11.164}$$

For finite cylinder boundary conditions with 0<z ≤2δ, 0<r≤σ_b and with heat production at a constant rate S per unit volume per unit time, a solution was given by Carlslaw [Carlslaw 1959] as

$$T = \frac{Sz(2\delta - z)}{2k} - \frac{16\delta^2 S}{k\pi^2} \sum_{n=0}^{\infty} \frac{I_\circ[(2n+1)\pi r/2\delta]}{(2n+1)^3 I_\circ\left[(2n+1)\,\frac{\pi\sigma}{2\delta}\right]} \sin\frac{(2n+1)\pi z}{2\delta} \quad . \tag{11.165}$$

The maximum temperature occurs at r=0, z=δ. Using only the first term (n=0) this reduces (assuming an adiabatic interface on the upbeam face of the window) to

$$\Delta T = \frac{s\delta^2}{2k}\left[1 - \frac{1.032}{I_\circ\left(\frac{\pi\sigma}{2\delta}\right)}\right] \tag{11.166}$$

where $I_\circ$ is a modified Bessel function of order zero. S=$0.393 \cdot P'_{\rm av}$=$13.9\,{\rm kW}/cm^3$. Then, we find

$$\Delta T_{\max} = \frac{13.86 \times 10^3 \times 0.1^2}{2 \times 3.9}\left[1 - \frac{1.032}{I_0\left(\frac{\pi \times 0.05}{2 \times 0.1}\right)}\right] \sim 2^\circ{\rm C} \quad . \tag{11.167}$$

Thus, for this case, lateral conduction is of great help and there are no worries about the heat flux into the water approaching the critical heat flux for the transition from nucleate to film boiling, even if we did not average over 2σ_b and instead used the real Gaussian distribution.

The last item to be checked for the window is the temperature rise per one bunch train (since we have already concluded that thermal relaxation is too slow to have any effect on individual bunches). Again, we find S=13.9 kW/ cm^3 and $\rho c|_{\rm Cu}$=3.45 J/ cm$^3{}^\circ$C

$$\Delta T_{{\rm T},W} = \frac{S}{\rho c f_{\rm rep}} = \frac{13.9 \times 10^3}{3.45 \times 120} \sim 33^\circ{\rm C} \quad . \tag{11.168}$$

In a fully restrained body and for E$\alpha|_{Cu}$= 282 psi/°C, we obtain a thermal stress rise

$$\sigma_{th} \approx E\alpha\Delta T = 282 \times 33 \approx 9330\ {\rm psi} \quad . \tag{11.169}$$

This would be excessive for fully annealed copper but should not lead to thermal fatigue in semi-hard or hard copper with fatigue strengths ranging from ~13,000 to 18,000 psi for N>10^8 cycles. Of some concern is the effect of cumulative radiation damage in the copper lattice due to the formation of clusters of vacancies and interstices with resultant embrittlement. For the expected current densities every atom in the lattice will "play musical chairs" and change its lattice site several times daily. It may be prudent to schedule a window change on a regular basis and not wait for a perforation to occur.

11.A.3 Isotope Production

The dominant radioisotopes produced in water by these high-energy electron beams are by photospallation on ^{16}O [Neal 1968]. They are in order of decreasing abundance and increasing half life:

$$
\begin{array}{ll}
^{16}\mathrm{O} \rightarrow {}^{15}\mathrm{O} & \mathrm{T}_{1/2} \sim \ 2 \text{ min} \\
\rightarrow {}^{13}\mathrm{N(small)} & \sim \ 10 \text{ min} \\
\rightarrow {}^{11}\mathrm{C} & \sim \ 20 \text{ min} \\
\rightarrow \ {}^{7}\mathrm{Be} & \sim \ 53 \text{ day} \\
\rightarrow \ {}^{3}\mathrm{H} & \sim 12.3 \text{ year}
\end{array}
$$

For practical purposes, we can neglect ^{15}O since it has mostly disappeared in ∼20 min. The isotope which influences tunnel entry and maintenance work the most is ^{11}C. A delay of three hours will allow most maintenance work. ^{7}Be is primarily removed by filtration in the demineralizer loop (not by ion exchange). Tritium builds up very slowly as $^{3}H_2O$ and can be managed by regularly planned disposal of the water as radioactive waste, albeit at a cost. Clearly, compared to solid power absorption materials (metals), water has significant advantages when viewed in light of residual radioactivity.

11.A.4 Radiolysis and Hydrogen Evolution

The deposition of the electron beam's energy in water causes radiolysis and dissociates the water molecule, thereby producing hydrogen, hydrogen peroxide, and oxygen per the following two relationships

$$
\begin{aligned}
\mathrm{H_2O} \xrightarrow{e^-} H_2^{\uparrow} \ &+ \ \mathrm{H_2O_2} \\
\mathrm{H_2O_2} \xrightarrow{e^-} O_2^{\uparrow} \ &+ \ \mathrm{H_2O} \quad .
\end{aligned}
\tag{11.170}
$$

After saturation and in the absence of oxygen removal processes such as corrosion, hydrogen and oxygen will evolve in gaseous form in places like the surge tank at the stoichiometric mixture ratio.

The measured rate of evolution of free hydrogen [Walz, 1967] is ∼0.3l H_2/MW-s [G(H_2)∼0.14 molecules H_2/100 eV]. Thus, for P_{av} ∼16 MW, the expected rate of evolution would be ∼5l H_2, less an allowance for power not directly deposited and dissipated in water or lost from the dump boundaries. The lower explosive limit (LEL) of hydrogen in air is 4%. A conservative maximum allowable concentration would be ∼1%. Thus, some 500l of the 1% H_2 in-air mixture has to be disposed of or processed per second. Since this gas contains most of the radioisotopes given above, release into the atmosphere is not an option.

One could build a delay line including holding tanks which would take care of the ^{15}O isotope. But ^{11}C with its $T_{1/2}$=20.3 min which evolves in the form of $^{11}CO_2$ is not manageable in a cost-effective way by this method. We have also studied the absorption of $^{11}CO_2$ in regenerative lime columns, but that method also proved to be cumbersome. Therefore, we developed at SLAC a catalytic H_2-O_2 recombiner with a capacity to process the radiolytically-evolved hydrogen for average beam powers up to 2 MW [Neal 1968, Walz 1969]. Four such recombiner facilities have been in use for more than 25 years with very low maintenance. Such a recombiner can be readily scaled up to process the gas from a 16-MW beam.

The recombiner uses a platinum-palladium catalyst through which the gases are pumped. The water vapor generated in the recombination of H_2 and O_2 is condensed out in the water spray which is generated by an ejector pump that provides the motive power to recirculate the gases. It is a closed-loop system and all gases and radioisotopes are always fully contained.

11.A.5 Summary

We showed that a beam dump to absorb and dissipate the power of 750 GeV/16-MW NLC beams is feasible using water as the primary power absorption medium. Such a dump could be a cylindrical vessel housing a potential vortex-like flow of water with the beam impinging at a radius of ∼0.3 m and a vortex flow velocity ∼1–1.5 m/s normal to the direction of the beam momentum vector. The vessel dimensions might be on the order of 1.5-m diameter and ∼6.5 m (∼18 X_o)-long for the water section followed by ∼1 m of water-cooled solid materials of higher Z to adequately attenuate the electro-magnetic cascade. The beam would enter the dump through a thin window (∼1-mm-thick copper would do). Isotope production in the water appears manageable relative to tunnel access requirements for maintenance and seems preferable to what can be expected for solid materials. The only drawback is production of ∼5l H_2/16 MW from radiolysis in water, but solutions exist to deal with this, notably the catalytic recombination of H_2 and O_2 into water, allowing for a closed-loop system containing all the radioisotopes.

References

[Burke 1995] D. Burke, "Proposal for New NLC Operating Plane", SLAC memorandum, unpublished (1995).

[Adolphsen 1995] C. Adolphsen, private communication (1995).

[Bernadini 1963] C. Bernadini *et al., Phys. Rev. Lett.* **10**, 407 (1963).

[Bambade 1989] P. Bambade *et al.,* "Observation of Beam-Beam Deflections at the SLAC Linear Collider", *Phys. Rev. Lett.* **62**, 2949 (1989); F. Zimmermann *et al.,* "Performance of the 1994/95 SLC Final Focus System", SLAC–PUB–95–6790, in *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Bane 1995] K. Bane, C. Adolphsen, F.-J. Decker, P. Emma, P. Krejcik, F. Zimmermann, "Measurement of the Effect of Collimator Generated Wakes on the Beams in the SLC", SLAC–PUB–95–6865, in *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Berz 1989] M. Berz, "Differential-algebraic description of beam dynamics to very high orders", *Particle Acclerators* **24**, 109 (1989).

[Bini 1991] C. Bini *et al., Phys. Lett.* **B** 135 (1991).

[Brinkmann 1990] R. Brinkmann, "Optimization of a Final Focus System for Large Momentum Bandwidth", DESY–M–90–14 (1990).

[Brown 1977] K. Brown, F. Rothacker, D. Carey, C. Iselin, "TRANSPORT - A Computer Program for Designing Charged Particle Beam Transport Systems", SLAC–91 (1977).

[Brown 1979] K. Brown and R. Servranckx, "Chromatic Correction for Large Storage Rings," SLAC–PUB–2270, in *IEEE Trans. Nucl. Sci.* **26**, 3598 (1979).

[Bula 1995] C. Bula, K.T. McDonald, E.J. Prebys (Princeton U.), C. Bamber, S. Boege, T. Kotseroglou, A.C. Melissinos, D.D. Meyerhofer and W. Ragg (Rochester U.), D.L. Burke, R.C. Field, G. Horton-Smith, A.C. Odian, J.E. Spencer, D. Walz (SLAC), S.C. Berridge, W.M. Bugg, K. Smakov and A.W. Wiedemann (Tennessee), "Observation of Nonlinear Effects in Compton Scattering", Submitted to *Phys. Rev. Lett.* , (Nov. 1995).

[Carlslaw 1959] Carlslaw, H.S. and Jaeger, J.C., "Conduction of Heat in Solids", (Oxford University Press, London, 1959).

[Carey 1982] D.C. Carey *et al.,* "DECAY TURTLE", SLAC Report 246 (March 1982).

[Chen 1993] P. Chen, "Disruption Effects from the Collision of Quasi-Flat Beams", SLAC–PUB–6215 (1993).

[delCorral 1982] J. del Corral, J.E. Spencer, S. St. Lorant, "Effects of Strong Transverse Fields on Permanent Magnets", SLC-CN-189 (1982).

[Dragt 1976] A.J. Dragt and J.M. Finn, "Lie Series and Invariant Functions for Analytic Symplectic Maps", *J. Math. Phys.* **17**, 2215–2227 (1976).

[Emma 1992] P. Emma, "Beam-Based Alignment of Sector-1 of the SLC Linac", *Proc. of 3rd EPAC*, Berlin (1992)

[Field NIM] Clive Field, "Wire Scanner System for the FFTB", NIM, To Be Published. See also SLAC–PUB–6717.

[Forest 1989] É. Forest, M. Berz and J. Irwin, "Normal form methods for complicated periodic systems: a complete solution using differential algebra and Lie operators", *Particle Acclerators* **24**, 91 (1989).

[Forest] The code DESPOT was written by É. Forest; the code TRACY by J. Bengtsson.

[Gross 1994] G. Gross, J. Spencer and SLAC's MMG, "Experience with the SLC Permanent Magnet Multipoles", Fourth European Part. Acc. Conf., London, England (1994). Also see SLAC–PUB–6558.

[Hartman priv] S. Hartman and T. Shintake, private communication.

[Hartman priv] S. Hartman, T. Shintake, N. Akasaka "Nanometer Resolution BPM Using Damped Slot Resonator", SLAC–PUB–95–6908, presented at 1995 Part. Acc. Conf., (1995).

[Hendrickson 1995] L. Hendrickson, private communication and this ZDR, Appendix D (1995).

[Hertzbach 1995] S. Hertzbach, private communication (1995).

[Hirata 1989] K. Hirata, B. Zotter, and K. Oide, *Phys. Lett. B* **224**, 437 (1989).

[Hollebeek 1981] The primary reference on disruption effects and their simulations is: Robert Hollebeek, "Disruption Limits for Linear Colliders", *Nucl. Instr. and Methods* **184**, 333 (1981).

[Irwin 1991] J. Irwin, "Final Focus System Optimization", *Proc. of Linear Colliders. 3rd International Workshop*, LC 91, Protvino, USSR, **Vol. 3**, V. Balakin *et al.*, (ed.) (1991).

[Irwin 1992] J. Irwin, "The Application of Lie Algebra Techniques to Beam Transport Design", SLAC–PUB–5315, published in *Nucl. Instr. and Methods* **A298**, 460 (1990).

[Irwin 1995a] These conditions were derived by J. Irwin (1995).

[Irwin 1995b] J. Irwin, presentation at SLAC (1995).

[Juravlev 1995] V.M. Juravlev, P.A. Lunev, A.A. Sery, A.I. Sleptsov, K. Honkavaara, R. Orava, E. Pietarinen, "Seismic Conditions in Finland and Stability Requirements for the Future Linear Collider", (1995).

[Keller 1993] L. Keller, "Muon Background in a 1-TeV Linear Collider", SLAC–PUB–6385 (1993).

[Lomperski 1993] M. Lomperski, "Compton Scattering off Blackbody Radiation and other Backgrounds of the HERA Polarimeter", *DESY* **93-045** (1993).

[Neal 1968] Neal, R.B., Editor, "The Stanford Two-Mile Accelerator", Chapter 20, (W.A. Benjamin, Inc., New York, 1968).

[Norem 1995] J. Norem *et al.*, "Tests of a High Resolution Beam Profile Monitor", *Proc. of the 1995 Part. Accel. Conf.*, Dallas, Texas, (1995).

[Oide 1988] K. Oide, "Synchrotron-Radiation Limit on the Focusing of Electron Beams", *Phys. Rev. Lett.* **61**, 15 (1988) 1713.

[Oide 1992] K. Oide, "Final Focus System with Odd-Dispersion Scheme", *KEK Preprint* **92**–58 (1992).

[Palmer 1990] R.B. Palmer, "Prospects for High Energy e^+e^- Linear Colliders", *Ann. Rev. Nucl. Sci.* **40**, 529 (1990).

[Pitthan 1995] R. Pitthan, "Re-Alignment: It is the Tunnel Floor which Moves, isn't It?", SLAC–PUB–95–7043, invited talk at 4th Int. Workshop on Acc. Alignment, Tsukuba (1995).

[Piwinski 1985] A. Piwinski, "Beam Losses and Lifetime", in CERN Accelerator School, Gif-sur-Yvette, France, *CERN* **85–19** Vol. II (1985).

[Raimondi 1995] P. Raimondi, F.J. Decker, P. Chen, "Disruption Effects on the Beam Size Measurement," SLAC–PUB–95–6882.

[Raimondi 1993] P. Raimondi, P.J. Emma, N. Toge, N.J. Walker, V. Ziemann, "Beam Based Alignment of the SLC Final Focus Superconducting Final Triplets", SLAC–PUB–95–6212. in *Proc. 1993 Part. Accel. Conf.*, Washington, DC (1993).

[Raimondi 1995] P. Raimondi, private communication (1995).

[Rohatgi 1987] R. Rohatgi, H.A. Schwettman and T.I Smith, "A Compact Energy Recovered FEL for Biomedical and Material Science Applications", *Proc. 1987 Part. Accel. Conf.*, Washington, DC, 1, 230 (1987).

[Rossi 1952] B. Rossi, *High-Energy Particles* (Prentice Hall, New York, 1952).

[Ross 1991] M.C. Ross *et al.*, "Wire Scanners for Beam Size and Emittance Measurements at the SLC", SLAC–PUB–5556 (May 1991).

[Ross 1994] M. Ross, *Proc. of Advanced Accelerator Concepts Workshop*, Lake Geneva (1994).

[Roy 1992] G. Roy, "Analysis of the Optics of the Final Focus Test Beam Using Lie Algebra Based Techniques", SLAC–397 (1992).

[Sands 1985] M. Sands, SLAC/AP–47 (1985).

[Servranckx 1990] R. Servranckx, K.L. Brown, L. Schachinger and D. Douglas, "User's Guide to the Program DIMAD", SLAC–0285 (1990).

[Shintake 1994] T. Shintake, "First Beam Test of Nanometer Spot Size Monitor Using Laser Interferometry", KEK-PREPRINT–94–129 (1994).

[Spencer 1995] J.E. Spencer, private communication (1995).

[Spencer 1995a] J.E. Spencer, "The SLC as a Second Generation Linear Collider", *Proc. 16th IEEE Part. Accel. Conf. and Int'l. Conf. on High Energy Accel's.*, Robt. Siemann ed., Dallas, TX, (1995).

[Spencer 1995b] J. Spencer, J. Irwin, D. Walz and M. Woods, "The SLAC NLC Extraction and Diagnostic Line", 1995 Part. Accel. Conf., Dallas, TX (1995).

[Spencer 1995c] J.E. Spencer, and "Beam-Beam Effects and Generalized Luminosity", SLAC–PUB–7051, in *Proc's: Workshop on Electron-Electron Colliders*, C. Heusch ed., Santa Cruz, CA, (1995);

[Telnov 1987] V.I. Telnov, "Scattering of Electrons on Termal Radiation Photons in Electron-Positron Storage Rings", *Nucl. Instr. Methods A* **260**, 304 (1987).

[Tenenbaum 1995] P. Tenenbaum, D. Burke, R. Helm, J. Irwin, K. Oide, and K. Flöttmann, "Beam-Based Magnetic Alignment of the Final Focus Test Beam", SLAC–PUB–95–6769 (1995).

[Walker 1993] N.J. Walker, J. Irwin, M. Woodley, "Global Tuning Knobs for the SLC Final Focus", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC (1993).

[Walker 1987] R.P. Walker, "Calculation of the Touschek Lifetime in Electron Storage Rings", Proceedings of IEEE PAC 1987, Washington, DC, 491 (1987).

[Walz, 1967] D.R. Walz and E.J. Seppi, "Radiolysis and hydrogen evolution in the A-beam dump radioactive water system", Report No. SLAC–TN–67–29, Stanford Linear Accelerator Center, (October, 1967).

[Walz 1967a] D.R. Walz *et al.,* "Beam Dumps, energy slits and collimators at SLAC—their final versions and first performance data", *IEEE Trans. Nucl. Sci.* **3** NS-14, 923–927 (1967).

[Walz 1969] D.R. Walz and L.R. Lucas, "The `sphere dump'—a new low-cost high-power beam dump concept and a catalytic hydrogen-oxygen recombiner for radioactive water systems", *IEEE Trans. Nucl. Sci.* **3** NS-16, 613–617 (1969).

[Wilson 1995] P. Wilson supplied the scaling laws of Table 11-38, and the analysis of the consequences of thermal changes of the waveguide.

[Woodley 1994] M.D. Woodley, unpublished (1994).

[Woods 1994] M. Woods, "Polarization at SLAC", SLAC–PUB–6694, (October 1994).

[Yokoya 1986] K. Yokoya, ABEL, "A Computer Code for the Beam-Beam Interaction in Linear Colliders", *Nucl. Instr. and Methods* **B251** 1 (1986); Toshiaki Tauchi *et al., Particle Acclerators* **41**, 29 (1993). We thank Mike Ronan for providing the latest version of ABEL (Analysis of Beam-beam Effects in Linear colliders).

[Yokoya 1988] K. Yokoya and P. Chen, "Depolarization Due to Beam-Beam Interaction in Electron-Positron Linear Colliders", SLAC–PUB–4692 (1988). See also: US-CERN School on Part. Accel's. (1990).

[ZEUS/H1] See numerous papers by ZEUS or H1 collaborations on photoproduction in HERA, *e.g.,* C. Kiesling, "Physics from the First Year of H1 at HERA", DESY–94–137 (1994).

[Zimmermann 1995] F. Zimmermann, R. Helm and J. Irwin, "Optimization of the NLC Final Focus System", SLAC–PUB–95–6791, presented at 1995 Part. Acc. Conf., (1995).

[Zimmermann 1995b] F. Zimmermann *et al.*, "Performance of the 1994/95 SLC Final Focus System", SLAC–PUB–95–6790 (1995). presented at 1995 Part. Acc. Conf., (1995).

Contributors

- Gordon Bowden
- Paul Emma
- Karl Brown
- Dave Burke
- Leif Eriksson
- Dick Helm
- Stan Hertzbach
- John Irwin
- Eugene Kraft
- Robert Messner
- Yuri Nosochkov
- Katsunobu Oide
- Tor Raubenheimer
- Jim Spencer
- Peter Tenenbaum
- Fang Tian
- Dieter Walz
- Perry Wilson
- Mike Woods
- Mark Woodley
- Frank Zimmermann

12

The Interaction Region

Contents

12.1 Introduction

Table 12-1 summarizes some of the machine design parameters important to the physics capabilities of the NLC and to the design of the interaction region (IR) and detector. Three sets of parameters are considered for each of the two machine energies. They define a volume in parameter space within which the luminosity is roughly constant.

Depending on the exact performance of each of the NLC's subcomponents, we will find ourselves somewhere in that space. The rf power system controls the bunch charge and the number of bunches that can be accelerated each machine cycle. The bunch compression system and requirements on momentum spread will determine the minimum bunch length, σ_z. The final doublet β functions, β_x and β_y, must be larger than σ_z. The invariant emittance in y depends on the performance of the damping rings and on how well the linac and final focus can transport a low-emittance beam. The final value of ϵ_y achieved, together with β_y, will determine the value of the y spot size, σ_y. This in turn will set the scale for vibration tolerance and field stability of the final quadrupoles. Assuming the damping rings provide the specified ϵ_x, β_x is a semifree parameter, adjusted to give the desired luminosity while keeping deleterious beam-beam interaction effects to an acceptable level.

The choice of X-band rf for the NLC sets the interbunch separation at 1.4 ns (or 42 cm). To have each bunch interact only with its partner, there must be a crossing angle at the IP. This crossing angle is put into the NLC at the big bend, just after the collimation section located after the linac and before the final focus. The bend angle then also helps to reduce the muon flux that results when the beam interacts with the collimators. The size of the collimator apertures are measured in terms of the number of beam widths. Physical limitations and wakefield effects imply minimum collimator apertures corresponding to $7\sigma_x \times 35\sigma_y$. Beam tails and these apertures determine the level of muons produced. The apertures, the final-focus lattice, and assumptions on what the non-Gaussian profile of the beam may be determine the production of synchrotron radiation (SR). The design of the masking system and the value of detector's solenoidal field control the backgrounds caused by the SR photons.

The high charge density in each bunch causes particles in one bunch to interact with the overall field of the opposing bunch. This beam-beam interaction results in a luminosity enhancement as the beams are attracted to each other. The resulting acceleration, however, results in the copious production of photons, which will smear the luminosity spectrum as a function of $\sqrt{s}$. The photons can themselves interact coherently with the field of the opposing bunch, or interact with the individual $e^{\pm}$ of the opposing bunch to produce e^+e^- pairs. While predominately produced at low p_t, these pairs can cause problems in the detector and must also be controlled by the solenoidal field and the masking. The $\gamma\gamma$ interactions can cause the production of jets of hadrons with high p_t. The problem is exacerbated by the 1.4-ns bunch structure of the NLC. Particle detectors with good timing resolution will be required to separate the background hits arising from other bunches in the train from those hits produced in the interaction that caused the trigger. Trigger schemes may need to be developed to control the rate at which these backgrounds trigger the detector.

The issue of dealing with the 3.6–7.0-nm y-spot sizes strongly affects the discussion of the IR. The source terms in the problem are the naturally occurring ground vibrations and ground motion driven by local laboratory sources, such as pumps and the flow of cooling fluids. The support structures which stabilize the final doublet against vibration at the nm level will reside within the detector. Any device that might be required to sense or control the inertial or relative movement of the the quadrupoles must be accommodated by the detector. Finally, detectors usually use various fluids to cool their magnet coil and electronics; the influence of fluid flow on the final-doublet vibration will need to be understood.

	500 GeV			1 TeV		
	A	B	C	A	B	C
Trains/s		180			120	
Bunches/train		90			90	
Interbunch spacing (ns)		1.4			1.4	
Electrons/bunch (10^{10})	0.65	0.75	0.85	0.95	1.1	1.25
σ_x-spot size (nm)	256	286	286	202	221	256
σ_y-spot size (nm)	4.23	5.25	6.71	3.39	4.11	5.15
σ_z-spot size (μm)	100	125	150	125	150	150
x beam divergence (μrad)	32.0	28.6	28.6	20.2	18.5	16.0
y beam divergence (μrad)	33.8	35.0	33.5	27.1	27.4	25.8
Minimum collimator apertures	$7\sigma_x \times 35\sigma_y$			$7\sigma_x \times 35\sigma_y$		
$\mathcal{L}$ (geometric) (10^{33} cm^{-2} s^{-1})	5.0	4.8	4.9	11.3	11.4	10.2
H_D - Luminosity Enhancement Factor	1.46	1.45	1.52	1.41	1.43	1.53
$\mathcal{L}$ (with pinch) (10^{33} cm^{-2} s^{-1})	7.3	7.0	7.4	16.0	16.4	15.6

Table 12-1. *List of the NLC parameters that define the operating range of the interaction region and detector.*

12.2 The Luminosity Spectrum

Ideally, the beam energy at the NLC would be a delta function at the design energy. It could then trivially be used as a constraint in any physics analysis. The finite energy spread of the NLC, initial-state radiation effects, and the production of photons in the beam-beam interaction will degrade the delta function to a spectrum. The spectrum can be unfolded from any physics analysis if it is not too broad, where "too" is determined by the specific physics channel of interest, and if we can accurately measure the spectrum.

12.2.1 Beam Energy Spread

Figure 12-1 shows the expected beam energy spread at the IP for the 1-TeV machines. The shape comes from the bunch compression that sets the beam spot size in z. It is similar for the 500-GeV machine.

12.2.2 Initial State Radiation

At the NLC, the effects of initial-state radiation (ISR) will be about the same as that at SLC or LEP II. The scale of the problem is set by the variable $L = \ln \frac{s}{m_e^2}$. L varies from 24.2 to 27.6, 29.0, or 29.8 as $\sqrt{s}$ goes from m_z to 500, 1000, or 1500 GeV. The amount of ISR is irreducible and therefore sets the scale for how small the luminosity smearing due to the beam energy spread and the beamstrahlung must be. The electron energy distribution, $D_e(z,s)$, in the presence of ISR has been calculated by Fadin and Kuraev [Kuraev, 1985]. (See their Eqs. (20) and (21)). Here $\sqrt{s}$ is the center-of-mass energy of the interacting e^+ and e^- and z is the fractional e^+ or e^- energy following ISR.

The cross section has the form,

$$\sigma(s) \propto \int\int dx_1 dx_2 D(x_1,s) D(x_2,s) \sigma(s x_1 x_2). \tag{12.1}$$

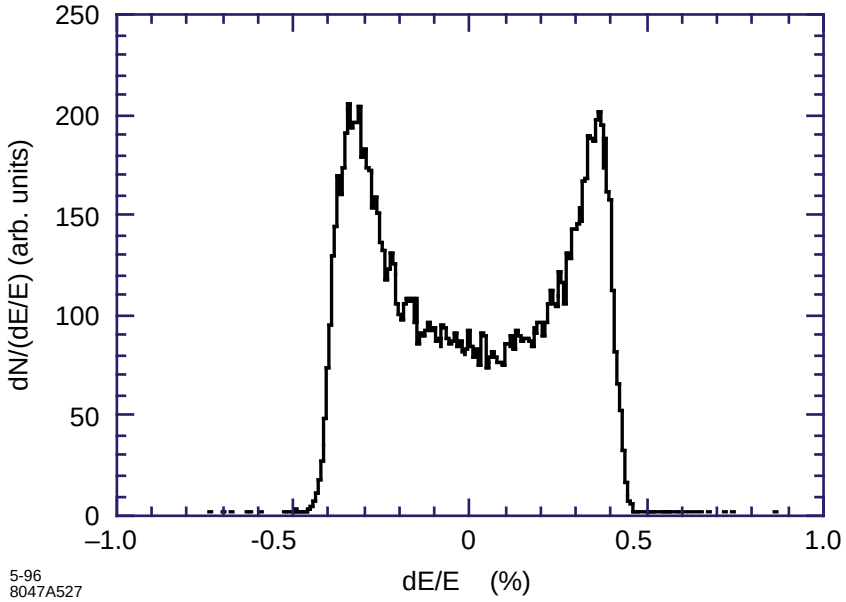


Figure 12-1. *Beam energy spread for the 1-TeV lattice.*

	m_Z	500 GeV	1 TeV	1.5 TeV
Mean e^+e^- c.m. energy loss	4.19%	4.77%	5.00%	5.14%
Rms e^+e^- c.m. energy spread	11.2%	11.9%	12.2%	12.4%
% of $\mathcal{L}$ within 0.1% of $\sqrt{s}_{\rm nom}$	54.7%	50.1%	48.3%	47.3%
% of $\mathcal{L}$ within 0.5% of $\sqrt{s}_{\rm nom}$	65.2%	61.2%	59.6%	58.7%
% of $\mathcal{L}$ within 1% of $\sqrt{s}_{\rm nom}$	70.2%	66.6%	65.2%	64.4%
% of $\mathcal{L}$ within 5% of $\sqrt{s}_{\rm nom}$	82.8%	80.5%	79.6%	79.1%
% of $\mathcal{L}$ within 10% of $\sqrt{s}_{\rm nom}$	88.3%	86.7%	86.0%	85.6%

Table 12-2. *Effect of ISR on e^+e^- luminosity spectra.*

Defining $sx_1x_2 = (1-x)s$ and integrating at fixed x yields,

$$\sigma(s) = \int dx \sigma(s(1-x)) F(x,s), \tag{12.2}$$

where $F(x,s)$ (Eq. (28) of Fadin & Kuraev) represents the luminosity distribution function.

The dashed curve in Figure 12-2 shows the electron momentum distribution function, $D_e(z,s)$, plotted for $\sqrt{s}$=500 GeV as a function of $z = p_{\rm beam}/p_{\rm max}$, the fractional lepton momentum. The solid curve shows the luminosity spectrum $F(x,s)$ for $\sqrt{s}_{\rm max} = 500$ GeV, plotted as a function of $z = \sqrt{s}/\sqrt{s}_{\rm max}$. Here $1-z = \sqrt{1-x}$. Each curve is normalized to unit area. When plotted as in Figure 12-2, one cannot distinguish any difference in either the electron momentum distribution function or in the luminosity distribution over the range of $\sqrt{s}$ considered. Any small differences can be seen by examining Table 12-2, which shows the fraction of the luminosity within 0.1%, 0.5%, 1%, 5%, and 10% of the nominal center-of-mass energy, as well as the mean energy loss and the rms energy spread, for each of the four values of $\sqrt{s}$ considered.

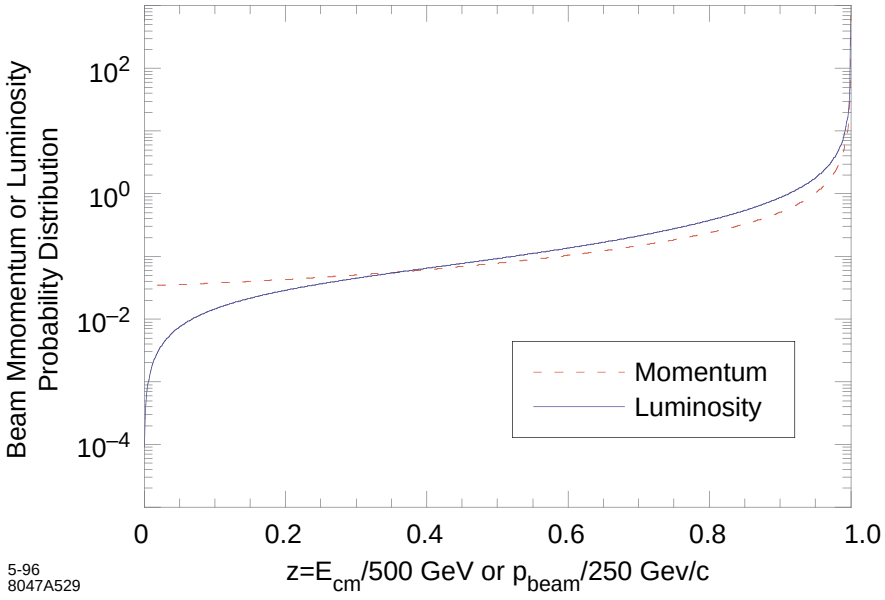


Figure 12-2. *The dashed curve shows the electron momentum distribution as a function of the normalized electron momentum after ISR. The solid curve shows the luminosity distribution after ISR as a function of the normalized center of mass energy. While the plot corresponds to a center of mass energy of 500 GeV, no difference would be visible for any center of mass energy between m_Z and 1.5 TeV.*

12.2.3 Beamstrahlung

As one bunch of beam particles passes through the electromagnetic field of the opposing bunch, radiation will be emitted. The photons thus produced are called "beamstrahlung" photons. The subject has been extensively discussed in the literature. The results herein are taken from Ref. [Chen, 1990] and Ref. [Chen, 1992].

The production of beamstrahlung photons is a stochastic process. The probability that a given particle will radiate and the characteristic energy of the radiation are determined by the field density or equivalently by the number of electrons or positrons per bunch and the bunch dimensions. When this density is high relative to the critical electric field, beamstrahlung is more likely to occur. The critical electric field, $E_{\rm critical}$, is defined as that field which, when an electron travels one Compton wavelength in it, does an amount of work equal to $m_e c^2$:

$$eE_{\rm critical}\lambda_c \equiv m_e c^2 \tag{12.3}$$

The magnetic field equivalent is:

$$B_{\rm critical} = \frac{\alpha e}{r_e^2} \sim 4.4 \times 10^9 \text{ Tesla} \tag{12.4}$$

The field strength is measured in terms of the dimensionless quantity

$$\Upsilon \equiv \frac{\gamma B_{\rm bunch}}{B_{\rm critical}} = \frac{E_{\rm bunch}}{E_{\rm critical}} \tag{12.5}$$

where

$$B_{bunch} \cong \frac{5eN_e}{6\sigma_z\left(\sigma_x+\sigma_y\right)} \quad . \tag{12.6}$$

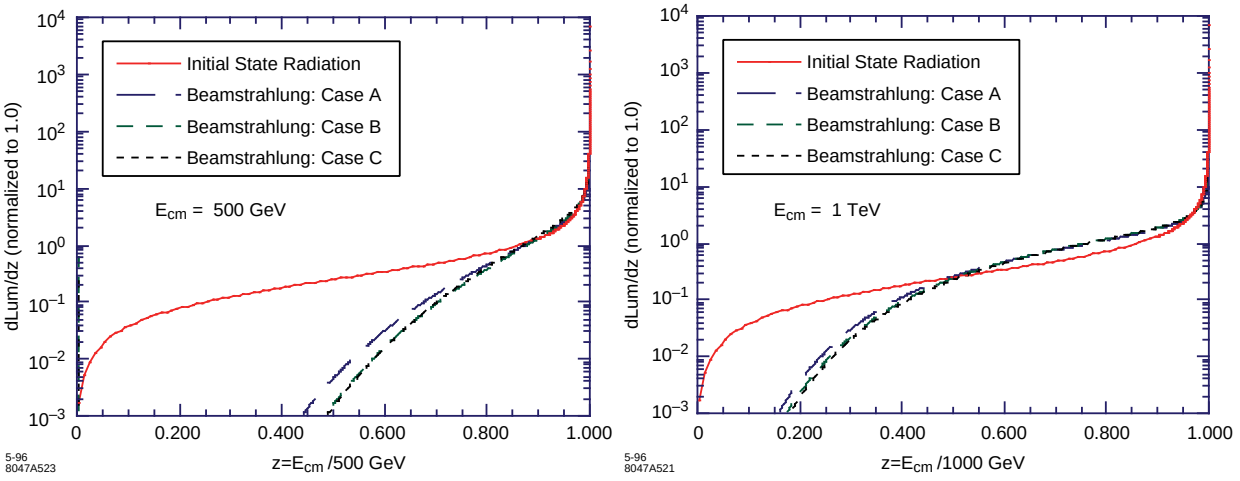


Figure 12-3. *The luminosity spectrum after beamstrahlung at 500 GeV and 1 TeV for the three parameter sets under consideration. The spectrum due to initial state radiation is presented for comparison.*

In Ref. [Chen, 1992], Chen derives an expression (see his Eq.(24)) for the differential luminosity which is composed of three parts, corresponding to the cases when either no, one, or more than one photon is emitted. Each part is expressed in terms of N_γ, the mean number of photons produced per electron,

$$N_\gamma = \frac{5}{2}\frac{\alpha\sigma_Z}{\gamma\lambda_c}\Upsilon\left(1+\Upsilon^{2/3}\right)^{-1/2} . \tag{12.7}$$

The shape of the curve is generally described by N_γ, by the average energy loss per electron, δ_b

$$\delta_\mathrm{b} = \frac{1}{2}N_\gamma\Upsilon\frac{(1+\Upsilon^{2/3})^{1/2}}{(1+(\frac{3}{2}\Upsilon)^{2/3})^2} \quad , \tag{12.8}$$

and by the fraction of the luminosity at the full nominal beam energy,

$$\frac{1}{N_\gamma^2}\left[1-e^{N_\gamma}\right]^2 \quad . \tag{12.9}$$

In Figure 12-3, we plot the luminosity spectra for the three parameter sets at 500 GeV and 1 TeV. The ISR curve is added for comparison. In Figure 12-4, the region of the luminosity spectra within 2% of the nominal beam energy is plotted for the three parameter sets at 500 GeV and 1 TeV. Figure 12-5 presents a comparison of Case A for the 500-GeV and 1-TeV machines.

Table 12-3 summarizes the situation by listing the values of Υ, N_γ, δ_b, and the fraction of the luminosity within 0%, 0.1%, 0.5%, 1%, 5%, and 10% of the nominal center of mass energy. The fraction at 0% is taken from the above expression, while the others are obtained by integrating the distributions in Figure 12-3,

The machine parameter sets have been chosen so that the beamstrahlung induced smearing of the luminosity spectrum is not substantially worse than the luminosity smearing due to ISR. As the ISR part of the smearing is calculable and small, it should be comparatively easy to unfold from the data. Comparing Tables 12-2 and 12-3 we find that at 500 GeV all three parameter set choices result in smearing below the level expected from ISR. At 1 TeV, there is roughly twice as much smearing as that from ISR.

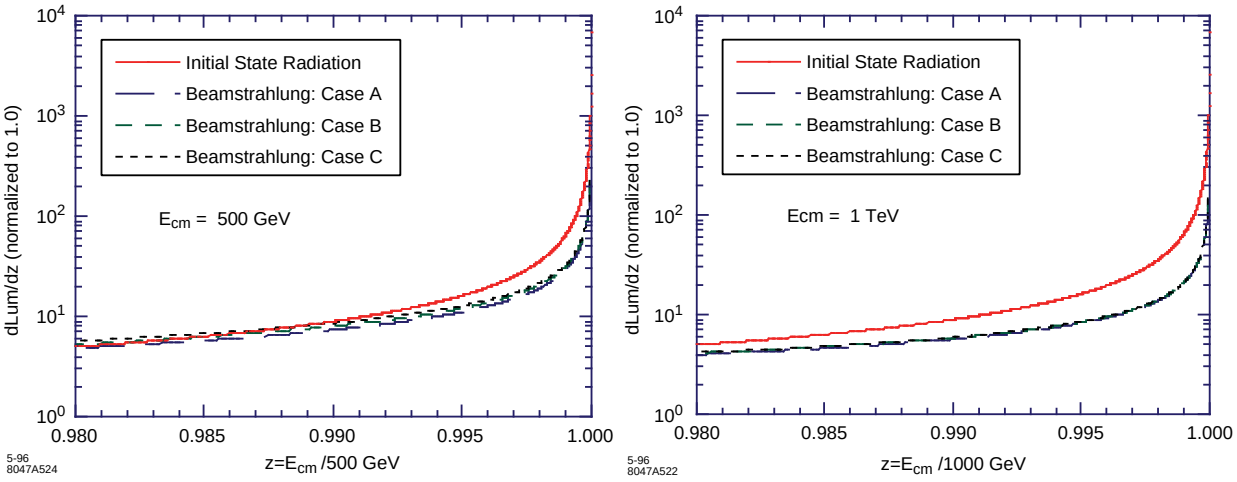


Figure 12-4. *A closeup of the region of the luminosity spectrum due to beamstrahlung near the nominal beam energy at 500 GeV and 1 TeV for the parameter sets listed previously. The spectrum due to ISR is presented for comparison.*

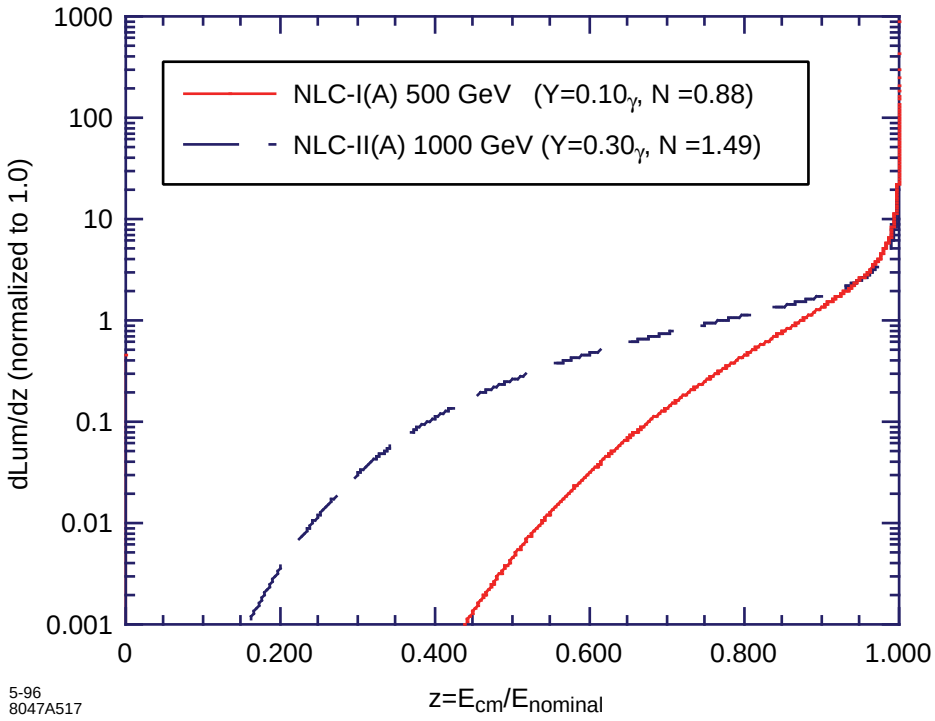


Figure 12-5. *A comparison of the beamstrahlung spectra for case A of the 500-GeV and 1-TeV machines.*

	500 GeV			1 TeV		
	A	B	C	A	B	C
Υ	0.112	0.0924	0.0878	0.334	0.296	0.289
N_γ	0.973	1.017	1.162	1.654	1.781	1.745
$\delta_{\rm b}$	3.55%	3.21%	3.52%	12.6%	12.6%	12.2%
R.M.S. e^+e^- c.m. energy spread	7.17%	6.44%	6.62%	15.9%	15.5%	15.2%
% of $\mathcal{L}$ within 0% of $\sqrt{s}_{\rm nom}$	40.9%	39.4%	35.0%	23.9%	21.8%	22.4%
% of $\mathcal{L}$ within 0.1% of $\sqrt{s}_{\rm nom}$	48.1%	47.0%	42.7%	28.6%	26.4%	27.2%
% of $\mathcal{L}$ within 0.5% of $\sqrt{s}_{\rm nom}$	54.5%	54.0%	50.0%	33.1%	31.0%	31.8%
% of $\mathcal{L}$ within 1% of $\sqrt{s}_{\rm nom}$	59.0%	58.8%	55.0%	36.3%	34.2%	35.1%
% of $\mathcal{L}$ within 5% of $\sqrt{s}_{\rm nom}$	75.6%	76.8%	74.4%	49.8%	48.1%	49.2%
% of $\mathcal{L}$ within 10% of $\sqrt{s}_{\rm nom}$	85.6%	87.2%	85.9%	60.2%	59.1%	60.1%
% of $\mathcal{L}$ within 100% of $\sqrt{s}_{\rm nom}$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 12-3. *Summary of quantities parameterizing the effect of beamstrahlung on the luminosity spectrum.*

12.2.4 Measurement of the Luminosity Spectrum

The significance of the luminosity smearing depends on the physics channel under study. For discovery physics, some measure such as the fraction of the luminosity with $\sqrt{s} > 90\%$ of nominal is probably the best figure of merit. However, attention is generally focused on the region within 1% of the nominal energy, as it is critical for an excellent determination of the top quark mass. D. Miller has argued [Frary] that the measurement of the top quark mass will require machine parameters that result in very little smearing and that the detector have excellent forward tracking so as to use the collinearity distribution of Bhabha events to unfold the luminosity spectrum.

In Figure 12-6, we show the cross section for $t\bar{t}$ production as a function of nominal center-of-mass energy for m_t = 180 GeV/c^2. The theoretical cross section, indicated as curve (a), is based on the results of Peskin and Strassler [Peskin, 1991] with $\alpha_s(M_Z^2) = 0.12$, infinite Higgs mass, and nominal Standard Model couplings. Each energy-smearing mechanism, initial-state radiation (b), beamstrahlung (c), and beam energy spread (d), has been successively applied. Hence, curve (d) includes all effects.

A comment on the beam energy spread is in order. The expected shape of the single-beam energy spread is given in Figure 12-1. The luminosity-weighted center-of-mass energy spread, $\Delta E_{\rm cm}/E_{\rm cm}$, is calculated from the single-beam distribution given its dependence on the bunch spatial distribution, as discussed in Chapter 4. The resulting center-of-mass energy distribution is given in Figure 12-7, where the calculation has been simplified by ignoring variations in energy in transverse space. In this case, the input single-beam energy spread has FWHM of 0.8%, corresponding to the distribution shown in Figure 12-1. The resulting distribution in $\Delta E_{\rm cm}/E_{\rm cm}$ is strongly peaked at zero with an rms of 0.38%, as indicated in the Figure 12-1. It is expected that the single-beam energy spread can be comfortably adjusted within the FWHM interval 0.6% to 1.0%. For the top threshold, it is clear that the smaller width is preferred, and curve (d) of Figure 12-6 was calculated using the 0.6% width. The large top mass of about 180 GeV/c^2 presents a relatively broad, featureless threshold shape which is not nearly as sensitive to the $\Delta E_{\rm cm}/E_{\rm cm}$ distribution as would be expected if the top mass were smaller. Figure 12-8 shows the change in shape of the threshold cross section as the single-beam energy spread is increased.

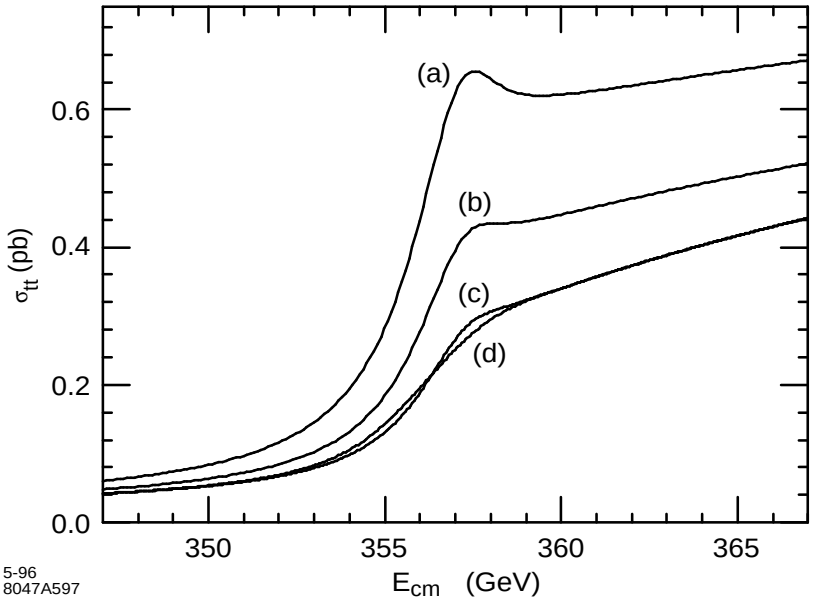


Figure 12-6. *Production cross section for top quark pairs near threshold for m_t = 180 GeV/c^2. The theoretical cross section is given by curve (a), to which the energy re-distribution effects have been applied. Curve (b): ISR; curve (c): ISR and beamstrahlung; curve (d): ISR, beamstrahlung, and beam energy spread.*

12.3 Detector Background Sources

12.3.1 Beamstrahlung-produced e^+e^- Pairs

As the beamstrahlung photons travel in the high-field region of the opposing bunch, they can produce e^+e^- pairs. If deflected to large enough angles, the pairs can enter the detector and cause unwanted backgrounds. The probability for pair production is set by the Υ parameter. When $\Upsilon \geq\sim 0.3$ the pairs are coherently produced; that is, the virtual electron-positron pairs accompanying the beamstrahlung photon can exchange energy-momentum with the strong electromagnetic field and be kicked on-shell. When $\Upsilon \leq\sim 0.3$ the incoherent pair creation processes become dominant. In these, the e^+e^- pairs are created from individual scattering of real beamstrahlung photons through the Breit-Wheeler ($\gamma\gamma \rightarrow e^+e^-$) and Bethe-Heitler ($e^\pm\gamma \rightarrow e^\pm e^+e^-$) processes, and from the scattering of virtual photons through the Landau-Lifshitz ($e^+e^- \rightarrow e^+e^-e^+e^-$) process.

This phenomenon is dealt with in two ways. Most importantly, as long as $\Upsilon \leq\sim 0.3$ the number of coherently produced pairs per bunch crossing is negligible. This has been a design constraint of all colliders to date and seems reasonably easy to achieve at the start up energy of 500 GeV. Note the values of Υ listed in Table 12-3 for the various collider designs. Figure 12-9 plots the contributions from each of the sources of pair production per bunch crossing versus Υ. Table 12-4 lists the number of pairs produced for the different parameter lists at 500 GeV and 1 TeV.

The preceding chapter introduced the ABEL Monte Carlo program used to simulate the beam–beam interaction. The ABEL program divides each bunch longitudinally, produces beamstrahlung photons and e^+e^- pairs as the bunches overlap, and propagates the particles through the high fields of the bunch overlap region. Figure 12-9 and Table 12-4

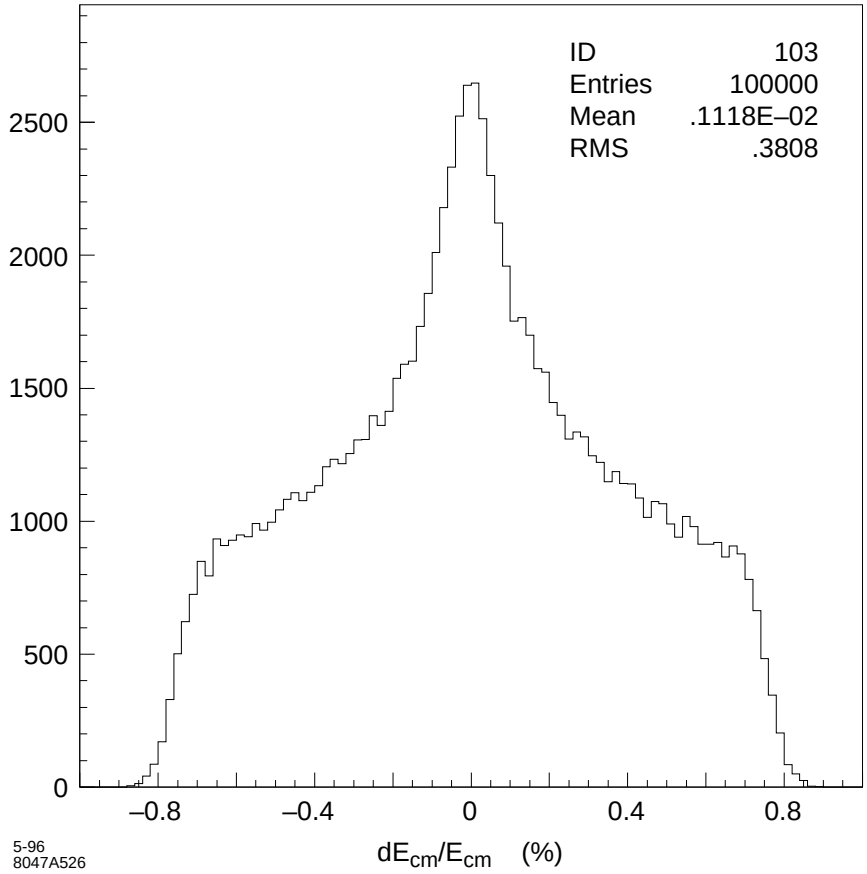


Figure 12-7. *Distribution of the center-of-mass energy, $\Delta E_{\rm cm}/E_{cm}$, due to the convolution of single-beam energy spread distributions for the two beams. The single-beam energy spread in this case has a FWHM of 0.8%, corresponding to the distribution given in Figure 12-1.*

are calculated using the parameterizations of ABEL described in Ref. [Chen, 1990]. In the last run of experiment E144 at the SLAC FFTB, values of $\Upsilon \sim 0.25$ were reached. Coherent pairs were observed at about the level predicted by the ABEL simulation. In its upcoming run, E144 should probe the region of $\Upsilon \sim 0.5$–0.6. This experimental confirmation of the background level anchors the estimates of detector hit densities arising as a result of pair creation which is discussed in later sections of this chapter.

Secondly, a masking scheme can be designed which, in conjunction with the detector's solenoidal field, protects the detector from the majority of the deflected pairs. The mask takes the form of a truncated cone which begins about 0.5 m from the interaction point and has an opening angle determined by the maximum kick the pair can receive from the field of the opposing beam. It defines a "deadcone" within which the detector is blind.

In coherent production, the pairs are not produced with significant intrinsic p_t; rather, roughly speaking, the e^+e^- are deflected by an angle proportional to $\sqrt{D_x/\epsilon} * \sigma_x/\sigma_z$, where D_x is the disruption parameter, ϵ is the ratio of the electron or positron momentum relative to the beam momentum, and σ_x and σ_z are the x and z beam spot sizes. The incoherent pairs have a more uniform energy distribution than the coherent process and can result in more particles with p_t and scattering angle θ outside the dead cone.

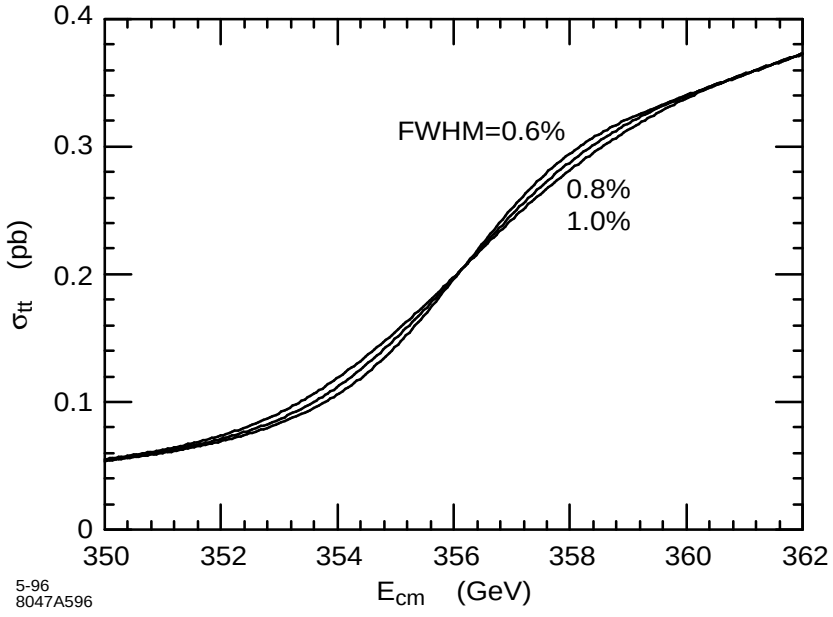


Figure 12-8. *Comparison of $t\bar{t}$ threshold shape, with all effects included, for different single-beam energy spreads. The three curves correspond to single-beam energy spread distributions with FWHM of 0.6%, 0.8%, and 1.0%, as indicated.*

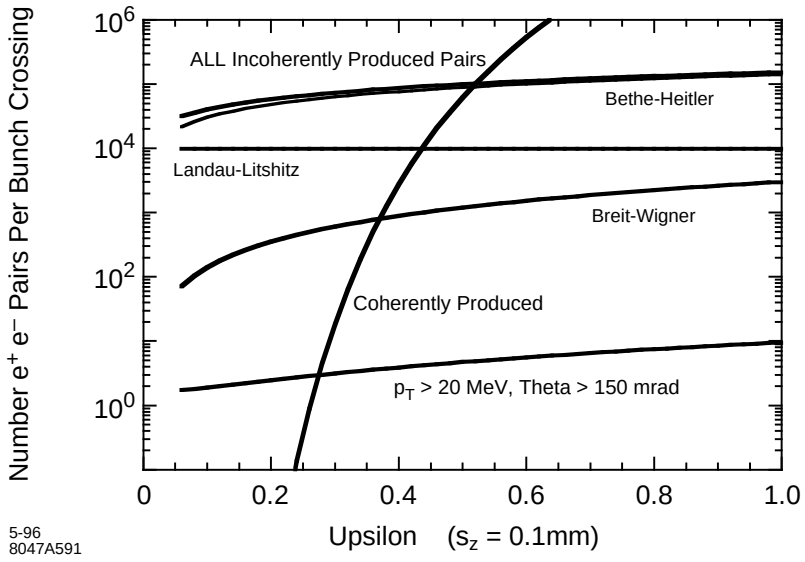


Figure 12-9. *Relative contribution of pairs from all sources as a function of Υ.*

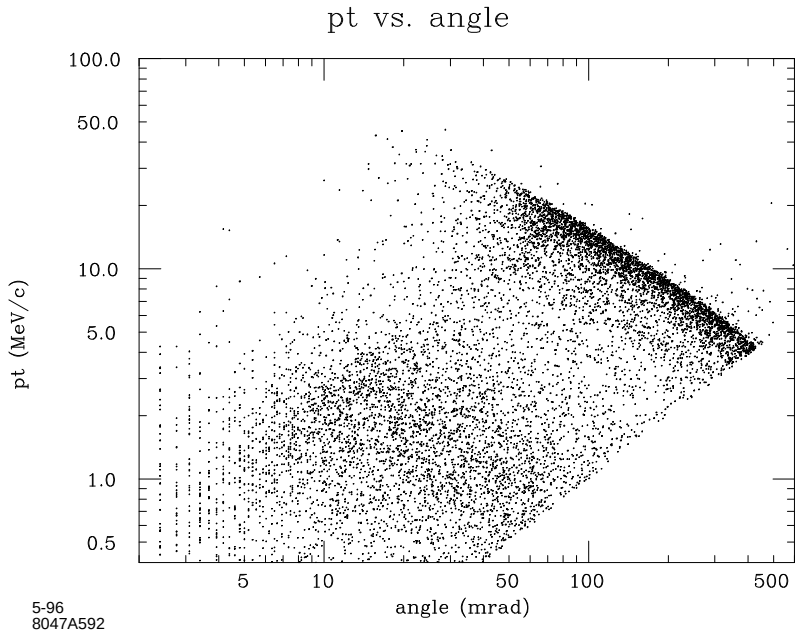


Figure 12-10. *p_t vs. θ distribution for pairs.*

The pair partner with the same sign as that of the bunch that produced it will tend to oscillate within the field of the oppositely charged opposing beam. In the flat beam designs under consideration, there will be more oscillations in the vertical plane than the horizontal plane. The vertical oscillations will tend to cancel and the exit angle of the particle will be predominately in the horizontal plane. The oppositely charged member of the pair will see a defocusing force from the opposing beam and be quickly deflected out of the beam. From outside the beam, it sees the field of an elliptic cylinder charge distribution; up to terms of order $\log(D_x/\epsilon)$ the result for the maximum scattering angle is the same as that for the same-sign particle, although in this case the deflection angles in both the x and y planes are comparable.

Figure 12-10 shows the scatter plot of transverse momentum and scattering angle of the pairs generated by ABEL. The simulation used a 10-MeV cut on pair-member energy and a 2-mr cut on angle. The two bands seen in the plot corresponds to the opposite sign partners in the higher p_t region and the same sign partners in the lower p_t region.

The particles travel in a helix from the IP. The solenoidal B-field strength and the distance of the conical mask from the IP determine the maximum particle momentum, and thus ϵ, that needs to be considered. All particles with lower momenta will curl up inside the cone. The maximum deflection angle can then be calculated from the formula alluded to above and the required radius of the conical mask determined. The dead cone is on the order of 100 mr and is tabulated for the various machine designs in Table 12-4. The thickness of the mask must be determined by detailed EGS or GEANT simulations to be adequate to stop the debris of the interacting electrons and positrons. Typically, the angle defining the outer dimension of the mask is 150–200 mr. Table 12-4 tabulates, for each of the parameter sets, the number of particles with p_t and θ large enough so that they fall outside the mask and hit the detector. Dealing with these particles depends on the time or bunch structure of the machine and on the timing capabilities of the detector.

	500 GeV			1 TeV		
	A	B	C	A	B	C
Detector dead-cone (mrad)	83	80	79	95	94	100
e^+e^- pairs per bunch X from coherent production	2.6E-13	1.4E-17	9.9E-19	76	13	9.1
e^+e^- pairs per bunch X from incoherent production	16.4E3	16.9E3	17.2E3	57E3	59E3	57E3
$N_{beamstr-e^\pm}$ /bunch $\theta > 150$ mr, p_T >20 MeV	2.1	2.1	2.4	6.8	7.4	7.1

Table 12-4. *Summary of quantities parameterizing the backgrounds, before they interact in the detector.*

12.3.2 Hadronic Backgrounds from $\gamma\gamma$ Interactions

In addition to the electromagnetic processes discussed above, the beamstrahlung photons can interact to produce hadrons and jets of hadrons. By folding the beamstrahlung spectrum into cross section estimates using an equivalent photon approximation, these hadronic rates can be estimated.

Hadronic events produced by beamstrahlung $\gamma\gamma$ annihilation are expected to be benign. Most of the hadronic events are minimum-bias events with small transverse momentum and small center-of-mass energy. Furthermore, the two photons involved in the collision usually have very different energies so that the hadronic system is highly boosted along the beam direction. Monte Carlo studies have demonstrated that the mean energy deposited in a detector from an hadronic background event will be 8 and 11 GeV for colliders with $\sqrt{s}$=500 and 1000 GeV, respectively. Here it is assumed that the detector has typical electromagnetic and hadronic calorimetry over all solid angle with the exception of a hole (dead cone) with $|\cos\theta| > 0.985$. The energy deposition with a larger solid angle dead cone of $|\cos\theta| > 0.900$ will be 3.3 GeV and 4.4 GeV for $\sqrt{s}$=500 and 1000 GeV, respectively.

12.3.3 Quadrupole and Bend Synchrotron Radiation

Linear collider designs obtain a small beam spot at the IP by using strong focusing magnetic quadrupole lenses close to the IP. SR generated by particles passing through these quadrupoles and bend magnets in the final focus is a potential source of background in the detector. SR backgrounds in the SLD detector at the SLAC Linear Collider (SLC) are in reasonable agreement with calculations. The SLD/SLC model implies that the nominal Gaussian beam core generates a small SR background in the final "soft" bend magnet, and negligible SR background in the quadrupoles. However, the non-Gaussian beam tail or beam halo can result in significant background due to synchrotron radiation generated in the quadrupoles (QSR). This source is controlled by collimation of the beam and by limiting the angular divergence at the IP. The QSR background is quite variable, and measures that control it can reduce luminosity.

Apertures in the beam line near the IP are matched to the design angular divergence and collimation. The SLD has a carefully designed set of internal masks to limit the SR background in the detector, especially the drift chamber. The SR background at SLD is reasonably well modeled by a nominal Gaussian core and a beam tail at the IP that is a very broad Gaussian containing 1% of the nominal charge. This "1% flat tail" is purely arbitrary and difficult to measure directly. The SR background generated by the tail depends directly on the fraction of the beam it contains and how it is collimated.

SR backgrounds in the SLD were modeled with EGS4, starting with SR photon fluxes calculated with a modified version of QSRAD. QSRAD was written to study SR generated by the PEP-I beam in the last two quadrupoles prior to the IP. The program traces weighted rays from a Gaussian beam profile through the specified magnetic optics and produces a geometric fan of synchrotron radiation with uniform power density and constant critical energy for each

magnetic element. These fans are then traced, and a tally is made of the fraction of each fan that strikes opaque surfaces with specified apertures. The distribution of photon critical energies is accumulated for each surface and converted to a photon energy distribution. The code provides additional information to characterize the SR photons incident on each surface, and this is the input data for an EGS4 model of the masking and detector.

At SLD the SR due to the final soft bend has a critical energy of 70 keV and results in 7×10^8 photons per pulse incident on a mask 1.3-m down beam from the IP. This is $\approx 1.5 \times 10^{10}$ keV per pulse at 120 Hz. EGS4 calculations predict that this would result in 280 photons entering the SLD drift chamber per pulse. This corresponds to about 0.5% wire occupancy. The major component of SLD background from quadrupoles in the model is due to photons striking the same mask. There are 1.5×10^8 photons per pulse with a critical energy of 1 MeV. This is about 5×10^{10} keV per pulse and, according to EGS4, results in 73 photons in the drift chamber. However, this background, due to the tail, is very sensitive to the model, and can be made considerably larger by varying the assumptions about the IP divergence angle and the beam collimation, thereby introducing new sources. We believe that variations in these parameters contribute to the fluctions in backgrounds actually observed. The energy spectra for the NLC calculation contain far more SR energy than calculated for the SLD case. The spectra in Figure 12-11 correspond to a critical energy of 200 keV and contain between 1.2×10^{11} keV and 1.5×10^{11} keV per train of 90 bunches.

Preliminary calculations of NLC synchrotron radiation have been performed for the 1-TeV machine with 500-GeV beams using the currently available final focus optics [Helm, TLCFF28]. The QSRAD code has been enhanced to include bend magnets in addition to quadrupoles, and to allow SR from individual magnets to be turned on or off. Optical elements within 300 m of the IP have been included in the calculation. SR from more distant magnets cannot reach the IP region directly because of the bends centered at 183 and 261 m from the IP. The beam at the IP was taken to be 226 nm $\times$ 3.57 nm with angular divergence 22.6 μr $\times$ 28.6 μr, corresponding to the 1-TeV parameter set A of Table 12-1. Calculations were done for a single bunch containing 9.5×10^9 electrons, and, unless otherwise indicated, all results quoted here are for a single bunch. Collimation was modeled as perfect rectangular collimators at $7\sigma_x$ and $35\sigma_y$. SR from beam tails was calculated for an arbitrary 1% flat tail, 9.5×10^7 electrons per bunch uniformly distributed over the collimator aperture, and the reader can scale this up or down as desired. In reality, detector elements will integrate backgrounds over a bunch train, and the single-bunch numbers should be multiplied by 90 bunches per train for background considerations.

The first studies of SR in the NLC final focus lead to the relocation of the final bends and other modifications. In the current design the final bend ends 151 m from the IP, apertures in the final focus have been increased, and the quadrupole closest to the IP (QFTA), with the smallest aperture, is shorter than in the previous lattice.

Our experience with EGS4 simulations of SLD backgrounds and initial calculations of SR at the NLC indicate that the most serious source of SR background in this model of the NLC is likely to be the SR photons incident on the inner surface of the beam pipe through QFTA, the magnet closest to the IP. The calculations reported here are for this source only. We consider here the inside of a cylindrical surface of length 1.8 m, between 2.0 and 3.8 m from the IP. This includes the interior of QFTA and a mask that is likely to precede the magnet. The calculation is done for several quad apertures, as discussed below.

The Gaussian beam results in 1.2×10^7 photons above 10 keV energy incident on the inside of a cylindrical QFTA of 4.5-mm radius. These photons are all from the final "soft" bend (SBSR), and have a critical energy of 200 keV. The number of incident photons changes by only 10% if the radius of the magnet aperture is changed by 0.5 mm. The energy spectra are shown in Figure 12-11. These numbers are independent of assumptions about the beam, other than parameters that are well defined. Furthermore, the interior of QFTA can be shielded from the SBSR by a mask in the horizontal plane 40 m from the IP. The half-aperture of this mask would be 3 mm, about $10\sigma_x$, whereas the beam is collimated at $7\sigma_x$. Although the mask is required on only one side of the beam, it would be symmetric to minimize wakefield effects, which must still be evaluated before this mask is accepted as part of the design.

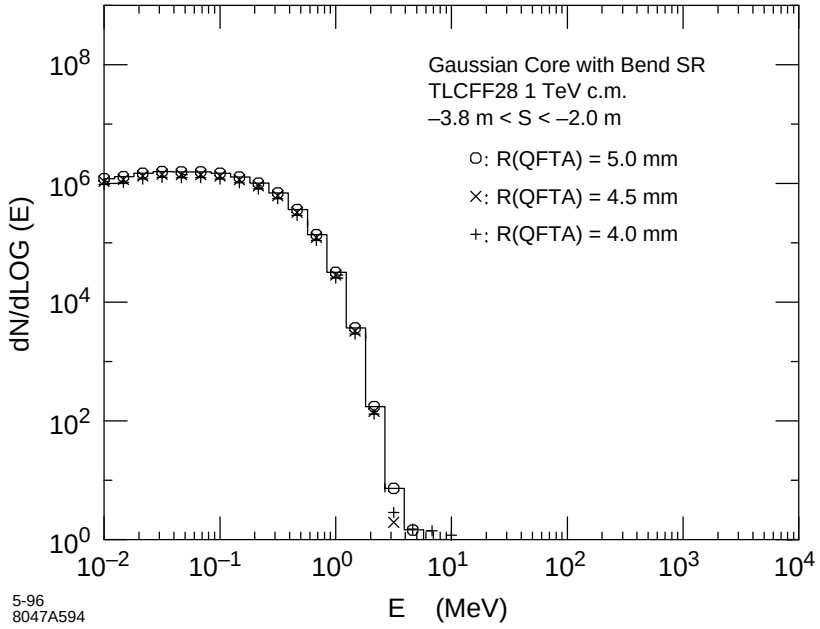


Figure 12-11. *Energy distribution of synchrotron radiation photons, from particles in the Gaussian core of the beam, striking the inner aperture of the QFTA quadrupole. There are only minor variations in the range of radii considered.*

The effect of the tail is calculated without the SR from the bends, because that has been included with the Gaussian core. Results for the flat tail are very dependent on assumptions about collimation and the population of the beam tail. With the assumptions above, the 1% tail calculation results in 3.7×10^5 photons per bunch with an average energy of 5.3 MeV incident inside QFTA if the aperture is 4.5 mm in radius. This SR is all from QFT5 and QFT6, 134 m and 150 m from the IP.

There is also a large flux of 5.1×10^7 SR photons per bunch, incident where the aperture decreases between QFT1 and QFTA. Calculation shows that this can be reduced by an order of magnitude, shifting the flux upstream to a protection collimator 10 m from the IP. This SR is primarily in the vertical plane, and 90% can be intercepted by a 5-mm mask, which is large compared to the 2-mm $35\sigma_y$ beam envelope at this location.

If the radius of QFTA is increased to 5.0 mm, there is a small increase in the number and average energy of the incident photons. The energy spectra are shown in Figure 12-12. However, Figure 12-12 also shows that reduction of the radius to 4.0 mm results in a large increase in both the number and energy of the incident photons. The total SR energy incident on the inner bore of QFTA increases by more than two orders of magnitude. This is because the smaller aperture intercepts a large flux of energetic QSR generated by the beam tail in the superconducting quadrupoles QFT1 and QFT2. This must be considered as the design evolves, because the only way to compensate for a smaller QFTA aperture is with tighter vertical collimation. In this lattice the collumator aperture must be reduced by one σ_y to compensate for each 125-μm reduction in QFTA radius.

The exit aperture to the dump line across the IP will have to be larger than QFTA in order to avoid a large flux of QSR incident on the face of the first dump line quad. The non-zero beam crossing angle makes this possible. For a 4.5-mm radius QFTA the first dump line quad aperture should have a 7-mm radius. Figure 12-13 shows the $7\sigma_x$ and $35\sigma_y$ beam envelopes. In the figure the optical elements are subdivided and the SR fan produced by each element traced to the IP. It is clear that the SR fan in the vertical plane from the final doublet is what sets the exit aperture. The radiation that does hit QFTA is seen to come from Q6, approximately 150-m upstream of the IP.

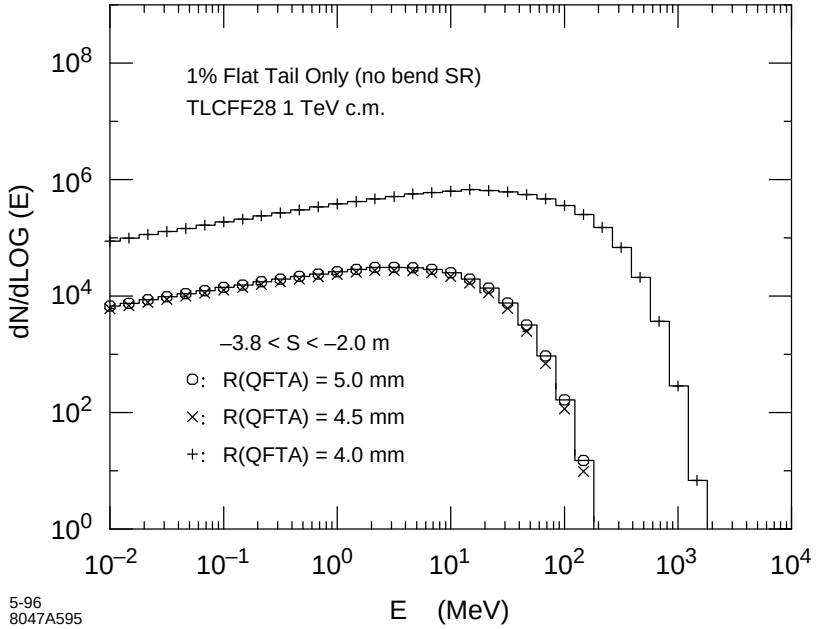


Figure 12-12. *Energy distribution of synchrotron radiation photons, from particles in the assumed 1% flat tail of the beam, striking the inner aperture of QFTA, the innermost quad. Photons produced by bends and photons produced by the Gaussian beam core are not included in this plot. The spectra for 5.0-mm and 4.5-mm radius do not differ significantly. When the aperture is small enough to intercept QSR from the superconducting quadrupoles the incident QSR is much larger, as shown for a 4.0-mm radius.*

These studies should continue in conjunction with design of the final focus and the dump line. More detail must be included in the models of the apertures to identify those sources which are most likely to result in detector backgrounds. It is not obvious that the 1% flat tail assumption, which seems to work at the SLC, is reasonable for the the NLC. Because of its importance in estimating detector backgrounds, an effort should be made to estimate the likely beam halo at NLC.

The results of these SR calculations are used as input to EGS4 and GEANT models of a detector and its masking, and the first results are described later in this chapter. This study is far from complete, and detailed results from the EGS calculation will be used to propose further modification of the lattice and magnet apertures. The severity of the backgrounds modeled here may influence the choice of detector technology. The machine design has been able to incorporate suggestions motivated by the background calculations, and considerable progress has been made in reducing the potential for backgrounds due to synchrotron radiation.

12.3.4 Muon Backgrounds

The same collimators which limit beam phase space and protect the detector from SR backgrounds produce considerable numbers of muons as the incoming beam interacts with them, primarily through the Bethe-Heitler process $e^{\pm}N \rightarrow e^{\pm}\mu^{+}\mu^{-}N$. The problem was first encountered in 1988 with the Mark-II detector at the SLC. To quantitatively study the production and transport of the muons the program MUCARLO [Feldman] was written. It successfully reproduced the experimental results for the number and spatial distributions of muons hitting Mark-II. The backgrounds were reduced to acceptable levels by moving the primary collimation to a point at the end of the linac, approximately

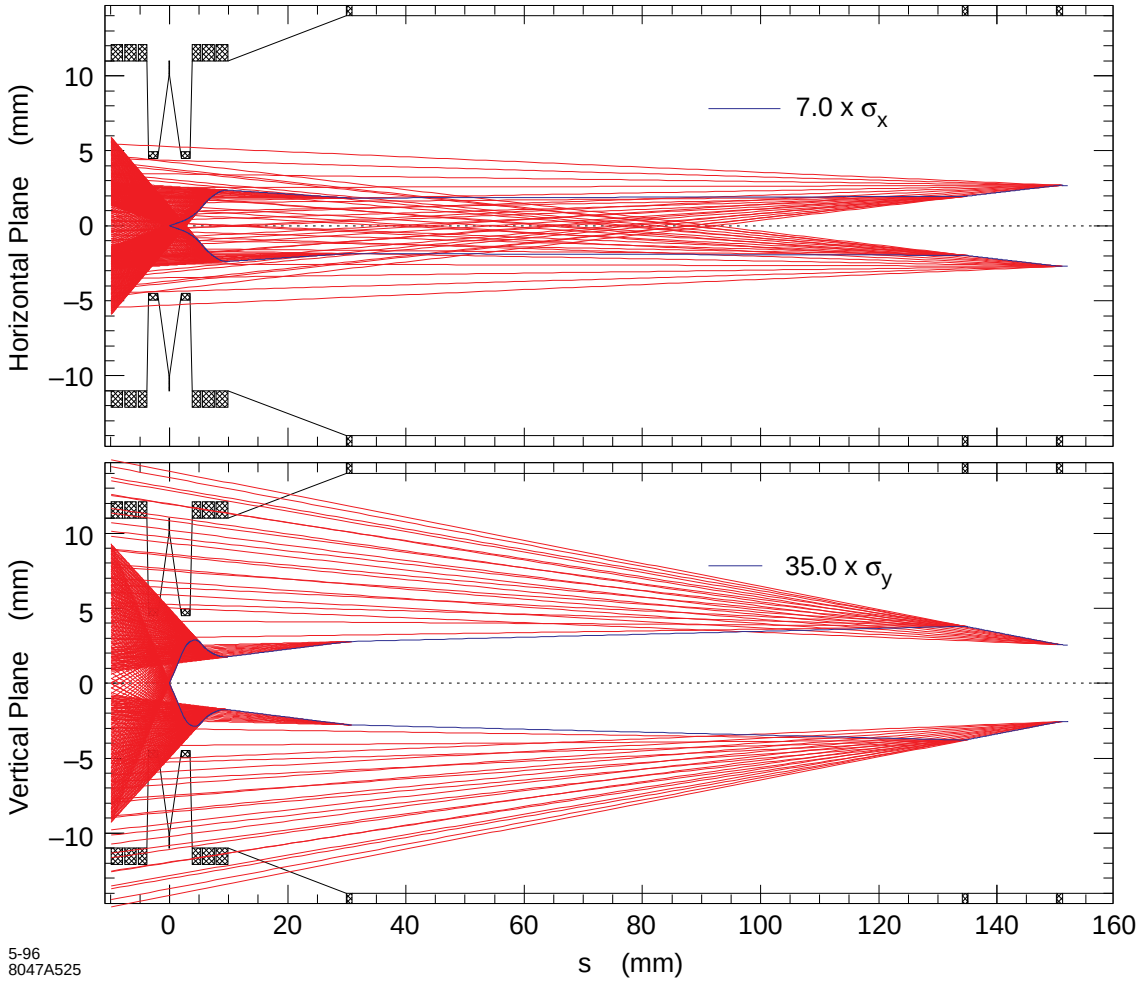


Figure 12-13. *The $7\sigma_x$ and $35\sigma_y$ beam envelopes and the SR fans emitted for the lattice used for these calculations.*

1500-m from the IP and separated from the detector by the SLC arcs, and installing magnetized iron spoilers in the final focus tunnels. At the NLC, this philosophy is duplicated by locating the collimation section at the end of the linac, ~ 5.2-km from the interaction point, and placing the final focus in a tunnel section dug at 10 mr with respect to the linac, the "Big Bend."

In 1990, the MUCARLO program was modified [Keller, 1991] for use with beams of up to 250-GeV energy. Using the TRANSPORT deck for a 500-GeV-c.m. collider available in June 1990 (FFN09, R. Helm), the placement of five toroidal spoiler magnets was optimized to maximize the number of electrons that could hit a collimator before producing one muon that would make it to the detector.

In 1993, MUCARLO was again modified. Beams of energy up to 500 GeV were allowed, and muon production by direct e^+ annihilation, $e^+e^- \rightarrow \mu^+\mu^-$, and photopion production, $\gamma A \rightarrow X\pi(\pi \rightarrow \mu\nu)$, were included. Using the TRANSPORT deck for a 1-TeV-c.m. collider available in June 1992 (TLCFFN5, R. Helm), the muon background study was repeated [Keller]. Most of the figures presented in this section are from that study. Figure 12-14 shows a schematic of the collimation and final focus beam transport sections being considered at that time. The collimation section has a series of six Hi-Z collimators and a total bend of 6.14 mr followed by a big bend with a total bend of

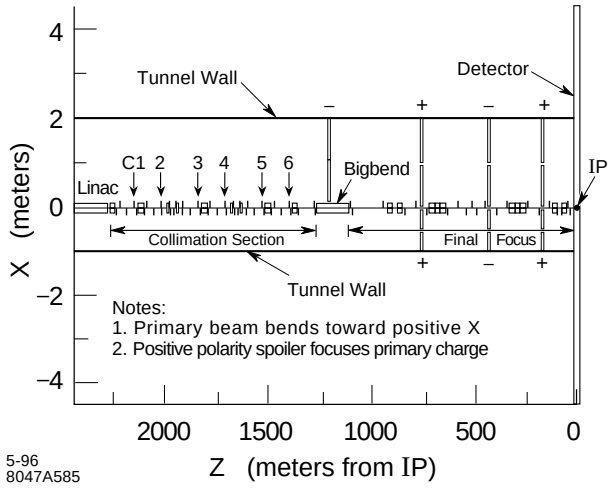


Figure 12-14. *Plan view of the 1992 NLC beam line in the tunnel with magnetized iron spoilers.*

10 mr followed by the final focus with a reverse bend chromatic correction section of ± 2.21-mr bends. The model includes a 3.05-m-square cross section concrete tunnel through sandstone, concrete support girders under the beam elements, and dipoles and quadrupoles which include return flux in the iron and pole tips. The detector is assumed to have a 4.5-m-radius cross section centered on the IP. There is a series of magnetized iron spoilers of alternating polarity distributed through the final focus.

Figure 12-15 shows how the spoilers are arranged in the tunnel at a given location. The field in the iron was modeled using the two-dimensional program POISSON.[7] Each spoiler is 9.1-m long with a winding slot width and height of 3.2 cm and 126 cm, respectively. Each set of "tunnel-filler" spoilers weighs 750 tons and would cost \$2–3 million installed. For magnetized iron with a field of 16 kg, the ratio of bend angle to scattering angle is $\theta_{\rm bend}/\theta_{\rm MCS} \approx 3\sqrt{L}$, where L is the length of the spoiler in meters.

An alternative to large tunnel-filler magnetized iron spoilers has been proposed, consisting of nested iron cylinders with opposite-polarity azimuthal magnetic fields as shown in Figure 12-16. The idea is that the nested cylinders are located downstream from each muon source and are long enough to either range out muons or cause enough energy loss so that the muon is unlikely to reach the detector. A version of the nested cylinder idea was tried in MUCARLO, and the results are presented in Figure 12-17.

After the muon exits the source, the Monte Carlo program swims it in 30-cm steps through the tunnel. When material is encountered, the muon scatters, loses energy, and bends (if a magnetic field is present). The trajectory of each muon is followed until the muon either stops or reaches the IP. For the purposes of this study, the details of the detector, *e.g.*, trackers, calorimetors, and muon walls, are not included. A muon which reaches the IP within a radius of 4.5 m from the beam line is counted as a detector hit.

Figure 12-17 shows the number of beam particles which must hit a collimator to produce one muon in the detector as a function of source location in the beam line. As a worst case, an e^+ beam was chosen for this study so that direct annihilation production could be included. The source points include the six Hi-Z collimators in the collimation section and collimators at four high-beta points in the final focus, which are potential scrapers of beam-gas coulomb scattering or beam-gas bremsstrahlung.

From the curve labeled "No spoilers" it is seen that a beam loss of 10^{10} per bunch train anywhere in the collimation section will result in a muon background which is many orders of magnitude away from the design goal. The curve

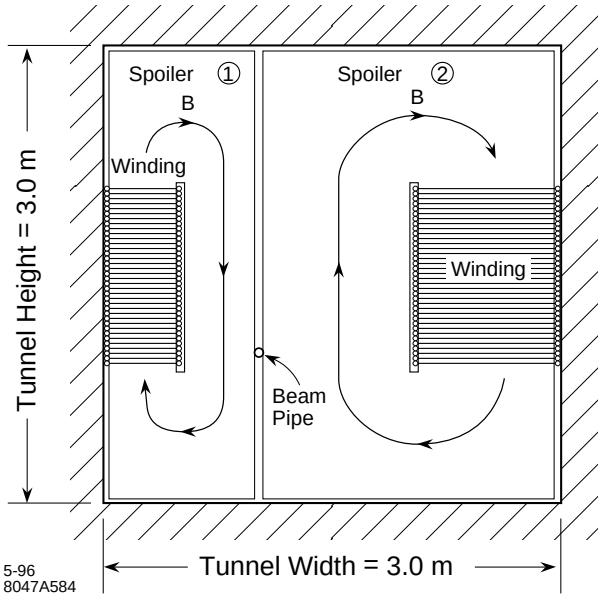


Figure 12-15. *Two side-by-side magnetized iron spoilers filling a* 3×3*-m beam tunnel.*

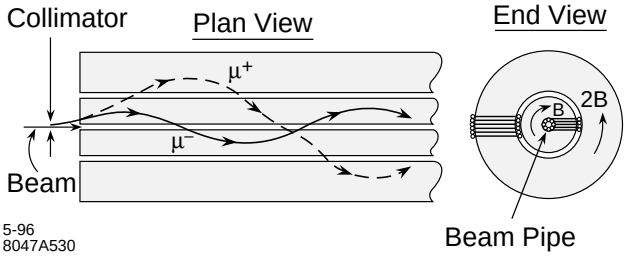


Figure 12-16. *Nested, magnetized iron cylinders of opposite polarity to channel both* μ^+ *and* μ^-*.*

labeled "Magnetized spoilers fill tunnel" in Figure 12-17 is the result of adding magnetized iron spoilers which fill the tunnel at three locations in the final focus and one magnetized iron piece in the tunnel aisle next to the 10 mr big bend. It is seen that for all six collimators in the collimation section, more than 10^{10} beam particles must be lost to produce one muon in the detector. This satisfies the design goal of allowing a 1% continuous beam loss in the collimation section.

The curve labeled "Magnetized cylinders" in Figure 12-17 shows the results of filling all drift spaces in the collimation section with magnetized iron cylinders described above. The result is considerably worse than for magnetized iron spoilers which fill the tunnel and does not meet the design goal. This is because the magnetized cylinders must be interrupted by beam elements, especially dipoles in the chromatic correction sections, which disperse muons away from the cylinders and therefore disrupt the channeling orbits.

For the case of the magnetized iron tunnel-fillers, Figure 12-18 shows histograms of muon-production momenta and final momenta for muons which hit the detector from the six sources in the collimation section. It is seen that the initial muon momentum must be greater than 300 GeV/c to hit the detector from sources outboard of the big bend.

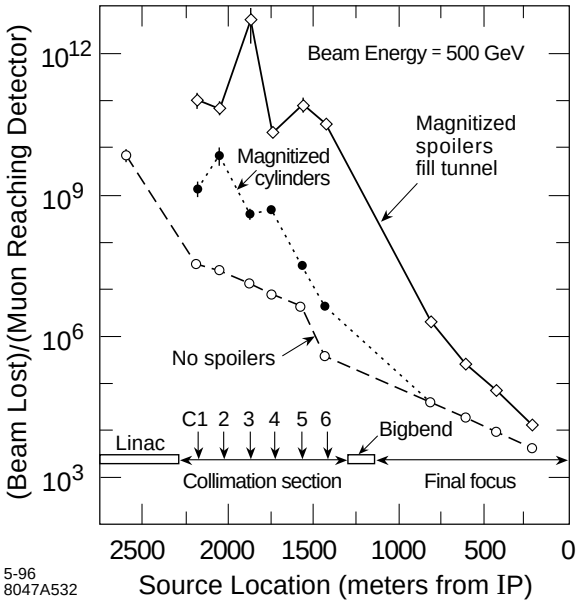


Figure 12-17. *Results of muon Monte Carlo calculations, based on the 2-km final-focus design of 1992, for three conditions: no spoilers, and two types of magnetized iron spoilers.*

Figure 12-19 shows the spatial distribution of muons from the collimation section which reach the IP. It is seen that increasing the detector dimensions by 2–3 m in the horizontal direction would intercept significantly more muons.

For sources in the collimation section of a 1-TeV center-of-mass linear collider, a system of magnetized iron spoilers which fills the tunnel and meets the design goal of allowing a continuous 1% beam loss, or 10^{10} beam particles per bunch train, was found. This is more than a three orders of magnitude improvement over the case with no muon spoilers. For the case of magnetized iron cylinders which fill all drift spaces in the collimation section, the design goal was not met.

These studies were recently repeated using the current design for the collimation section, big bend, and final focus (TLCBD01B, R Helm). While the tunnel size, magnet design, and magnet support are important inputs to the program, we still assume the original 3×3-m-square tunnel cross section and SLC-like magnet and support. Figure 12-20 shows the results for two configurations: no spoilers and with tunnel filling magnetized spoilers. It can be seen that the newer design, with its longer beamline, allows for a larger number of beam particles to be lost per muon arriving in the experimental hall. Nonetheless, muon spoilers are still required to attain the design goal. The spoiler locations used here were determined by scaling those of the older, shorter final focus design. Additional optimization of the spoiler system is in progress.

12.4 Detector Issues

There has been much less thought devoted to the NLC detector than to the accelerator itself. In physics studies conducted in Japan and Europe, the detector is envisioned to follow the solenoidal architecture typified by OPAL, ALEPH, or SLD, scaled appropriately for the increase in center-of-mass energy and having, depending on the study,

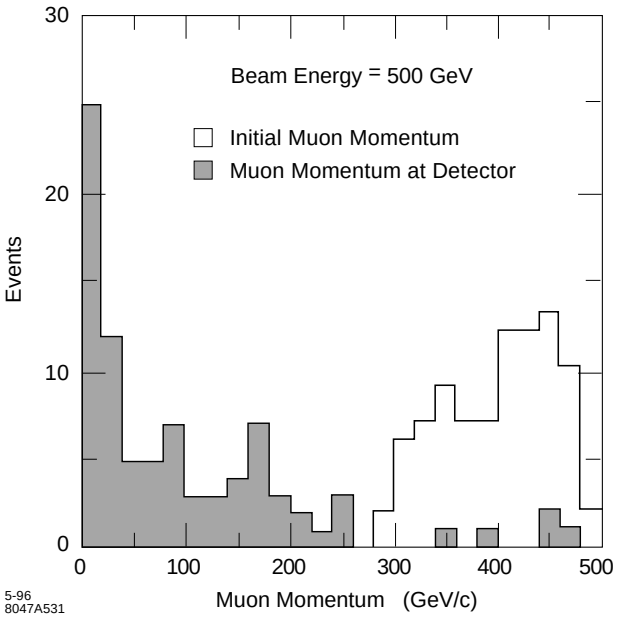


Figure 12-18. *Momentum distribution of muons which hit the detector from six sources in the collimation section. The normalization is arbitrary.*

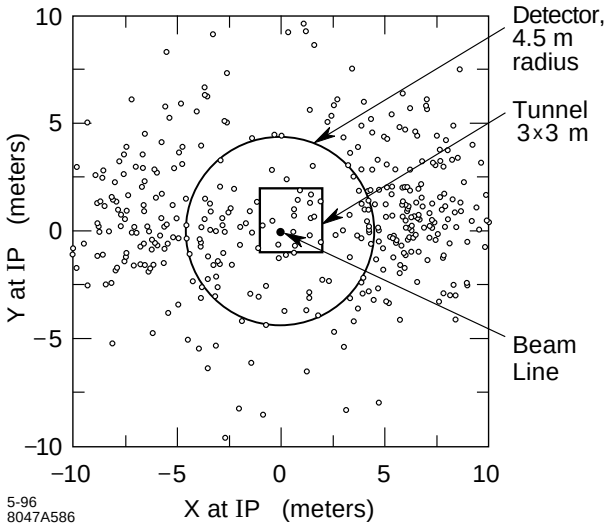


Figure 12-19. *Spatial distribution of muons which reach the IP from six sources in the collimation section. The normalization is arbitrary.*

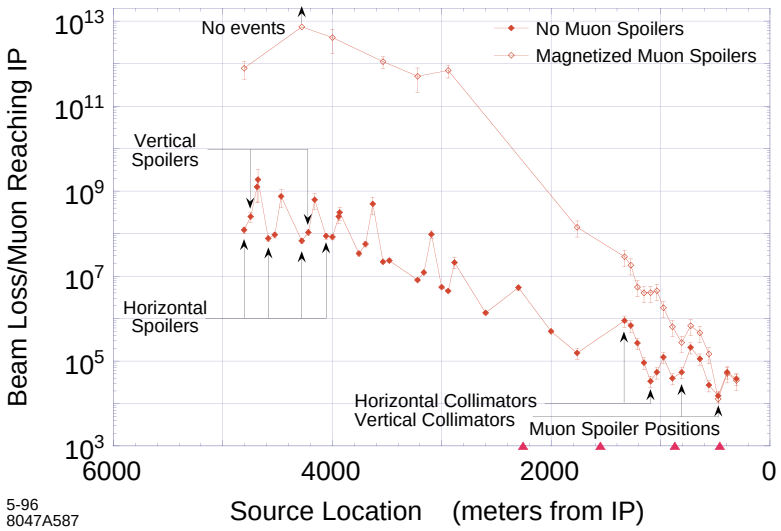


Figure 12-20. *Results of muon Monte Carlo calculations updated for the collimation and final focus design available in early 1996. Data are presented for calculations with and without a muon spoiler system. The data from the previous calculation (Figure 12-17) are also plotted for reference.*

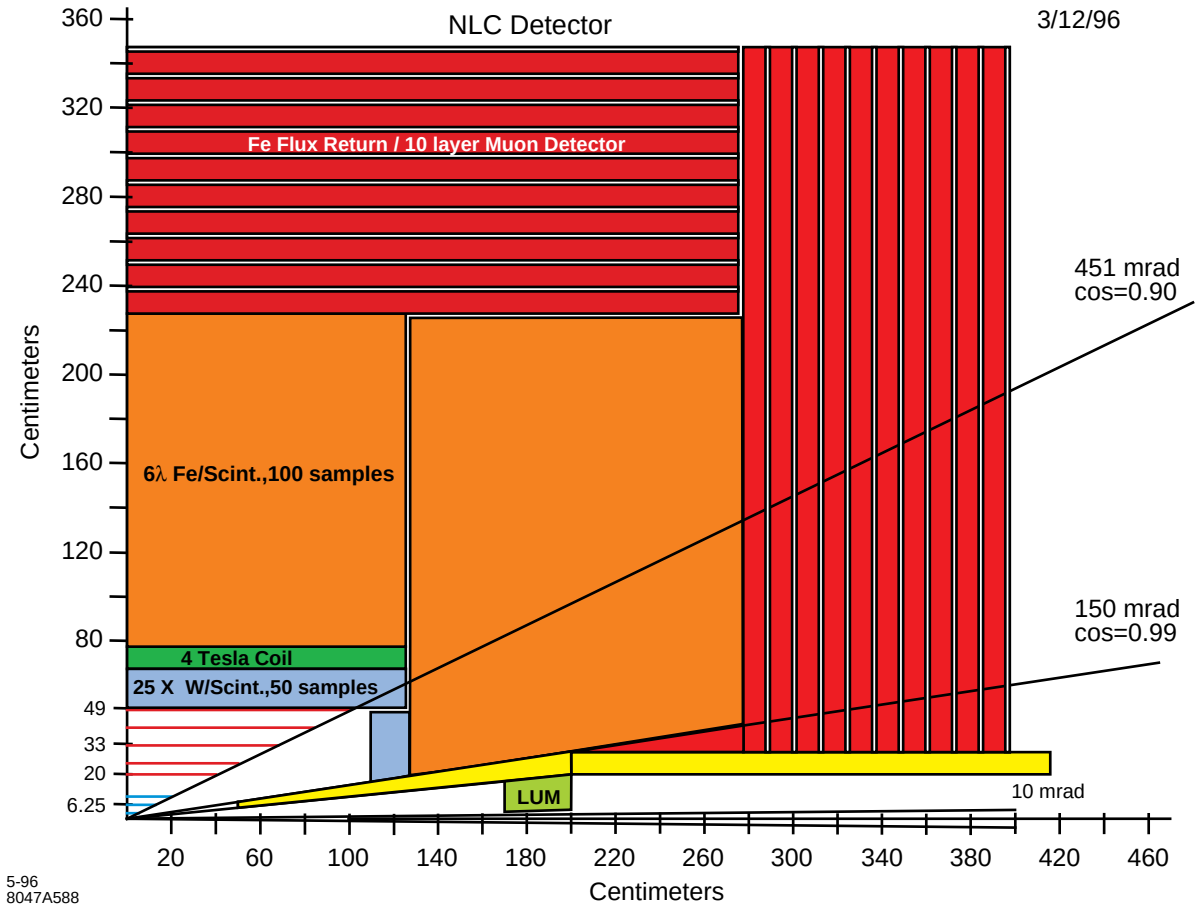


Figure 12-21. *A compact NLC Detector, robust against backgrounds. The scale is in cm.*

improved tracking, vertexing, or calorimetry. Table 12-5 gives a list of possible physics-motivated specifications as summarized in a recent review.[TRC 1995]

Two issues appear to drive the conceptual design of the detector: the required momentum resolution of the charged particle tracking system and the requirements for "special" support systems for controlling vibrations of the final focus. A pixel vertex detector (presumably CCDs), a moderate resolution electromagnetic and hadronic calorimeter, and a muon identifier of instrumented steel plates are non-controversial.

Recently, a "straw-man detector" design has been presented [SLAC 1996] which is based on a compact five-layer tracking system using double-sided 5-μm-resolution silicon strips in 4T magnetic field. Such a system would have a momentum resolution around $0.03\% \times \mathrm{P(\,GeV)}$, and would be resistant to the expected backgrounds and instabilities of the NLC. By keeping the coil inside the electromagnetic calorimeter it remains relatively thin (2–3 radiation lengths) and short (2–3-m in length). Figure 12-21 shows a quadrant this detector.

The nominal nanometer vertical size of the NLC focus drives the question of the quadrupole support system. The design possibilities include a rolling support of the quadrupole from a retractable end door such as in the SLD; a static steel end door supporting the quadrupole with retractable muon system and calorimeter for interior access; and actual intrusion of a static steel and concrete buttress into the door region for quadrupole support. Measurements

	LEP/SLC-Style	ee500 1991	ee500 Typical	JLC Detector	1 TeV Detector	Units
Tracking						
$\frac{\delta p_t}{p_t{}^2} = C$						
$C =$	8	5-100	10	1	2	$\times 10^{-4}$ gev/c^{-1}
E-M Calorimeter						
$\frac{\delta E}{\sqrt{E}} =$	0.2	0.02-0.15	0.1	0.15	0.10	$\sqrt{\text{GeV}}$
Hadronic Calorimeter						
$\frac{\delta E}{\sqrt{E}} =$	0.9	0.3-1.0	0.8	0.40	0.65	$\sqrt{\text{GeV}}$
Energy Flow						
$\frac{\delta E}{\sqrt{E}} =$	0.65	0.3-0.8	0.5	0.3	0.4	$\sqrt{\text{GeV}}$
Vertexing						
$\delta(IP) = A \bigoplus \frac{B}{p}$						
$A =$	25	5-20	10	11	10	μm
$B =$	100	50-100	50	28	50	μm GeV/c
Hermetic coverage						
$\lvert\cos\theta\rvert <$	0.96	0.70-0.99	0.95	0.98	0.98	

Table 12-5. *Examples of detector performances used in physics studies.*

of the differential motion of the quadrupole end of SLD have been made (see Section 12.4.5). It appears that, with appropriate care, seismic concerns will not overly compromise the detector design.

12.4.1 Effect of Backgrounds on the Detector

The zeroth-order masking design for the IP is shown in Figure 12-22. This geometry has been programmed into EGS4, and full simulations of the effects of the interactions of beamstrahlung-produced pairs, QSR-produced photons, and lost beam particles are in progress. Tungsten is used for Mask M1, M2 and the luminosity monitor. The beam pipe made of 750-μm-thick beryllium is located at r = 1 cm near the IP and of 500-μm-thick stainless steel for the rest. The beam pipe is pulled back to a larger radius as soon as it is past the maximum z required by the vertex detector. The rf shield made of 200-μm-thick copper is necessary to avoid the wakefield effect. The solenoid field in the detector and the quadrupole field in the final quads are included, and charged particles are transported properly in the magnetic field. However, non-axial field components of the solenoid or fringing field of the quads are not considered. The vertex detector is modeled with four layers of sillicon+beryllium at the radii of 1.2, 2.7, 3.8, and 4.8 cm with an angular acceptance of $|\cos\theta| < 0.9$. The layer thickness and the locations of the outer three layers are chosen after the new SLD vertex detector. In addition to the vertex detector layers, there are massless scoring planes at r = 10, 15, 20, 25, and 30 cm with $|\cos\theta| < 0.9$. The tracking chamber is located at r = 30 cm and $|z| < 100$ cm.

The complete history of secondary electrons, positrons, and photons produced from the incident e^+e^- pairs and synchrotron photons was traced using EGS, and the background level in the detector was calculated.

The figure of merit we have traditionally used for the maximum allowable background level is one charged particle hit/mm^2/train-crossing for a pixel-based vertex detector, and 10,000 photons/train-crossing for a drift chamber.

This work has resulted in changes to the final-focus lattice and to the design of the final quadrupole doublet. The process has been iterative and is still in progress; the reader is warned not to assume that we have presented a unified and consistent set of results throughout the relevant sections of the design report.

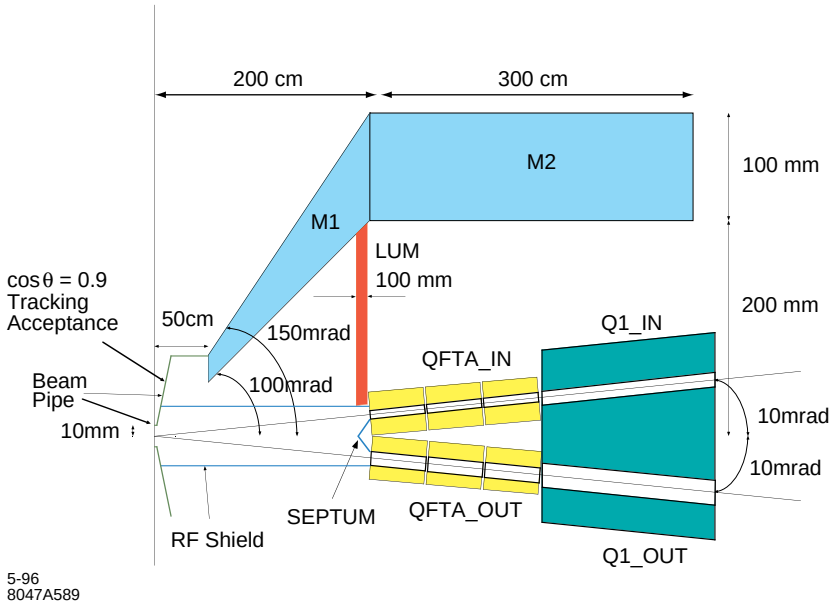


Figure 12-22. *First pass at NLC masking:*

M1 Mask	Tapered Tungsten Cone beginning at z=0.5 m and ending at z=2.0 m, with inner and outer angles of 100 and 150 mr, respectively.
M2 Mask	Tungsten annulus with 10-cm wall and inner radius 20 cm, $2.0 < z < 5.0$ m.
QFTA_IN/OUT	Incoming/Outgoing Sm2Co17 FF quads rotated 10 mr in x–z plane. QFTA_IN - Inner/Outer radii = 4.5/20.0 mm QFTA_OUT - Inner/Outer radii = 7.5/20.0 mm Longitudinally divided into three 5.0-cm-long segments with 2.0-cm gaps. L^* = line at 10 mr to IP is 2.0-m long.
Q1_IN/OUT	Superconducting magnet with Q1_IN aperture = 5 mm; Q1_OUT aperture = 8 mm. Extends from 3.5 to 5.0 m.
Beam Pipe	750-μm Be beam pipe with 100-μm Titanium liner at 1.0 cm, at $z \pm 2.1$ cm, which is joined onto a 500-μm Stainless sectioned flared at 451 mr until $r = 7.56$ cm, after which it proceeds to the M1 mask and follows its inner contour.
RF Shield	200-μm Cu extension at $r = 2.75$ cm, $z = 2.1$ cm, extending at constant radius to $z = 165$ cm, then proceeding via two "legs" to join electroplated inner radius of in-/out-going quads.
SEPTUM	200-μm Cu cone beginning at $z = 189$ cm and proceeding to $z = 2.0$ m at an angle θ, where $\theta = \arctan((20\,\text{mm} - 7.5\,\text{mm})/11\,\text{cm})$.
LUM	10-cm Tungsten with back end at 195 cm, outer radius tapered at 150 mr at M1, inner radius tapered at 10 mr at radius of outgoing quad.

Beamstrahlung-produced pairs

The ABEL simulations of the beam-beam interaction have been described in the final-focus chapter (Chapter 11) of this report. We have concentrated on the data set which describes the 1-TeV NLC parameter set A machine design listed in Table 12-1. The simulations used a 10-MeV cut on pair-member energy. This cut is responsible for the hard edge at the lower right corner of Figure 12-10. As the bremsstrahlung probability for high-Z materials is not negligible at 10 MeV, the cutoff energy should be lowered to 1 MeV in ABEL. The effect of lowering this cut is under study.

Particles with a large p_t will hit the vertex detector directly, while particles with a small p_t will traverse the detector following the axial magnetic field and hit the final quad face producing secondary charged particles which come back to the IP and hit the vertex detector. Figure 12-23 shows the average hit densities of e^+e^- per train at r = 1 and 2 cm as a function of z for two solenoid field strengths. The solid histogram represents the particles directly hitting the layer, while the dashed histogram is for the backscattered particles. Since the interaction of e^+e^- at a few MeV is very dependent on the detailed geometry and materials, the calculation presented in Figure 12-23 used massless scoring planes at r=1 and 2 cm. Since no interactions were simulated at these layers, one particle could contribute multiple hits. Therefore, the numbers should be considerd as an upper-limit. With this caveat, the hit density of directly hitting particles is about 2–3/mm^2/train at r = 1 cm and less than 1/mm^2/train at r = 2 cm; an acceptable level for a pixel-based vertex detector. The hit density of backscattered particles is, however, about ten times higher. Low Z coatings on the quadrupole and luminosity monitor faces will substantially reduce the flux of soft backscattered particles. Improvements to the IR layout and masking to reduce the backscattering probability are underway.

This study has tried to find the minimum feasible radius for a vertex detector. The reader must remember that at $r = 3$ cm, the direct pair background is essentially zero in a 4-T field. Furthermore, at 500 GeV-center-of-mass, backgrounds are down by another factor of 4.

Figure 12-24 shows the number of photons per train striking a scoring plane at $r = 30$ cm, the nominal position of the inner wall of a central tracking chamber as a function of the solenoidal field of the detector. The number of electrons crossing the scoring plane is negligible. Since the majority of photons are produced by the interactions of pairs at the Mask M1 face and at the rf shield, the background can be reduced substantially by using higher solenidal field strength so that more pairs are curled up within the deadcone. However, even at 4 T, the total number of photons is 80K/train, which may not be tolerable for a drift chamber.

Figure 12-25 shows the number of photons per mm^2 per train striking the vertex detector layers and scoring planes from $r = 1.2$ cm to $r = 30$ cm for three solenoid fields of 2, 3, and 4 T. These hit densities can be used to optimize the position of a vertex detector or intermediate tracker. As Figure 12-25 shows, 10/ mm^2 as the maximum *photon* hit density per train, *i.e.*, before conversion, the background would seem to be completely manageable for a pixel-based vertex detector.

At this point, the lesson we can draw from these studies is that having a beam-pipe radius less than 2 cm will require a field strength greater than 2 T. Most of the photons produced come from the interaction of the pairs with the beam pipe and from backscattering from the surfaces which are the closest in z to the IP, namely the luminosity monitor at $z = 190$ cm, the front face of the quadrupole doublet at 2 m, and the septum of the rf shield, currently at 189 cm. The IR design should pull back the the beam pipe to a larger radius as soon as it is past the maximum z required by the vertex detector, limited only by wakefield considerations to have a smooth transition from the FF quadrupole inner aperture to the beam pipe. If, when all residual questions as to the normalization of the simulation are understood, the effect of the photon hits in the central tracker are supportable, we may be able to decrease the distance between the IP and the quadrupole face, which would make the FF optical design much easier. At present the conical tungsten mask M1 between 100 and 150 mr seems completely adequate to shield against pair induced backgrounds. The long tungsten skirt around the quadrupole pair, mask M2, appears necessary in the z area around the quadrupole face. It does not appear to be necessary that M2 neither be 3-m long, nor 10-cm thick.

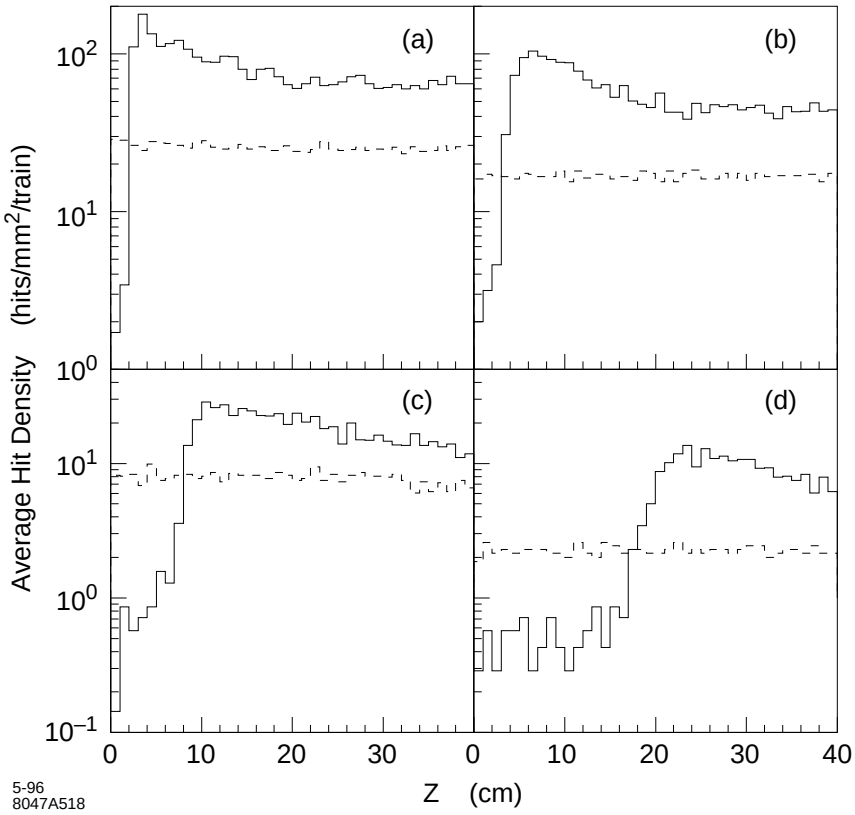


Figure 12-23. *The electron pair hit density per mm^2 per train of 90 bunches as a function of z. The solid line counts hits in scoring planes as the pairs leave the IP; the dashed line counts hits in the scoring planes resulting from the low p_t particles that have gone forward and hit the quad faces, septum, and luminosity monitor. The four parts correspond to: a) $r = 1$ cm and $B = 2T$, b) $r = 1$ cm and $B = 4T$, c) $r = 2$ cm and $B = 2T$, and d) $r = 2$ cm and $B = 4T$. The result is an EGS4 calculation using as input the ABEL beam-beam simulation at a center-of-mass energy of 1 TeV. In each case the scoring plane has an angular acceptance corresponding to $\cos\theta = 0.90$.*

Synchrotron Radiation Backgrounds

Section 12.3.3 discusses in detail the photon flux incident on the inner aperture of QFTA for the various lattice and final quadrupole options under discussion. Since, as described there, the synchrotron radiation from the Gaussian beam can be reduced to a negligible level by using an upbeam collimator, the detector backgrounds were calculated only for the synchrotron photons from the 1% flat-tail. The photon flux and energy spectrum of the tail described in Section 12.3.3 were used as an input for the EGS simulation.

Most of the photons incident on the inner surface of QFTA are absorbed in the quad, and particles penetrating the magnet can be stopped by Mask M1 and M2. However, the synchrotron photons hitting within about 10 cm of the quad exit produce secondary photons and electrons/positrons which come out of the quad and become a potential source of the detector backgrounds. The number of e^+e^- coming out of QFTA is 400K/train and the number of photons is 2×10^6/train. Those electrons/positrons that come out of the quad traverse the detector following the solenoid field and hit the vertex detector, while those secondary photons with angles between 10 and 30 mr with respect to the beam line interact with the vertex detector and rf shield within $z = \pm50$ cm or hit the downstream M1 face, contributing major backgrounds in the detector. Figure 12-26 shows the number of photon and electron/positron hits per mm^2 per train-

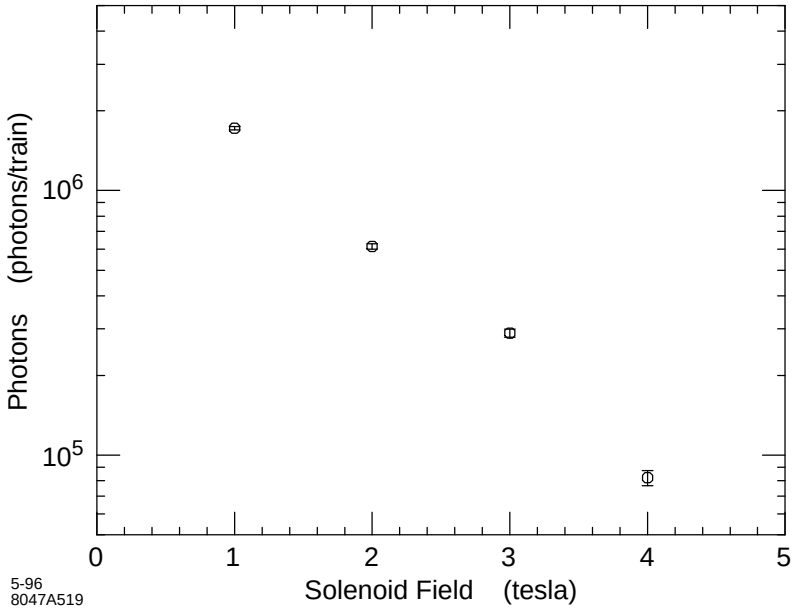


Figure 12-24. *The number of photons that strike a scoring plane at $r = 30$ cm arising from the interaction of e^+e^- pairs produced by the beam-beam interaction. The number of hits is shown as a function of the detector's solenoid field for center-of-mass energy of 1 TeV.*

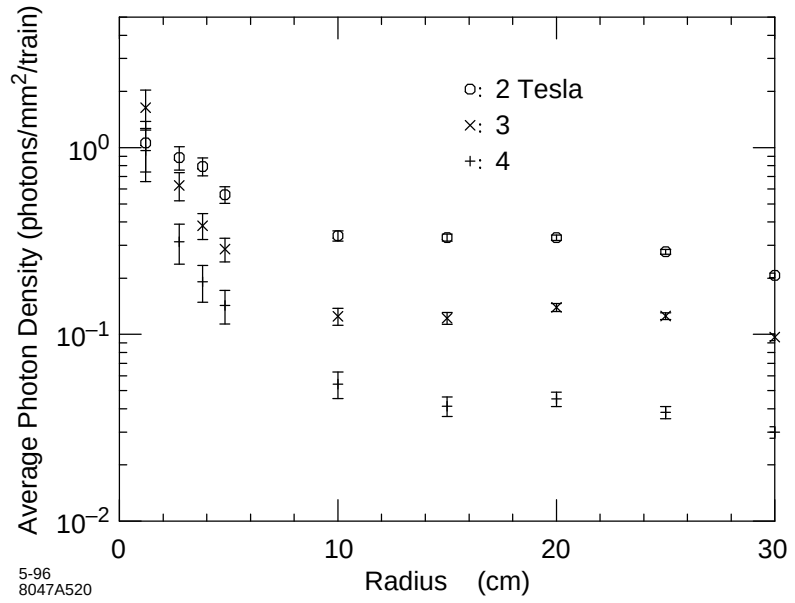


Figure 12-25. *The number of photons per train crossing that strike scoring planes at the indicated radii arising from the interaction of e^+e^- pairs produced by the beam-beam interaction. The number of hits is shown as a function of radius for three solenoid field strengths and 1 TeVcenter-of-mass energy.*

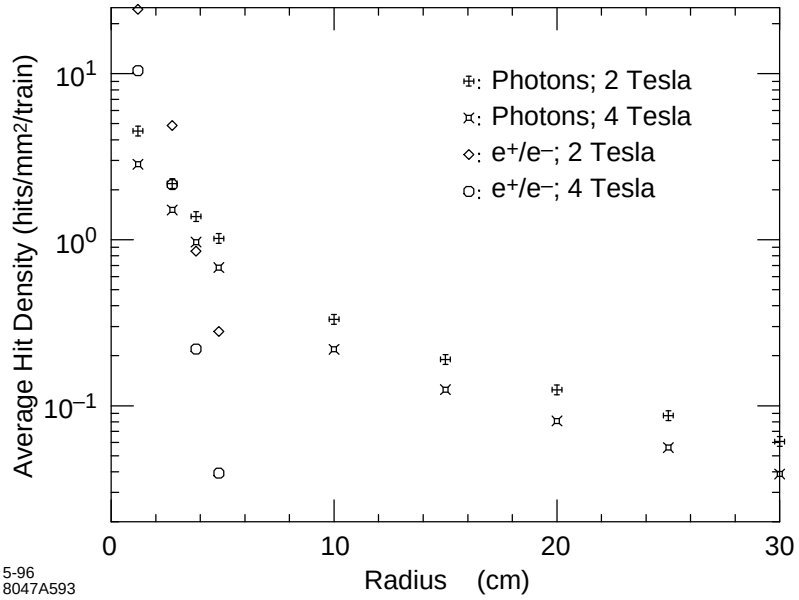


Figure 12-26. *The number of photons and e^+e^- hits per train crossing that strike scoring planes at the indicated radii arising from the interaction of photons in the inner bore of the QFTA quadrupole. These hits are produced by the quadrupole SR in the final focus lattice at 1-TeV-c.m. energy.*

crossing striking the vertex detector layers and massless scoring planes from r = 1.2 cm to r = 30 cm. The simulations were made for 2 and 4-T solenoid fields. At r = 1.2 cm, the e^+e^- hit density reaches as high as 10–20 hits/ mm^2 per train-crossing, and it may be intolerable even for a pixel-based vertex detector. While the photon density in the vertex detector is less than 5 photons/ mm^2 per train-crossing and is tolerable, the total number of photons striking the tracking chamber at $r = 30$ cm and z=±100 cm is 167K for 2 T and 110K for 4 T per train-crossing. The large photon flux may not be tolerable if we want to use a drift chamber.

The preliminary calculation described above has indicated that the SR backgrounds are potentially serious at r = 1 cm even for a pixel-base vertex detector and if a drift chamber is used. However, at 20–25-mm radius, such as that used by SLD at SLC, the situation is very much easier. While the backgrounds from the beam-beam interaction can be reduced by using a higher-strength solenoidal field, as shown in Figure 12-24, the SR backgrounds are not affected in the same way. Rather, the machine collimation and final-focus scheme must be designed so that the photon flux striking the inner surface of QFTA is substantially reduced. Furthermore, at 500 GeV-center-of-mass-energy, the likely startup point for an NLC, the number of SR photons goes down a factor of 2 and the average energy of the photons goes down a factor of 6, reducing the severity of this background source substantially.

Track Reconstruction Limits to the IR design

We will assume for the present that the innermost tracking detector will be a pixel-based device. The SLD has found that its vertex detector, VXD2, effectively composed of two layers of CCDs with 22 μm × 22 μm pixels, is very robust against the backgrounds seen at the SLC, which result in average occupancies at the level of 10^{-4}, or 0.4 hits per mm^2, summed over its readout of 19 beam crossings. SLD is currently installing a new vertex detector, VXD3, consisting of three layers of larger, less massive CCDs with the same pixel size. The extra layer will allow VXD3 to be a self-tracking device. VXD3 serves as an excellent model for an NLC detector inner tracking device.

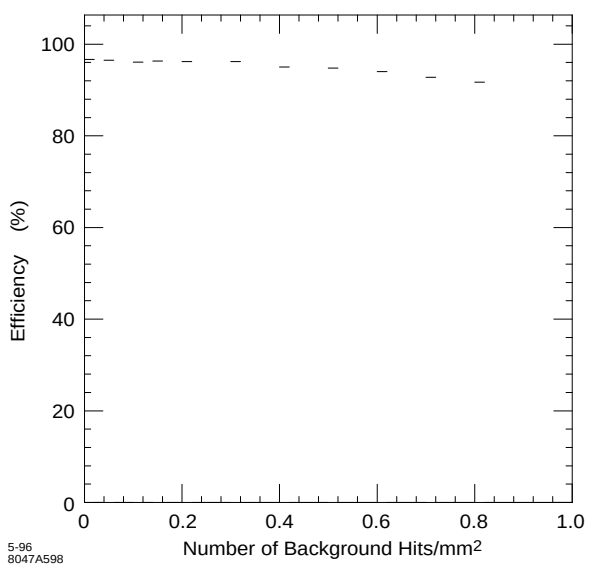


Figure 12-27. *Track efficiency vs. background level.*

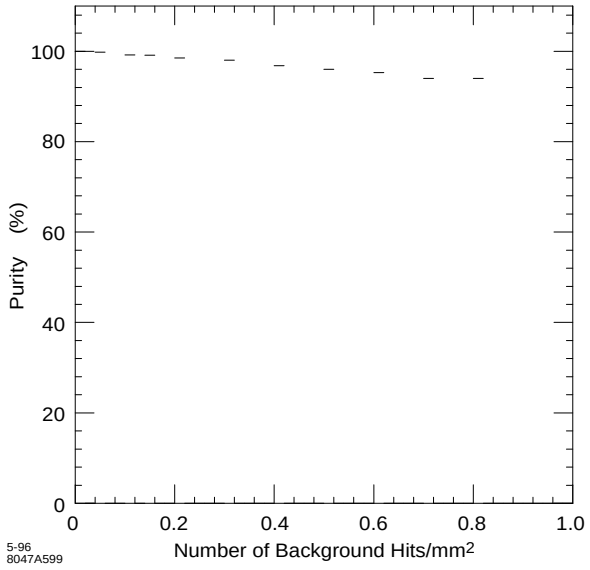


Figure 12-28. *Track purity vs. background level.*

To begin to estimate quantitatively the effects of backgrounds on track reconstruction, we have employed the detailed VXD3 Monte Carlo simulation and the standard SLD track reconstruction code to study the efficiency and purity for reconstructing the correct tracks resulting from hadronic Z^0 decays at resonance as a function of the density of randomly generated background hits. The mean number of tracks in the VXD3 acceptance is approximately 15. The standard SLD track reconstruction code begins with a track found in the SLD central drift chamber. For the purposes of this study, the Monte Carlo tracks were smeared appropriately and track banks formed to simulate a central tracking device. These tracks are then extrapolated to the VXD3 and linked to pixel clusters layer-by-layer using a Kalman Filter technique. Pixel clusters (hits) in the VXD3 resulting from background can affect the reconstruction procedure in two ways: background hits can merge with real hits acting to worsen the resolution, and background hits can cause track mislinks which can pull the track away from the real hits. The purity of the VXD3 linked hits, defined as the ratio of the number of real Monte Carlo hits to the total number of hits linked per track, measures the former effect. The efficiency of linked VXD3 hits, defined as the ratio of the number of hits linked to the number of hits expected to link per track, measures the latter effect. Figures 12-27 and 12-28 show how the efficiency and purity vary as a function of background hit density ranging from 0 hits/mm^2 to 1 hit/mm^2. Both the purity and efficiency drop only a few percent across this range display the robustness of this procedure against random backgrounds in the VXD3. This work is currently being extended to study the ability to tag B mesons resulting from top and Higgs production at 500-GeV-c.m. energies. This study will continue to use VXD3 as a model vertex detector, but will use a field strength of 2 T as well as using typical resolution and efficiency parameters for the NLC detector central tracking device. It is hoped to employ the B mass tagging algorithm which has become successful at the SLD.

12.4.2 Vibration Suppression for the Final Focus Quadrupoles

A schematic of the interaction region is presented in Figure 12-29, and indicates the four quadrupoles nearest the IP for the incoming beam. Also indicated are approximate locations of the detector tracking, calorimetry and muon identification systems, and the detector solenoidal coil. The quadrupole nearest the IP, QA, is a permanent magnet;

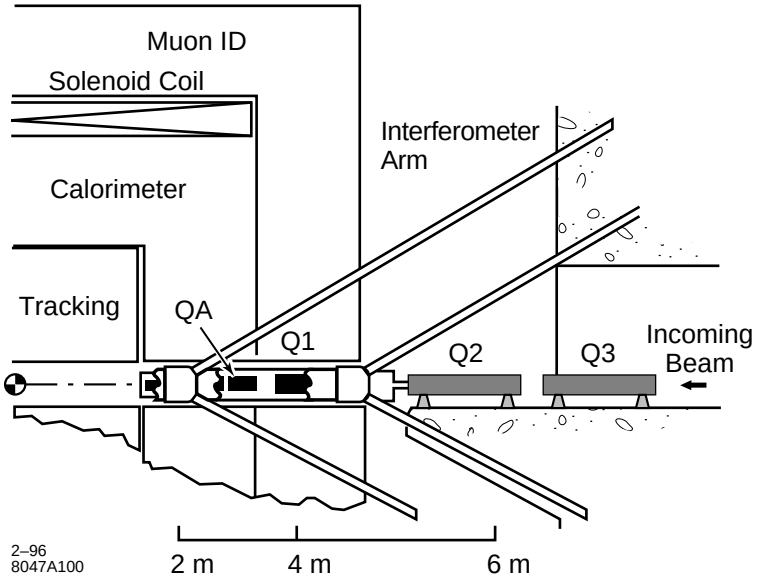


Figure 12-29. *IR final-focus quadrupoles.*

Q1 is a superconducting magnet; and the two Q2 magnets are normal magnets. QA and Q1 are vertically focusing, while the Q2s are horizontally focusing. For a 5-nm vertical spot size at the IP, the tolerances for uncorrelated vertical vibrations of QA and Q1 are at the nm level. The vertical vibration tolerances for the Q2 magnets are about a factor of 3 less severe.

QA is chosen to be a permanent magnet to achieve a transversely compact magnet with the high-focusing gradient field needed. (Note that the separation of the incoming and outgoing beams is only 40 mm at z = 2 m from the IP.) Q1 is superconducting to allow tuning of the final focus at different beam energies. The drift distance from Q1 to the IP, l^*, is chosen to be 2 m due to geometric constraints and to allow space for vertex and luminosity detectors. QA must not be located in the fringe field region of the detector solenoid, since strong transverse fringe fields can demagnetize it. QA and Q1 are located well within the detector volume and will need to be supported from it.

The Q2 magnets are located outside the detector volume. Though their vibration tolerances are rather severe, we will assume that these magnets can be anchored mechanically to bedrock in the tunnels leading to the IP. As discussed in Appendix C and below, seismic motion in bedrock has a negligible impact on the NLC luminosity. We will not consider further the vibration stability issues for Q2, but will now focus our attention on the more difficult situation for QA and Q1. Figure 12-29 proposes that QA and Q1 be linked by optical transport arms of a laser interferometer to bedrock. This will be discussed below.

Appendix C gives a detailed description of the theory and measurement of ground motion. As is discussed there, the quadrupole alignment tolerances are very insensitive to disturbances with wavelengths which are long compared to the local lattice betatron wavelength. The dominant seismic effect is due to the microseismic peak, which has an amplitude of about 100–200 nm and a frequency of about 0.15 Hz. The wavelength of this disturbance, however, is many kilometers and so does not cause a relative misalignment of the final focus quadrupoles. Disturbances with frequencies below 1 Hz show highly correlated motion over separations of up to hundreds of meters, and the residual uncorrelated motion for the opposing final focus quadrupoles should be correctable by means of slow feedback. This feedback can monitor the deflection angles of the colliding beams to determine corrections to be made to the incoming beam trajectories. For frequencies above 1 Hz, seismic motion at quiet sites is less than 1 nm and therefore less than

the tolerance for the final focus quadrupoles. Thus, if one could fix the FF quadrupoles to bedrock tens of meters below the earth's surface, seismic motion would have a negligible effect on colliding beam luminosity.

In practice, the final quadrupoles will be mounted inside a large particle detector. The structural characteristics of such a detector weighing thousands of tons and housing many layers of particle-detection apparatus in a strong 2–4-T magnetic field are similar to those of a large cube of gelatin at the nanometer scale. It will not be possible to construct independent supports for the final quadrupoles since the detector encloses nearly 4π of solid angle around the collision point. The final quadrupoles will have to be supported from the detector. Changing thermal gradients, cooling-system noise, cultural noise, and amplification of noise sources by the mechanical structure of the detector and quadrupole supports, will all contribute to mechanical noise.

Measurements made at SLAC (Section 12.4.5) and elsewhere indicate that even with good mechanical design, one may expect residual vibrations for the final quadrupoles at the level of 50-nm rms for frequencies above 1 Hz. This level of vibration will have to be measured and corrected to better than 1 nm.

Both optical and inertial measurement techniques are feasible for detecting sub-nm motion. The inertial devices are described in Appendix C. and these have been used extensively at SLAC and elsewhere to characterize ground motion. However, it may not be feasible to use these devices for the final quads that are buried inside the magnetic field of a large particle detector. An alternative method is to use a laser interferometer. A description of how to measure sub-nm displacements with an interferometer is given in Section 12.4.3, and a preliminary description for how to use such a device to stabilize the final quads is described in Section 12.4.4.

If motion of the final quadrupoles buried inside the detector can be measured with respect to bedrock, there are several approaches to stabilization. One indirect method would apply these signals to an external correction magnet to stabilize beam trajectory either by dithering its current or by piezoelectric positioning. A more direct approach would be active piezoelectric positioning of the final quadrupoles themselves. This approach has all the advantages of linearity and stability intrinsic to direct closed-loop feedback regulation around a null.

A number of activities are currently planned to address vibration issues for the IR. These include more measurements of vibrations of quadrupoles in the FFTB. The FFTB is also planning to commission an rf BPM capable of measuring beam position and beam jitter to better than 10 nm. If this is successful, they can try to correlate the beam jitter with the quadrupole jitter, and if a correlation is observed to try to correct for it. More vibration measurements are needed to try to quantify effects of cultural noise and vibration amplification by mechanical structures. Simulation work is needed to understand the vibration measurements being made and to develop engineering guidelines for a mechanical design of the IR and Detector. In Sections 12.4.3 and 12.4.4, we describe an interferometer scheme to optically anchor the FF quadrupoles to bedrock. We plan to propose developing such a scheme with a mockup of an FF quadrupole to demonstrate that 50-nm vibration jiter can be measured and corrected for. This proposal will require developing independently a laser interferometer to measure the jitter and a piezo-mover system to correct it. The complete mockup would integrate a simulated FF quadrupole with its interferometer and piezo-mover systems.

12.4.3 Measurement of Sub-nm Displacements by means of a Laser Interferometer

The relative displacement of two objects at the subnanometer level can be measured by a laser interferometer. The LIGO experiment [Abramovici, 1992] proposes to measure the relative displacement of two masses, separated by 10 km, to a precision of 10^{-18} m. At the NLC, we only need to measure the relative displacement of the final quadrupole doublets, separated by 10 m, to a precision of 10^{-9} m. This may seem easy by comparison, but the geometry and accessibility of the final quadrupoles, buried inside the NLC detector, is much more complex. LIGO sensitivity will be at frequencies greater than 100 Hz, while the NLC is sensitive to frequencies below 50 Hz, where motion amplitudes are much greater. The measurement technique proposed by LIGO uses a laser interferometer

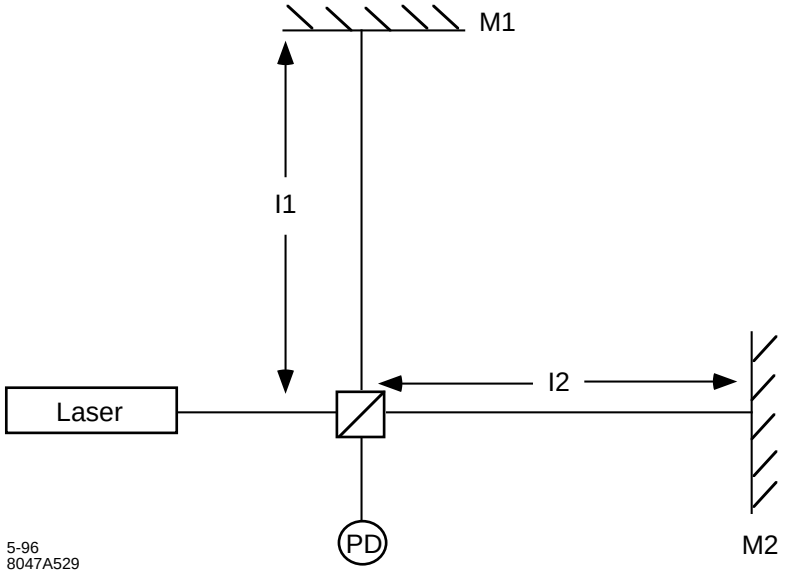


Figure 12-30. *A simple laser interferometer.*

somewhat more complex than the one indicated in Figure 12-30, but this figure serves to illustrate the measurement technique.

In Figure 12-30, we wish to detect relative changes in the lengths of the optical arms, l_1 and l_2, by monitoring intensity changes on the photodiode. The photodiode signal is given by

$$\begin{aligned} I &= N_\gamma(1+\cos\Phi) \\ \Phi &= \Phi^0 + \delta\Phi \end{aligned} \tag{12.10}$$

where N_γ is the number of incident photons, Φ^0 is the equilibrium phase difference between the two interferometer arms, and $\delta\Phi$ is due to changes in $(l_1 - l_2)$. The intensity change at the photodiode due to a change in $(l_1 - l_2)$, is given by

$$\begin{aligned} \delta I &= N_\gamma \sin\Phi^0 \delta\Phi \\ \delta\Phi &= 2\pi \cdot \frac{\delta(l_1 - l_2)}{\lambda} \end{aligned} \tag{12.11}$$

where λ is the wavelength of the laser light. One can choose to set $\sin\Phi^0 = 1$, in which case one gets

$$\frac{\delta I}{I} = \delta\Phi \quad . \tag{12.12}$$

For $\lambda = 633\,\text{nm}$, and $\delta(l_1 - l_2) = 1\,\text{nm}$, one will have $\delta\Phi = 0.02$ and this will cause a 2% intensity change on the photodiode.

There are many tolerances to be satisfied by the laser and optical system in this interferometer. These include tolerances on the laser's intensity stability, frequency stability and power. There are also tolerances on pressure and temperature fluctuations in the optical transport arms, which require that the transport arms be evacuated.

The laser intensity stability is required to be less than 2% for sub-nanometer displacement measurements. (Actually, the photodetector and its signal processor can be configured to detect and correct for laser intensity fluctuations. However, lasers with 1% intensity stability are commercially available.)

To determine the frequency stability required, we assume that the two optical path lengths can be equalized to better than 1 mm. Then

$$\begin{aligned} \frac{\Delta f}{f} &< \frac{10^{-2}}{2\pi} \times \frac{633 \times 10^{-9}}{2 \times 10^{-3}} \\ \frac{\Delta f}{f} &< 5 \times 10^{-7} \end{aligned} \tag{12.13}$$

The laser must have sufficient power that photon statistics will not cause significant noise in the measurement. For frequencies up to 100 Hz, we require

$$\begin{aligned} P &> \frac{(10^{-2})^2 \gamma}{10^{-2}\text{sec}} \times \frac{2\text{eV}}{\gamma} \times (1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}}) \\ P &> 3.2 \times 10^{-13}\,\text{W} \end{aligned} \tag{12.14}$$

Pressure fluctuations in the laser transport arms will cause a change in the index of refraction, n, and hence a change in the optical path length. In air, the dependence of the index of refraction on pressure and temperature is given by $(n-1) = 1.2 \times 10^{-4}[P(\text{Torr})/T(^\circ\text{K})]$. An optical path length of 15 m and a temperature of 300°K, will require $\Delta P < 10^{-4}$ Torr.

Temperature fluctuations will cause a change in the optical path length due to the strain coefficients for the mirrors and beam splitter. If we assume a strain of 10^{-6} per °C, then a 1-cm-thick optic can change the path length by 10nm/°C. Thus, the temperature will need to be stabilized to 0.1°C.

A laser which satisfies these requirements is a commercially available frequency-stabilized HeNe, with the following characteristics:

$$f = 473.61254\,\text{THz} \tag{12.15}$$

$$\frac{\Delta f}{f} < 10^{-8} \tag{12.16}$$

$$\frac{\Delta I}{I} < 5 \times 10^{-4} \tag{12.17}$$

$$P = 1\,\text{mW} \tag{12.18}$$

$$\text{Cost} < \$4000 \tag{12.19}$$

12.4.4 An Optical Anchor for the Final Quadrupoles

To gain the necessary stability for collision of nanometer beams, the final quadrupoles must be rigidly connected to bedrock outside the detector. This has been done in the past by optical interferometry. [Wyatt, 2982] The final quadrupoles could be anchored to points buried deep into the walls of the detector hall by building arms of an interferometer out from the final quadrupoles to corner cube reflectors mounted to stable rock. Optical paths are angled at 60° as shown below in Figure 12-31. (Note that the choice of a 60° angle is not critical, and that smaller angles can be chosen to work equally well.)

For a 60° geometry, the differential change in the optical path between the two arms d equals the transverse motion of the beam splitter S. Each quadrupole would require individual interferometers to measure y transverse motion (we assume that measurement and correction of x vibrations is not necessary). Optical transport arms from the quadrupole to bedrock retroreflectors would be 15–20-m long and would require vacuum pipes 60–80-mm in diameter out through

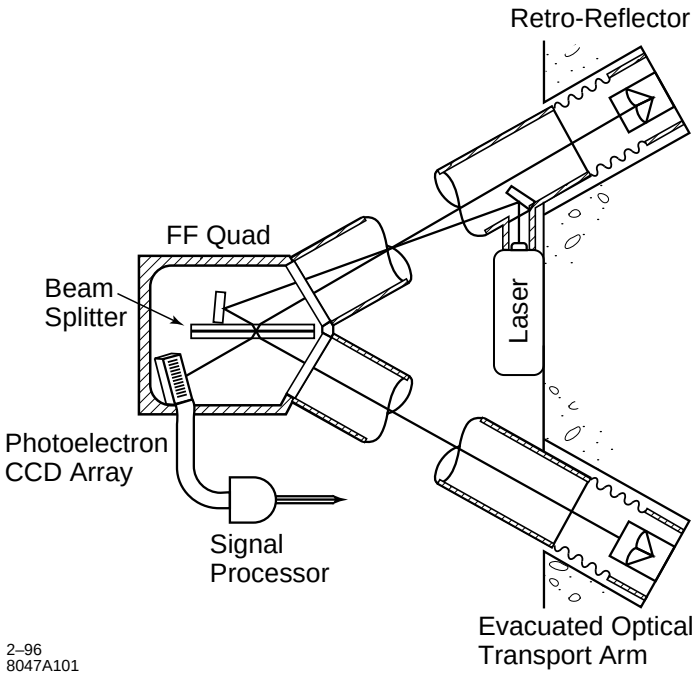


Figure 12-31. *Interferometer for Optical Anchoring of an FF quadrupole.*

the detector. For each QA or Q1 quadrupole, we assume that the quadrupole acts as a rigid body (not a trivial assumption for nanometer vibrations of a 1.5-m-long quadrupole; we will need to investigate over what lengths this assumption is valid), and that independent interferometer systems will be needed to stabilize each end of a quadrupole.

As described in Section 12.4.3, a commercial frequency-stabilized HeNe laser can be used for this system. The photodetector and its signal processing system are required to detect phase changes corresponding to less than 1/500th of a fringe (*i.e.,* less than 1 nm). We are currently investigating commercial systems to do this, but it is possible this will need to be custom-built. There are at least two companies who appear to make appropriate piezo movers with sub-nanometer resolution and accuracy, and we are investigating the specifications for these.

12.4.5 SLD Final Focus Quadrupole Vibration Measurements

Preliminary results are available from vibration measurements of the SLD final focus quadrupoles. The SLD final focus quadrupoles are a superconducting triplet supported from SLD's endcap door. The measurements were made with the STS-2 seismometers described in Appendix C.

In the final-focus tunnel leading to the collider hall and SLD, the seismometers measured about 10 nm of rms vertical motion above f = 1 Hz. On the triplets themselves, the measurements were in the range of 30–50 nm of vertical rms motion above $f = 1$ Hz. These measurements were made with the solenoid cooling water on and the cryogenic He flow on. There was no observed effect from the He flow, though there appeared to be an effect due to the cooling water. The solenoid was not powered during these tests.

It appears that triplet vibrations do not cause loss of luminosity at the SLC where the beam sizes are greater than 500 nm. And it appears that one might expect vibrations for the final focus magnets inside an NLD to be on the order of 50 nm or less for $f > 1$ Hz. (One might expect less at a quieter geographic site, and when a support design is engineered giving consideration to vibrations at the 10-nm level.) If such is the case, this vibration would need to be measured and corrected for to about 1-nm accuracy. It also appears that superconducting magnets are acceptable in the final focus, so that one is not limited to permanent magnets.

12.4.6 Measurement of Polarization and Beam Energy

One key advantage of doing physics at the NLC is that the electron beam will be highly polarized. The experimental challenge is to measure this polarization accurately enough for the physics channel under study. At the SLC, this is accomplished to an accuracy of $\simeq 0.5\%$ by colliding the electron beam with a longitudinally polarized photon beam of known polarization and using the spin asymmetry in the cross section of the resulting $e\gamma$ Compton-scattering interaction to determine the beam polarization.

Two major mechanisms induce depolarization during the beam-beam interaction. They are discussed in Ref. [Yokoya, 1988]. These are the classical spin precession under the collective field of the oncoming beam, and the spin-flip effect from beamstrahlung (Sokolov-Ternov effect). The latter is typically the most serious problem. For the range of parameters of the collider designs under discussion, one can place the limit $\Delta P < 0.04 n_\gamma$. We need to determine the precision to which the polarization needs to be measured (this will later be moved to the Physics section being developed). The details of the polarization measurement are described in the Final Focus Chapter 11 Section 11.8 on the Extraction Line.

The SLC measures the beam momenta to 20 MeV, or about 0.05%, using a spectrometer-like system which measures the separation and width of two synchrotron radiation stripes produced before and after the beam is bent by the field of a very well measured magnet. Again we need to specify the precision to which this needs to be measured. The details of the measurement are described in the Final Focus Chapter 11 Section 11.8 on the Extraction Line.

12.5 Conclusions

More work is needed at all levels. For example, the same tools used for the other background sources can be used to place limits on the allowable number of particles in the beam tails which may strike the quadrupole aperture. We have not discussed the possibility of using fast detector timing as a means of resolving background processes, nor the general question of how backgrounds will affect trigger rates. Backgrounds have not been discussed in the context of radiation damage to detectors.

References

[Abramovici, 1992] A. Abramovici *et al., Science* **256**, 325 (1992).

[Barklow, 1992] T. Barlow, P. Chen, and W. Kozanecki, "Beamstrahlung Spectra in Next Generation Linear Colliders", SLAC–PUB–5718 Rev., (April 1992).

[Chen, 1990] P. Chen and K. Yokoya, "Beam-Beam Phenomena in Linear Colliders", US/CERN School on Particle Accelerators (1990).

[Chen, 1992] Pisin Chen, Differential Luminosity Under Multiphoton Beamstrahlung, *Phys. Rev. D* **46**, 1186 (1992).

[Feldman] G. Feldman, MUCARLO, Version 1.0 (1988).

[Frary] M.N. Frary and D.J. Miller, in Ref. [Munich, 1991], p. 379.

[Helm, TLCFF28] R. Helm, Transport Deck TLCFF28.

[SLAC 1996] R. Jacobsen, *Proc. 1996 SLAC LC Physics Workshop*, (1996).

[TRC 1995] International Linear Collider Technical Review Committee Report 1995, G.A. Loew, Chairman, SLAC–R–95–471, (December 1995).

[Keller, 1991] L.P. Keller, "Calculation of Muon Background in a 0.5-TeV Linear Collider", SLAC–PUB–5533 (April 1991).

[Keller] L.P. Keller, Muon Background in a 1.0-TeV Collider, SLAC–PUB–6385 (October 1993).

[Kuraev, 1985] E.A. Kuraev and V.S. Fadin, *Sov. J. Nucl. Phys.* **41**, 466 (1985).

[Mokhov] N. Mokhov, MARS13 muon production and transport code.

[Peskin, 1991] M.E. Peskin and M. Strassler, *Phys. Rev. D* **43**, 1500 (1991).

[Wyatt, 2982] F. Wyatt, K. Beckstrom, and J. Berger, *Bull. Seismological Soc. of Am.* **72**, 1707 (1982).

[Yokoya, 1988] K. Yokoya and P. Chen, "Depolarization due to Beam-Beam Interaction in Electron-Positron Linear Colliders", SLAC–PUB–4692 (1988).

[Munich, 1991] P.M. Zerwas, ed., e^+e^-*-Collisions at 500 GeV: The Physics Potential, Parts A and B* (Workshops at Munich, Annecy, Hamburg, 1991), DESY 92-123A&B (1992).

Contributors

- G. Bowden
- M. Breidenbach
- R. Frey
- S. Hertzbach
- L. Keller
- T.W. Markiewicz
- T. Maruyama
- R. Messner
- T. Usher
- M. Woods

13

Multiple Bunch Issues

Contents

13.1 Introduction

Obtaining the full design luminosity in the NLC requires that a train of about 90 bunches be accelerated on each machine pulse while preserving the emittance and stability of the beam. Operating a linear collider in multibunch mode has an impact on the entire machine. Many of the multibunch issues have already been discussed at some length in previous chapters. In this chapter we give a general overview of multibunch issues, and we will point out the most crucial multibunch problems, *i.e.,* those problems that have significantly affected the overall machine design and required the most effort to solve. One important example is control of multibunch emittance growth in the main linacs; this strongly impacts the design of the accelerator structures. We summarize the proposed solutions to the problems posed by multibunch operation.

Multibunch issues exist in the SLC, since there are three bunches (the e^+ and e^- colliding bunches, and the e^- bunch used to produce the positron bunch for the next machine pulse) accelerated down the SLC linac on each machine pulse, and there are two bunches circulating in the damping rings during normal operation. However, this is a small number of bunches compared to the 90 bunches per pulse in the NLC design, and the bunches are closer together in the NLC (1.4-ns apart, as opposed to about 60 ns in the SLC linac). Multibunch issues also exist in long-pulse operation of the linac at SLAC. Multibunch beam break-up was encountered when the SLAC linac was turned on, and detuning of the dipole modes was used to help control it. Also, it was necessary to control the energy spread of the long-pulse beam.

Furthermore, the NLC main linacs are at higher frequency (11.424 GHz) compared to 2.856 GHz in the SLC linac. We have chosen to go to this higher (X-band) frequency in the main linacs of the NLC, because of the savings in power and the higher accelerating gradient obtainable. Even if we had not chosen X-band, control of multibunch beam break-up would still be an issue, but it is nevertheless much more severe at X-band. The X-band accelerator structure is smaller and closer to the beam, resulting in much stronger wakefields in the main linacs of the NLC, unless additional measures are taken to reduce these wakefields. This has been the major force driving the design of new types of accelerator structures for the NLC, namely the Gaussian-detuned structure and the damped detuned structure (DDS) discussed in Chapters 7 and 8. Once the design and fabrication techniques are developed for the main linac accelerator structures, it is also convenient to apply them to the design of the accelerator structures for the other lower-frequency linacs (S-band and L-band) that are part of the NLC design.

Regulation of the bunch charges is a second very important issue, because of its impact on multibunch energy control, particularly in the main linacs. As was discussed in Sections 7.4.5 and 8.2.8, the method chosen for multibunch energy control is to fill the accelerator structure with a field profile that simulates that of the beam-loaded steady-state in the structure. The ideal profile depends on the charge in the bunches. If the charge of each of the bunches in a train jitters by as little as a percent from pulse to pulse, this compensation of the beam loading in the main linac is upset. This places tight tolerances on the sources and may also necessitate collimation and feedforward systems to control the charge profile over the length of each train.

A third major multibunch issue is polarization of the electron beam. We wish to obtain a train of 90 bunches of electrons with at least 80% polarization. A gun with a strained GaAs cathode capable of achieving this is under development. Although such a gun is somewhat beyond what has been achieved at present, it is believed to be well within reach on the timescale needed for NLC.

These three problems—control of long-range transverse wakefields, regulation of average current from pulse to pulse, and obtaining high polarization from the e^- source—are the multibunch-related beam dynamics issues which we have identified as most critical. In addition, the development of new instrumentation, in particular new diagnostics such as structure beam position monitors, will be critical to the success of multibunch operation in the NLC. In the next section, we will turn to a survey of these and other multibunch issues starting from the beginning and proceeding to the end of the machine.

13.2 Major Impacts of Multibunching

Multibunch issues are in general closely tied to other issues in the NLC design. For example, interbunch and intrabunch dynamics cannot be considered completely independently of each other. As noted above, the transverse dipole wake left in a linac by the bunches at the front of a train exerts transverse forces on subsequent bunches, and thus directly affects the growth of the projected multibunch emittance at the end of the linac. However, other effects come into play in determining the final emittance. The longitudinal wake left by a given bunch affects the energy and energy spread of subsequent bunches in the train. The transverse single bunch emittance can be blown up by the combination of intrabunch energy spread and transverse kicks, since particles of different energies will filament onto different trajectories unless the dispersion is zero. On the other hand, single-bunch filamentation can damp the motion of the bunch centroid and thus reduce its effectiveness as a driver of the transverse wake. For these and other reasons, single-bunch and multibunch trajectory correction and emittance control are strongly interrelated and must often be considered together.

Because of the many inter-relationships between multibunch issues and other issues, and because it was logical to organize the bulk of this design report according to geographical regions, many of the studies dealing with multibunch issues have been discussed at length in other chapters. However, in this section, for the convenience of the reader who wants an overview of multibunch effects in the NLC, we summarize the major findings of these studies and refer the reader to sections of the design report containing further details. Since some features of the design having to do with multibunching rely heavily at this time on simulations, we summarize briefly some of the simulation methods used in obtaining the results presented here and in other chapters. We will describe the solutions found to the multibunch problems and indicate where we expect further work will be concentrated as the detailed implementation of these solutions evolves.

The final goal of the NLC is to obtain two opposing trains of bunches with suitable properties for doing experiments at the interaction point (IP). The nth bunch in each train must meet its counterpart in the other train sufficiently close to the nominal IP. Thus the transverse offsets in x and y from the nominal incoming orbits must be small compared to the respective transverse bunch sizes. The two bunches meeting at the IP at a given time should pass through each other with maximum overlap. The centroid energy deviation and the energy spread of each bunch should each be no more than a few tenths of a percent. Perhaps the most difficult tolerance is that the projected multibunch emittances in each train are to be kept close to the desired transverse single-bunch emittances of $\gamma\epsilon_x = 4 \times 10^{-6}$ m-radand $\gamma\epsilon_y \approx 10^{-7}$ m-rad. Obtaining these properties at the IP requires careful control of various parameters in the other regions of the machine. In addition to tolerances on machine components, there will be a need for feedback, feedforward, special instrumentation, and beam-based correction techniques to operate in multibunch mode.

13.2.1 Electron and positron sources

There are several important multibunch issues that must be considered in designing the e^- and e^+ sources (for details on the electron and positron sources, see Chapters 2 and 3). One is bunch-to-bunch charge uniformity within a given train. Another is train-to-train total charge uniformity. The electron gun must be capable of producing a train of 90 bunches that are only 1.4-ns apart, and the population of each bunch needs to be up to 2.8×10^{10} (NLC-IIc with 20% overhead). Furthermore, the polarization of the electron beam should be at least 80%.

A laser modulator using rf-driven resonant Pockels cells is used to turn the approximately 100-ns DC laser pulse into a train of nearly square pulses with period 1.4 ns and width 1 ns. It will also be possible to obtain trains with bunch-to-bunch spacing of 2.8 or 5.6 ns, rather than the nominal 1.4 ns.

As was noted in the previous section, the train-to-train charge jitter tolerance is very tight because of its effect on multibunch beam-loading and thus on keeping the overall energy spread of each train within tolerance. A conventional DC gun (with a strained GaAs cathode to produce polarized electrons) has been chosen for the baseline NLC design, but an rf gun, which could inject flat, smaller-emittance beams, is under consideration as a possible upgrade. One major reason (though not the only one—survival of the cathode is probably the most important reason) for choosing the DC gun over the rf gun is the difficulty of achieving the laser intensity stability tolerance in the higher-bandwidth laser that would be required in the rf gun, in order to obtain bunch trains with the required charge intensity stability. The tolerance of $< 0.5\%$ rms laser intensity stability is not easy to obtain even for the baseline NLC polarized electron source. However, it is expected that it can be obtained by improvements to a feedforward system of the type used in the oscillator for the existing SLAC polarized gun.

A intensity-limiting aperture that scrapes away about 17% of the beam before it enters the injector bunching section will be used to reduce the intensity jitter below the very small required tolerance of about 0.5%. This tolerance is what is required to achieve the desired beam loading compensation in the X-band main linacs.

Beam-loading compensation in the various accelerator structures that are part of the e^- and e^+ sources is another significant multibunch issue in this region. Two basic methods of beam-loading compensation were considered. One possible energy compensation scheme (Δt scheme) is to turn the beam on before the structure has completely filled. The additional filling of the structure while the beam is passing through compensates the linear part of the beam loading. The slope of the SLEDded rf pulse can also be adjusted to compensate the quadratic "droop" in energy over the train.

Another possible scheme (Δf scheme) utilizes additional cavities driven at $\pm\Delta f$ away from the nominal central frequency. The resulting variation in phase from bunch to bunch can be used to cancel some of the variation in beam loading.

The Δf scheme has the advantage of being relatively easy to tune (by changing the amplitude of the fields in the Δf accelerator sections). However, the Δt scheme was selected for most of the source linacs because the Δf scheme gives a single bunch energy spread that is too large. A combination of the Δt and Δf schemes will be used in the e^+ capture linac, the e^- capture section, and the bunch compressor S-band prelinac.

Long-range transverse wakefields must be kept small enough to prevent multibunch beam blow-up in the source linacs. Satisfactory control of the multibunch emittance was achieved by using Gaussian-detuned structures for the positron booster linac (see Section 3.6.1) and the electron injector linac (see Section 2.4.5).

Multibunching also presents additional demands on the positron target (see Section 3.4), which must be capable of withstanding the peak and average power in the e^- beam impinging upon it. The design will be based on that of the positron target of SLC, with improvements to allow higher beam power and better intensity stability.

13.2.2 Damping rings

The damping rings of the NLC are larger and more complicated to design than those of the SLC. Each SLC damping ring contains at most two bunches at a time, while the NLC damping rings each contain four trains of about 90 bunches each; the beam loading in the NLC damping rings is much heavier than in SLC. Also, since the bunches are not distributed uniformly about the circumference (there is a gap between trains to allow time for the kickers to inject and extract a train from the ring on each machine pulse), there is a variation in the synchronous phase along each bunch train.

The required emittances of the bunches extracted from the damping rings are $\gamma\epsilon_x = 3 \times 10^{-6}$ m-radand $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad. The number of electrons per bunch in the damping ring of NLC-I is about 1×10^{10}, and goes up

to about 1.3×10^{10} for NLC-II and NLC-III. The maximum charge per bunch in the most extreme design variations under consideration is about 1.5×10^{10}, which for four 90-bunch trains in a ring of circumference ∼220 m, leads to a maximum average current of about 1.2 A.

The two main damping rings (one for electrons and one for positrons) each damp four 90-bunch trains at a time, with one train being extracted and one train immediately injected in its place on each machine pulse. The reason for simultaneously injecting and extracting a train from each ring on each machine pulse is to minimize transients in the rf cavities that would be produced by changes in the average ring current. In addition to the main damping ring, there is a pre-damping ring for the positrons which damps three 90-bunch trains at a time.

Although the trains are separated by many buckets, they can still affect each other through long-range wakefields, unless these wakefields are quite heavily damped. Injection and extraction of bunch trains must be done with minimal disturbance of other trains in the ring. There are gaps of about 60 ns between trains, so the kicker rise and fall time must be comfortably less than this, and ringing of the kicker pulse must be minimized. The separation of 60 ns between trains is about equal to the rise and fall times of the kickers existing at present in the SLC. A flattop of about 130 ns is needed to accommodate the 90-bunch train. None of the requirements on the kickers are especially difficult, although the positron pre-damping ring kicker will need to kick more strongly due to the relatively large aperture.

Beam loading and synchronous phases

As noted above, the beam loading seen by the damping ring rf system will vary in time, due to gaps between trains, and the synchronous phases of the bunches in a train are different due to the different amount of beam loading seen by each bunch. Unless the bunch-to-bunch variation in beam loading is compensated within the ring (*e.g.*, by a special higher-harmonic cavity in addition to the regular rf cavities), it must be compensated further downstream, presumably in the bunch compressors. The variation in phases along the bunch train is very nearly linear if no phase compensation is performed in the damping ring. The present bunch compressor design is able to perform compensation for this phase variation (see Section 5.4.7). However, two methods of compensating the phases while still in the damping ring are also being considered, in an effort to simplify the requirements on the bunch compressors. One possibility is to vary the generator voltage as a function of time; this requires that the klystron have sufficient power and bandwidth and is currently under study. Another possibility is to use passive, lower-frequency harmonic cavities to partially compensate the phase variation (see Section 4.4.4); the main disadvantage of this scheme is that the pattern of synchronous phases versus bunch number becomes very nonlinear, and it would be difficult to remove the residual phase variation downstream, if this were necessary.

The variation in synchronous phases due to changes in charge of a bunch train has also been studied in simulations (see Section 4.4.4). Even for a change in average charge of 5%, the resulting phase variations could be easily compensated by a damping ring phase feedback system.

Coupled-bunch Instabilities

In addition to the effect on the synchronous phases, the longitudinal wakefields (both the fundamental and higher-order modes in the cavities) produce longitudinal coupled bunch instabilities. Preliminary rf cavity design and coupled-bunch simulations (see Section 4.4.5) indicate that it should be possible to damp the longitudinal higher-order modes (HOMs) to keep the threshold for longitudinal coupled-bunch instabilities comfortably below the radiation damping rate.

The transverse wakefields, due to both the rf cavities and the resistive wall of the vacuum chamber, can produce transverse coupled-bunch instabilities in the damping rings. Assuming that the cavity HOMs are damped to have Qs less than 300∼500, then the resistive-wall impedance dominates. A bunch-by-bunch feedback system will be needed

to damp any modes that are not suppressed by radiation damping, coherent head-tail damping, or Landau damping. Even if there is sufficient damping present that all the normal modes of oscillation are stable, interference between modes can produce transient blow-up of the beam. This transient behavior can be important in damping rings since the storage times are short. In addition, for sufficiently strong wakes and long trains of bunches, the transient could be large enough to cause beam loss at injection. A bunch-by-bunch feedback system along the lines of that designed for the PEP-II B-factory at SLAC will be used to suppress these effects.

Several coupled-bunch simulation programs have been used in calculating longitudinal and transverse coupled-bunch instabilities in the damping rings. Some are based on a semi-analytic, normal-modes approach, in which the bunches need not be symmetrically placed on the circumference. Interference between the modes can produce transient blow-up of the beam even if all these modes are long-term stable. Given the coherent frequencies and normal modes, the Laplace transform can be used to obtain the motion of the bunches, taking the initial conditions into account [Thompson 1991a]. Alternatively one may use a computer tracking method to obtain the offset of each bunch as a function of time. This is straightforward and computationally efficient provided that the number of bunches is not too large and the wakefields do not persist for too many turns. Several tracking codes are in use for NLC damping ring calculations [Thompson 1991b, Thompson 1991c, Byrd 1993].

A new code to investigate coupled-bunch mode-coupling was also developed [Berg 1995]; however, this turns out not to be a significant effect in the damping rings (see Chapter 4).

Ions and other effects

A possible multibunch issue in the damping rings (and also in the main linacs) stems from the fact that the bunches in a train are coupled not only by long-range wakefields but also by the fields due to ions in the beam line [Raubenheimer 1995]. If there is significant collisional ionization and if ions remain trapped between the passage of successive bunches, then bunch-to-bunch coupling can be mediated by the ions (similar to the way that surrounding structures mediate transverse wakefields). Ions can also produce a focusing variation between bunches, which may lead to filamentation of the trajectories of different bunches. Control of these effects may put stringent requirements on the vacuum in the damping rings and linacs, according to simulations of the beam dynamics with ions present. It should be noted that our concerns about ions are based mainly on simulation results, and experimental studies of the effects of ions on multibunch operation are needed.

Calculations predict that ions produced by the beam scattering with residual gas in the electron damping ring vacuum chamber can produce a fast transverse instability within a bunch train (see Section 4.4.6). Simulations and analytical estimates suggest that a vacuum pressure of 10^{-9} Torr or better may be required to control this instability. If this is not adequate, additional gaps in the bunch train may be used to clear the ions, but obviously this is a somewhat inelegant solution. Other solutions, such as "detuning" the ion frequencies or lowering the equilibrium emittance may be possible. More work, including experimental tests, is still needed.

A different multibunch instability may occur in the positron damping ring (see Section 4.4.7). This occurs when an electron cloud is produced in the vacuum chamber from photoelectrons and their collisions with the walls to produce secondary electrons. This electron cloud can couple to the transverse motion of the bunches and lead to an instability. There is some evidence for such an instability in the KEK Photon Factory and in CESR. Theoretical predictions of the coupled-bunch instability growth rates [Ohmi 1995] agree approximately with what has been observed. Present estimates for the NLC positron main damping ring give a characteristic growth time scale of about 200 ns. If this estimate is correct, it should be taken into account in the feedback systems being designed to combat coupled-bunch instabilities. Estimates also need to be made for this possible instability in the pre-damping ring.

Experience obtained in high-current, multibunch storage rings currently in operation or under construction is of course important to the design of the NLC damping rings. The electron-positron instability will be studied in the APS as it

begins operation with positrons and in the PEP-II Low Energy Ring (LER). The ion-electron instability will be studied at the ALS, the PEP-II LER, and the KEK Accelerator Test Facility.

13.2.3 Bunch compressors

The main multibunch issue in the bunch compressor is compensation of the multibunch beam loading. Multibunch beam break-up must also be controlled in the various linacs that are part of the compressor design; these are discussed in the next section.

Compensation of beam loading

As noted above, the bunch compressors may need to perform the compensation of phase offsets produced by differential bunch-to-bunch beam loading in the damping rings. In addition, there is beam loading in the bunch compressors themselves, and the resulting bunch-to-bunch energy differences must be kept sufficiently small. It is possible to compensate the multibunch beam loading in the bunch compressor by using two rf systems having slightly different frequencies; this was assumed in the initial design studies and satisfactory results were obtained (see Section 5.4.7). However this "Δf" scheme of beam loading compensation has the disadvantage of being somewhat nonlocal, since the beam energy spread grows between the off-frequency compensation sections. As in the injectors, a combination of this "Δf" method and the "Δt" (early injection) method will be used to obtain even better results.

13.2.4 Control of multibunch beam break-up in low-frequency linacs

The transverse emittance of the multibunch trains must be controlled in the low-frequency linacs, throughout the front end of the NLC (in the sources and compressor regions). The multibunch beam break-up can be controlled by using Gaussian-detuned or damped detuned structures, as in the main linac. The pre-linacs are at lower frequency than the main linacs, so the wakefields are not as strong. However the beam is at lower energy, which makes it more susceptible to wakefield kicks. Simulations show that methods similar to those studied for the main linacs, namely the use of detuned or damped detuned accelerator structures, will control the break-up.

In some of the S-band (2.856 GHz) linacs, detuning alone may not be quite sufficient to control multibunch beam break-up. Thus an S-band damped detuned structure (DDS) is being designed for use in all the S-band linacs. It will have a total detuning frequency spread of about 6%. The modes will be damped to Qs of about 1000.

In the L-band (1.428-GHz) positron linac, Gaussian detuning with a 10% total spread is sufficient to control multibunch beam break-up (see Section 3.6.1).

13.2.5 Main linacs

As was noted at the beginning of this chapter, one of the most important issues in the design of the main-linac accelerator structures is control of the transverse wakefield. The achievement of an X-band accelerator structure that will accomplish this has been one of the major efforts in the design of the NLC. This damped detuned structure (DDS) has been discussed at greater length in Chapters 7 and 8. The structures are detuned by varying the individual cell dimensions in such a way that there is an approximately gaussian (truncated) distribution of frequencies of the

fundamental dipole mode. The dipole modes in the structures are damped via ports leading into manifolds running parallel to the structures. Construction of the first DDS is nearing completion, and DDSs will be part of the complement of accelerator structures for the NLCTA.

Calculation and measurement of long-range wakefields

Calculation of the long range wakefields in the new accelerator structure designs being proposed for the NLC has been the focus of much effort. Over the past several years, increasingly sophisticated models of the wakefields in increasingly complex accelerator structures have been developed.

A fairly good representation of the long range wake was obtained in initial simple models of detuned structures, by regarding the structure as a collection of uncoupled oscillators corresponding to the synchronous modes of the periodic structures that one could construct from each of the cells in the structure. However, a more complete and accurate treatment includes the effects of the small couplings between the oscillators. A discussion of two such models [Bane 1993], a single-passband model and a model which takes into account the mixing of the two lowest dipole passbands, was given in Section 7.4.2. These are equivalent circuit models that give the best representation we have so far obtained for the wakefields in the detuned accelerator structure without damping (except for copper losses, which are taken into account in the models via perturbation theory).

These coupled, equivalent circuit models can be further extended to include the interaction between the accelerator structure and the damping manifolds in a damped detuned structure (DDS). The first such models treated the manifolds as coaxial lines, ignoring the periodicity introduced by the openings from the cells into the manifold [Kroll 1994]. Only a single passband was included in most of this work. A more complete model has now been developed that takes into account the mixing of the two lowest dipole passbands, as well as the periodicity of the manifold [Jones 1996].

A crucial element in the design of the NLC main linacs is an accelerator structure in which the transverse wakefields are greatly reduced below those that would occur in a conventional disk-loaded structure. Experimental verification of the performance of such structures is therefore essential. The Accelerator Structure Setup (ASSET) facility in the SLC has been used to measure the wakefields in a Gaussian detuned X-band structure (see Section 8.2.10). Damped detuned structures will also be tested in ASSET as they become available during the coming months.

Beam dynamics simulations in main linacs

The results of beam dynamics simulations in the main linacs, including multibunch effects, have been discussed extensively in Chapter 8, and we will not repeat that discussion here. In this section we will briefly describe the simulation tools used to obtain those results. These simulations incorporate the calculated long range wakefields, to calculate the multibunch beam blow-up to be expected, the bunch train injection tolerances, structure misalignment tolerances, and the effects of various correction schemes. As was noted in Chapters 7 and 8, it has been found that the structure internal misalignment tolerances are very tight, due to the effect on multibunch emittance growth. Also, control of the multibunch energy spread imposes tight tolerances on the variation of train current from pulse to pulse.

Several codes have been used to study multibunch beam break-up in the main linac and in the other linacs in the NLC. These include: LINACBBU [Thompson 1990], MBLINAC [Thompson 1991d], and LTRACK [Bane 1987]. A program that can handle combined single- and multibunch emittance control and trajectory correction was developed [Kubo 1995] and used to make the initial studies on these issues.

The main simulation tool used so far to study multibunch energy compensation in the main linacs is the program MBENERGY [Thompson, 1993]. In this simulation, one may take account of input rf pulse shaping and timing, the dispersion of the rf pulse as it transits the structure, the longitudinal distribution of charge within the bunches, the long

range wake (LRW) including both the fundamental (accelerating) mode and higher order modes (HOMs), the short range wake (SRW), and phasing of the bunches with respect to the crests of the rf.

Ion effects may be important not only in the damping rings, but also in the main linacs [Raubenheimer 1992]; the basis for this expectation is almost entirely calculations and simulations. Results of these simulations are described in Section 7.4.6 and suggest that the vacuum tolerances in the main linac will be very tight, although achievable.

A new linac code, the Linear Accelerator Research (LIAR) code has recently been developed [Assmann 1996] to do the many detailed tolerance studies that will be needed over the coming months. This code is designed to flexibly accommodate new features, and is now beginning to be used for multibunch tolerance studies. It will be possible with this code to do more complete simulations that incorporate multibunch effects along with other effects to get a more accurate assessment of the many tolerances required to preserve the emittance of the multibunch beam.

13.2.6 Final focus, interaction region, and beam dumps

There are not any multibunch problems in the final focus region that are as difficult as those that must be faced in other parts of the machine. However, multibunching does have an impact on the design of the final focus, interaction region and beam dumps.

"Parasitic encounters", *i.e.,* kicks on bunches in the incoming train due to bunches in the outgoing train, must be kept sufficiently small. This is the main reason for the introduction of a crossing angle at the interaction point. The crossing angle would entail a significant loss of luminosity since the bunches are long and thin, if the bunches were allowed to be non-parallel when they collide.

To avoid this loss in luminosity, special rf cavities are used to rotate the bunches away from their direction of travel just before they collide, so that the longitudinal axes of opposing bunches will be parallel when they pass through each other. It has been checked that these "crab cavities" do not produce unacceptable kicks due to bunch-to-bunch wakefields (see Section 11.7.2).

Another component affected by multibunch operation is the beam dump, which must be able to absorb the large amount of power in the multibunch beam. The requirements on the beam dumps are very stringent, due to the large amount of total charge, small emittance, and high energy of the beam. A dump using water as the main absorbing material has been designed (see Section 11.A) to handle the electromagnetic shower from the multibunch beam of up to 750 GeV.

13.3 Machine Protection and Operations

The long bunch trains at design emittance are capable of seriously damaging certain components (including the main linac accelerator structure) on a single pulse. Thus, start-up and recovery procedures will be strongly influenced by multi-bunch issues (see Chapter 16 and Section 7.8). The control system must be capable of handling a variety of bunch patterns and modes of operation, and the stability (*e.g.,* temperature of certain components) of the machine must be preserved as one cycles through the various modes of operation.

Even a single bunch at the design emittance may be capable of damaging the machine. It is anticipated that sacrificial titanium spoilers and collimators would be placed so as to prevent damage to crucial machine components. Such spoilers would be able to withstand a single bunch of nominal emittance and intensity, but not a full nominal-emittance

bunch train. One must begin with single bunches having relatively large emittance and work up from there as the machine is tuned.

Note also that once operation with the full bunch train has been established at nominal emittance, the repetition rate must be kept high enough that the beam trajectory cannot change too quickly between pulses. Changes must be monitored and the beam must be shut off or have its emittance blown up before the trajectory wanders far enough to damage accelerator structures or other critical components.

Controlling the emittance, energy, energy spread, and trajectories of all the bunches in a train of 90 bunches is not trivial, even in simulations. There are a number of aspects of multibunch running that will require detailed online simulation and control. One such example is fine-tuning the input rf pulse in linac accelerator sections, as part of a feedback system to improve the multibunch energy compensation. Obviously this is only one of many examples—the control system software will be required to do detailed online simulations related to emittance control, trajectory correction, feedback, etc throughout the machine.

13.4 Instrumentation Specifications

Meeting the required specifications at the IP also imposes certain requirements on instrumentation of various parts of the machine. These include: BPMs that can resolve bunches (or a few bunches) within a train, multibunch emittance measurement stations, multibunch energy measurement stations, and fast kickers (bandwidth sufficient to correct alignment of trains). Design of such instrumentation is underway, and discussions are given in Sections 7.3.4 and 7.10. and in Chapter 15.

The tolerances on the alignment of the structures with respect to the beam are very tight due to the need to control the transverse multibunch wake; the tolerance is only a few microns on some scales. Thus, one of the most important issues is instrumentation of the accelerator structure to measure its offsets with respect to the beam, via detection and analysis of the beam-induced dipole wakefield. This method of measuring the alignment of the structures by looking at signals derived from the dipole wake is discussed in Section 7.10.3. and experimental tests of the method are being carried out in the SLC.

13.5 Experimental tests related to multibunch issues

A crucial element in the design of the NLC main linacs is an accelerator structure in which the transverse wakefields are greatly reduced below those that would occur in a conventional disk-loaded structure. Experimental verification of the performance of such structures is therefore essential. The Accelerator Structure Setup (ASSET) facility in the SLC has been used to measure the wakefields in a Gaussian detuned X-band structure (see Section 8.2.10). Damped detuned structures will also be tested in ASSET when they become available.

We of course wish to verify the multibunch beam dynamics simulation results as soon as possible, and so a number of multibunch beam dynamics experiments will be done in NLCTA (see Section 8.2.9), including detailed tests of the multibunch energy compensation scheme and measurements of multibunch beam breakup. Studies of the latter will be greatly facilitated by an upgraded injector for NLCTA that produces a bunch train with a charge per bunch and bunch spacing similar to that in the NLC designs.

Experience obtained in high-current, multibunch storage rings currently in operation or under construction is of course important to the design of the NLC damping rings. Bunch-by-bunch feedback systems similar to that needed to

suppress transverse coupled bunch instabilities in the NLC damping rings are being built for the PEP-II B-factory and other high-current storage rings.

13.6 Summary and Conclusions

The proposed NLC design relies heavily on multibunching to achieve the design luminosity. Some of the resulting tolerances are very tight—two important examples are the alignment of the X-band accelerator structures in the main linacs, and variations in the bunch populations from pulse to pulse. However, we believe that they are achievable by the methods we have proposed.

Experience at the SLC provides considerable guidance in pursuing solutions to the problems posed by multibunch operation. While in some ways SLC experience is limited by the fact that the number of bunches per pulse is small compared to that in NLC, there is much that can be done in SLC that is directly relevant, particularly in the characterization of the long-range wakefields and in the development of instrumentation. Experience with high-current storage rings, such as PEP-II, which have average currents even greater than that proposed for the NLC damping rings will also be valuable, especially in refining the design of feedback systems for the NLC damping rings. As has already been emphasized, one of the most crucial elements of the NLC design is an accelerator structure for the X-band main linacs that adequately controls the transverse wakefields. Here we must rely on simulations to guide the design and ongoing experimental tests to verify that the structures work as planned. Tests in the SLC will also be important to the development of instrumentation, such as structure beam position monitors that use measurements of the induced dipole wakefields to infer the alignment of the structures with respect to the beam.

Our goal so far has been to find satisfactory conceptual solutions to the problems we will encounter in multibunch operation at the NLC. Considerable detailed design and engineering remains to be done, for example, on the various feedback systems that will be required to control the multibunch emittance and energy spread throughout the machine, as well as the associated instrumentation. Another major focus will be to continue the detailed engineering of the most practical ways to manufacture the large number of accelerator structures for the X-band linacs.

References

[Assmann 1996] R. Assmann, C. Adolphsen, K. Bane, T. Raubenheimer, R. Siemann, K. Thompson, SLAC–AP–Note 103 (1996).

[Bane 1987] K.L.F. Bane, 1987a, "Wakefield Effects in a Linear Collider", in *Physics of Particle Accelerators*, edited by M. Month and M. Dienes, *AIP Conf. Proc. No. 153* (AIP, New York, 1987).

[Bane 1993] K.L.F. Bane and R.L. Gluckstern, 1993a, "The Transverse Wakefield of a Detuned X-Band Accelerating Structure", SLAC–PUB–5783, in *Particle Acclerators* **42**, 123 (1993a).

[Berg 1995] S. Berg, 1995, "Transverse Coupled Bunch Mode Coupling and Growth Rates for the NLC Main Damping Ring", NLC–Note 11, (February 1995).

[Byrd 1993] J. Byrd, 1993, "Study of Coupled Bunch Collective Effects in the PEP-II B-Factory", Presented at 1993 Particle Accelerator Conference, Washington, DC (1993).

[Jones 1996] R.M. Jones, K. Ko, N.M. Kroll, R.H. Miller, and K.A. Thompson, submitted to 1996 European Particle Accelerator Conference.

[Kroll 1994] N. Kroll, K. Thompson, K. Bane, R. Gluckstern, K. Ko, R. Miller, R. Ruth, "Manifold damping of the NLC detuned accelerating structure", 6th Workshop on Advanced Accelerator Concepts, Lake Geneva, WI, (1994a).

[Kubo 1995] K. Kubo, K.L.F. Bane, T.O. Raubenheimer, K.A. Thompson, NLC–Note–14 (1995).

[Ohmi 1995] K.Ohmi, "Beam and Photo-electron Interactions in Positron Storage Rings", KEK preprint 94–198 (1995).

[Raubenheimer 1992] T. Raubenheimer and P. Chen, "Ions in the Linacs of Future Linear Colliders", SLAC–PUB–5893, presented at LINAC 92, Ottawa (1992).

[Raubenheimer 1995] T.O. Raubenheimer and F.Zimmermann, "A Fast Beam-Ion Instability in Linear Accelerators and Storage Rings", SLAC–PUB–6740, submitted to *Phys. Rev. E* (1995a).

[Thompson 1990] K.A. Thompson and R.D. Ruth, "Controlling transverse multibunch instabilities in linacs of high energy linear colliders", SLAC–PUB–4801, in *Phys. Rev. D* **41**, 964 (1990).

[Thompson 1991a] K.A. Thompson and R.D.Ruth, "Transverse coupled-bunch instabilities in damping rings of high-energy linear colliders", SLAC–PUB–4801, in *Phys. Rev. D* **43**, 3049 (1991).

[Thompson 1991b] K.A. Thompson, "Simulation of Longitudinal Coupled Bunch Instabilities", ABC–Note–24, (7 February 1991).

[Thompson 1991c] K.A. Thompson, "Extension of Coupled Bunch Ring Simulation to Transverse Case", ABC–Note–57, (31 December 1991).

[Thompson 1991d] K.A. Thompson, unpublished (1991).

[Thompson, 1993] K.A. Thompson and R.D. Ruth, "Simulation and Compensation of Multibunch Energy Variation in NLC", SLAC–PUB–6154, in *Proceedings of 1993 Part. Acc. Conf.*, Washington, DC (1993);

Contributors

- Kathy Thompson

14

Control System

Contents

14.1 Introduction

Control system implementation is notoriously dependent on available technology. “Notoriously dependent,” because its technology base—computers, communications, and electronics—is developing at the most rapid rate of all components used in accelerators. Thus design proposals for systems to be built in five to ten years from the date of a proposal are often little more than examples of how dated ideas can become in that period of time. In spite of that fact, some rough cut on control system design must be made now. The control system will interact with most other components of the accelerator. It is in everyone's interest to define the most important features—in terms of functionality, not implementation—of hardware and software interfaces early. This will promote uniformity where appropriate and will give other systems an opportunity for review.

In this chapter, we will acknowledge that proposals for a control system design are transient and focus on two points:

1. Although the implementation design of a control system may be transient, the functional requirements that such a system must meet are much less so. Operations at the SLC and the Final Focus Test Beam have given us valuable lessons for the control system needs of an NLC [Humphrey 1992]. Thus, the first section of this chapter will describe some of the key functional requirements for NLC operations that the control system must supply and the second section will discuss some of the implications of these requirements for the architecture.
2. A serious proposal to build NLC must include a cost estimate for all significant components. The cost estimate for the control system can only be based on a specific implementation; therefore, an implementation model, based on presently available technology, is proposed in the third section of this chapter.

14.2 NLC Requirements

A control system should be tailored to its users and the challenge they face. Much of its bulk is devoted to device controller interfaces, networks, consoles, etc.; those parts will not be discussed at length in this chapter, though some rough idea of the numbers of such devices will be needed for the cost estimate and for discussions of reliability (Chapter 17). But, there are many aspects of a linear collider that impose requirements on the control system that are distinct from normal accelerator control systems. Most of our understanding of these requirements comes from experience with the SLC and its control system, which has evolved considerably from its initial implementation. They include:

- Beam-based feedback
- Measurement synchronization with beam pulses and high-bandwidth data acquisition
- Special handling of damping ring applications
- Automated procedures for tuning
- Machine modeling and simulation
- Accelerator/detector coupling.

14.2.1 Feedback

Beam-based feedback, of the type described in Appendix D. should be considered a possible engineering solution for all tight tolerances which are static enough to be within the correction bandwidth of the feedback system. Furthermore, beam-based feedback should be considered for all systems whose initial fabrication errors place them outside of predicted tolerances. However, for systems with disturbances outside of the feedback bandwidth, or "within the bunch train," other solutions such as feedforward or dedicated, very high-bandwidth feedback systems should be considered.

There are several aspects to this overall approach:

- Use of the control system, where effective, to offset the cost of meeting tight tolerances by modifying or rebuilding the system affected.
- Development of feedback to control instabilities of all time scales, especially mid to long term. In the SLC "mid term" is the highest frequency presently addressed by the fast feedback system—120 Hz.
- Strong effort to counter fast instabilities with high-powered tools and feedforward.
- Consideration of high-bandwidth feedback for systems with very tight tolerances.
- Establishment of *treaty regions* with full separation of phase space and centroid stability.

It is critical to identify which technical issues can be addressed by these means and which cannot. Mechanical, thermal, slow magnetic fields, and low-bandwidth microphonic disturbances have time scales which can be addressed by such feedback systems. Pulsed devices such as kickers and modulators must be stabilized using other feedback techniques (not beam-based).

The present system of feedback at SLC was built after much of the lower-level design was mature. It evolved through several generations of development, each incorporating the lessons learned. The pervasiveness of the use of feedback throughout the accelerator controls was not fully anticipated and therefore the system is not optimally integrated. This requires attention in the detailed design of an NLC control system.

Fast Feedback Rate Considerations

The SLC feedback system is designed in a generalized, database-driven fashion, which contributes greatly to its flexibility. The system uses standard control-system hardware. As a result of this design, unplanned control loops can often be added with only database work, without requiring hardware or software changes. In addition to the beam position monitors and correctors used for steering control, the feedback system is capable of measuring and controlling a wide variety of devices. For example, it is as easy to add a steering loop in the linac as it is to provide control of laser gun timing in the injector. Special-purpose extensions to the linear feedback system have been added to accommodate non-linear cases, such as optimization feedbacks in which the measurement responds parabolically to actuator movement. The system also provides built-in diagnostic and analysis capabilities, and the many sample-only monitoring loops provide a wealth of diagnostic information. These design features have been key to the success of the SLC system, and the NLC design should be equally flexible and extensible in order to support unplanned controls needs.

It must be noted that for this generalized feedback system to perform properly sufficient resources must be available. The specific system used in the SLC employs dedicated point-to-point links in order to minimize communication overhead. A typical launch feedback loop at the end of the SLC Linac stabilizing angle and position in two planes for

both electrons and positrons absorbs approximately 70% of the available bandwidth of an INTEL 486 single board computer running at a 66-MHz clock rate.

The feedbacks for the NLC are planned to run at the full beam rate of 120 or 180 Hz. The steering loops control the average trajectory of the bunch train rather than individual bunches, so the Q BPMs are used, which measure the average of the train. Dipole corrector magnets are used, for which the feedback controls the magnetic field by setting a DAC (digital to analog converter) which alters the current from a power supply. Correctors and other actuators need to respond (to make 90% of a requested change) in a single 180-Hz period.

For the NLC, linac feedbacks will use typically use 10 BPMs and four correctors (two horizontal and two vertical). Communications and CPU must support processing of this local data at the 180-Hz rate. In addition, a cascade correction system for the NLC will require 180-Hz communications over long distances between linac loops (in diagnostic sections).

14.2.2 Synchronized measurement

The collider will be a pulsed machine with a relatively low repetition rate. A linear collider does not have the tendency to stabilize inherent in a storage ring machine: every pulse is a new beam. Therefore it is crucial that the facility exist to take measurements synchronized to particular pulses throughout the accelerator. These will include beam measurments, hardware diagnostics, and state information from the detector. This will allow pulse-to-pulse correlations of all different aspects of the collider and is the only way to efficiently trace sources of beam jitter and subtle hardware failures. A full sample – data for every pulse – over an extended time may be needed to see both high and low frequency instabilities.

Pulse-oriented sampling is also needed to provide an estimate of electromagnetic fields and other analog monitors at the beam pulse frequency. This means that the acquisition of analog signal data should be synchronized with the beam or line frequency.

14.2.3 Data acquisition and processing

High data-acquisition bandwidth is required to characterize pulses for optimization or feedback. Ideally, the BPM system should produce phase space centroid information of each bunch in the train at the full 180-Hz rate. This will require a control system data acquisition that has a throughput equivalent to the maximum beam pulse rate. The system should be able to acquire and process data more or less indefinitely. With a large number of samples acquired at the pulse rate, fine details of machine performance can be examined such as, for example, the frequency structure of narrow excitation lines. (Such lines have been observed in the SLC beam motion frequency spectrum, and are sometimes caused by the linac water-cooling pumps).

A more serious challenge is the measurement of phase space volume or emittance. In order to measure the bunch train's transverse and longitudinal volume with the sampling mode currently in use at SLC, a minimum of several hundred pulses is required. At low repetition rate (as when trying to diagnose and correct a problem) this is too slow to allow for effective tune up. A profile monitor system has been proposed that can be used to “fast” scan the entire train in a single pulse.

14.2.4 Special architecture for damping ring applications

At SLC, most of the data acquisition system and scheduling system is optimized for a pulsed linac machine. A special architecture is required that can be used to track the performance of the damping ring hardware throughout a single storage cycle and can track the behavior of the beam at the same time. The B-factory control system, for example, will have to have an integrated timing system that handles both the injected and stored beam. NLC damping rings will need controls that address the problems of injection, extraction and damping.

14.2.5 Tuning

Complex tuning will be required in the NLC. This is true in many places but is most prevalent in the final focus. To limit the luminosity loss due to the tuning procedures, the tuning scans and measurements must be highly automated and fast. Thus, the entire scan, which usually involves changing the strength of many magnets and then measuring the IP spot sizes with a beam-beam deflection scan, would probably be performed locally before the data can be shipped back and processed.

14.2.6 Modeling and simulation

An early decision in the design of the SLC control system was to base many aspects of machine setup, operation, and diagnostics on online accelerator models rather than a look and adjust method of interaction. The adoption of this approach has resulted in the development of a rich suite of applications that forms the foundation for near-automated operation of SLC. This modeling framework has been effectively used throughout the machine life cycle including commissioning, routine operation, diagnostics, and performance upgrade and optimization phases.

Given the complexity, strict tolerances, and the expansiveness of NLC, it is vitally important to have an online modeling environment with appropriate degrees of sophistication to facilitate machine commissioning and operation. At SLC offline simulation has been used to better understand machine behavior and to investigate alternative strategies; this trend should continue at NLC.

General areas of application for modeling and simualtion would include:

- Machine commissioning where the objective is to reconcile the model with the accelerator.
- Routine machine setup based on the design models and specifications.
- Model driven feedback system as described above.
- Near-automated diagnostic capabilities such as emittance measurement, and lattice diagnostics to quickly identify the sources of machine performance degradation.
- Model-based optimization tools such as orbit and lattice properties correction applications to allow rapid fixes for performance degradation.
- Creation of “multiknobs” which allow one to vary a parameter which may depend on many hardware values in a linear or nonlinear manner.
- Ability to use the same model base for online as well as offline machine physics studies.

14.2.7 Accelerator/Detector Coupling

The interaction region is precious not only to the physicists taking data with a detector but also to those running the accelerator. Certain diagnostic information can only be obtained there, but instrumentation is likely to impinge on the detector volume. The situation can be ameliorated somewhat by close coordination between shift personnel in the Accelerator Control Room and the active Detector Control Room as well as direct communication between the control system and the detector data acquisition.

At NLC provision should be made for at least the following:

- Shared timing. The Detector Acquisition must be synchronized with the bunch train. It should also have access to the configuration of the train (number of bunches, spacing). In order to make correlations offline each bunch train should have a unique identifier available to both the Control System and the Detector Acquisition.
- Interlocks. Detector components known to be vulnerable must be in a protected state during potentially damaging accelerator tuning.
- Tuning information. The Detector Acquisition should make available to the Control System a suitable collection of background signals at bunch-train rate during periods of good luminosity or fine tuning. This information is needed both for online tuning and for offline analysis.
- Veto information. The Control System should promote knowledge of anomalous pulses to the Detector Acquisition front-end.
- Polarization state. Both the Control System and the Detector Acquisition need access to this.
- Slow updates. Various quantities – state information, statistics – kept on one side may be of interest to the other.

14.2.8 Reliability and Availability

The availability of Control and Instrumentation Systems in existing accelerators (FNAL, SLC, CESR) has achieved the availability numbers required for the NLC (98-99%). The NLC is 10 times larger, and its construction comes 10 years (or more) after the existing machines. From the reliability point of view, the fact that the NLC is larger is the engineering challenge; the fact that its construction occurs a decade later is the engineering opportunity.

The NLC Control System will use components from the computer, communications and electronics industries. These industries have a long-standing record of improving their product reliability and availability in response to market needs. Thus, one can expect that industry will supply improved component reliability which will, in turn, get us part way to the achievement of our availability goals.

It would be wonderful if we could leave it up to the marketplace to solve this problem. Unfortunately, we cannot. We have already seen in existing accelerators that availability is heavily impacted by local design and operating decisions. Spares availability, and the scheduling of time for checkout and maintenance debugging are examples of local operating decisions. The use of redundancy in designs, and the design of equipment in ways to decrease the MTTR (Mean Time To Repair), such as hot swapping of spares, are examples of local design decisions.

We have not yet expended engineering effort in putting together a design and operational strategy to achieve the availability goals of the control system. However, the goals appear to be realistic in terms of the experience of presently existing accelerators. The technologies involved for the control system are well known and tested. There are many practitioners in the fields of reliability, availability, and maintenance of large complicated systems. In the fields

of electronics, communications and computers, this is a well-established engineering discipline. We know that effort in this area has to start early in the life of a project, since reliability goals must be set and given to designers early in the design phase.

14.3 Architectural implications

To zeroeth approximation, NLC is just a large accelerator requiring a large control system. To this extent NLC can use something like the control system "standard model" prevalent in large modern control systems. But it is also clear that the functional requirements discussed in the previous section can only be satisfied by making substantial perturbations to this model. These fall into the following categories:

Network bandwidth NLC's heavy dependence on feedback will make significant demands on realtime per-pulse bandwidth between device-controlling computers (called "micros" at SLC). There is also a need for real-time communication in order to synchronize measurements. Finally, to employ online modeling and provide information for simulation, large amounts of data must flow from the micros to an arena accessible to online servers and perhaps to a logger. That is, the control system will have a substantial data acquisition component.

Computational resources As pointed out in Section 14.2.1 a single fast feedback loop at SLC consumes the better part a micro. By contrast, less than 100 micros (mostly running at a lower clock rate) are used for standard device monitoring and control of the thousands of devices comprising the control system. The lesson for NLC, with its anticipated heavy use of feedback, is clear. Online modeling will be another cpu-intensive activity.

Interface with detector data acquisition In order to support the communication outlined above in Section 14.2.7, well-defined interfaces (hardware and software) and adequate network bandwidth between the NLC control system and the detector data acquisition will be needed. This kind of communication has proven essential for SLC/SLD (for example, parts of the detector readout are of great value as accelerator diagnostics). but was absent from the original design. The ad hoc methods currently employed lack flexibility and the amount of information which can be transmitted at 120 hertz is pitifully small.

Hierarchical software organization The cost of the control system for the SLC was much larger in proportion to total project cost than that of any other machine yet built. Already 200 to 300 person-years of software resources have been invested and further improvements are planned. This software development effort has been primarily an intellectual challenge as opposed to a bookkeeping or task management exercise. In many ways, the pace of evolution of the control system has been limited by the time required for the machine physicists to identify and understand a problem and specify a solution.

To appreciate the need for hierarchical organization within the NLC control system it is instructive to look at SLC as it has evolved. SLC's database is in many respects the heart of its control system. During the first years of its existence (1980–1986), approximately 80 kinds of objects (with typically many instances of each type) were defined in it. All but a handful describe hardware. Since that time about another 90 kinds of objects have been added to the database, but of these well over half describe something bearing little resemblance to a physical device, for example feedback loops or model parameters. The database, like the system and application software which access it, is both substantially larger and different in character from initial expectations.

This evolution at SLC was only possible because of the generic nature of the treatment of the lowest level components. Strict rules are followed in all low-level device interfaces, and this structure facilitates continued growth through the application of higher and higher layers. This implies that the design of the lower levels is critical for a control system which must support a complex variety of high- and very high-level applications.

With its tight tolerances which must be actively maintained, and opportunities for subtle interactions among its elements, NLC will continue this trend. In order to get their job done efficiently all who use the control system (operators, machine physicists, hardware maintenance personnel, programmers) must be able to access it at an appropriate level of abstraction so they can control, monitor and analyze the entities of interest to them.

14.4 The Control System Model

As noted in the Introduction, we will describe a control system implementation model that is, we believe, a realistic model for an NLC Control System, if it were being built today. It is realistic because it follows the main thread of control system design used in accelerators recently built (CEBAF, ALS) or under construction (PEP-II). However, we will state the standard disclaimer that this is a design based on current technology, and that the real NLC control system design will be based on the technology available at the time of its construction.

This model is based on what has become a fairly standard approach in modern computer control systems (Figure 14-1); it comprises a set of consoles and servers linked to each other and to a hierarchically lower set of front-end computers via high speed networks. These front-end computers are, in turn linked to devices via dedicated control networks. The devices themselves are expected to often include embedded computers; thus the front-end computers are really communicating with still another architecturally-lower layer of device computers. In the case of low to medium multiplicity devices (1 to 1000 devices), there may be a device interface crate (VME, VXI) which contains a module which controls the device. We expect that the truly high multiplicity devices (Beam Position Monitors, Magnet Movers; greater than 5000 devices) will probably have a cost-optimized design which includes dedicated embedded computers, and communicates to the front-end computers via a digital network link of some kind.

14.4.1 Operator Consoles

Console Computers will be located in the Control Room and will be the Primary Operator Interface (OPI) into the Control System. There will be a Graphical User Interface (GUI) driven by the EPICS Control System based on the X-protocol. The physical hardware will be comprised of a processor device driving (perhaps) two video heads with some pointing capability plus a standard keyboard (Figure 14-2). These workstation processors will have large internal memory and internal hard disks.

These machines are modeled as DEC Alpha Processors running the NT Operating System. These machines are capable of high throughput, excellent number crunching, and a fast network response.

Consoles will use the TCP/IP over Ethernets which are in turn connected to the FDDI backbones via switching bridges. Separate local networks will be used to subdivide the overhead status display screens and the pairs of screens associated with each console such that network or server failures will not bring down the control room facilities.

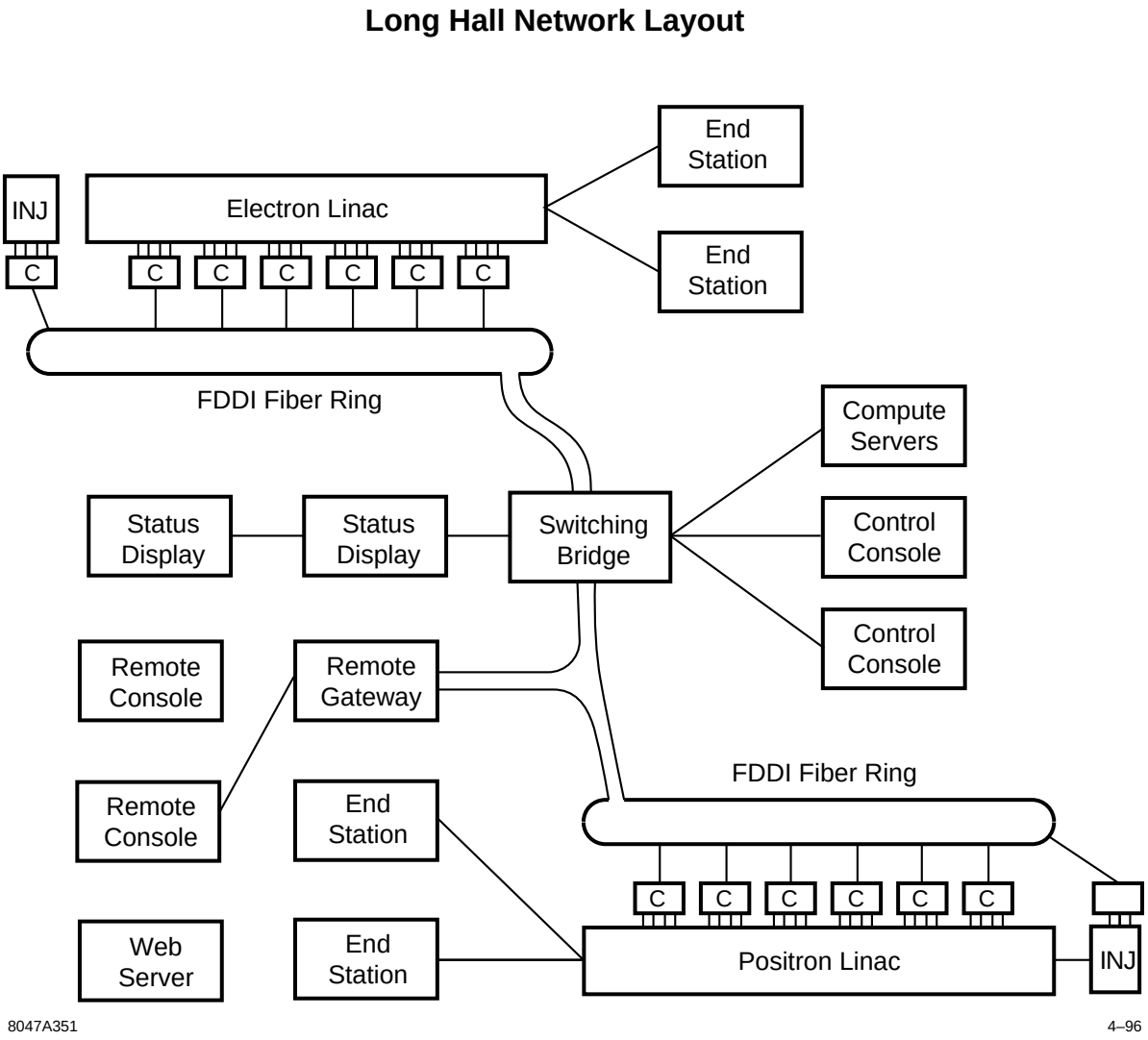


Figure 14-1. *Schematic of the control system layout.*

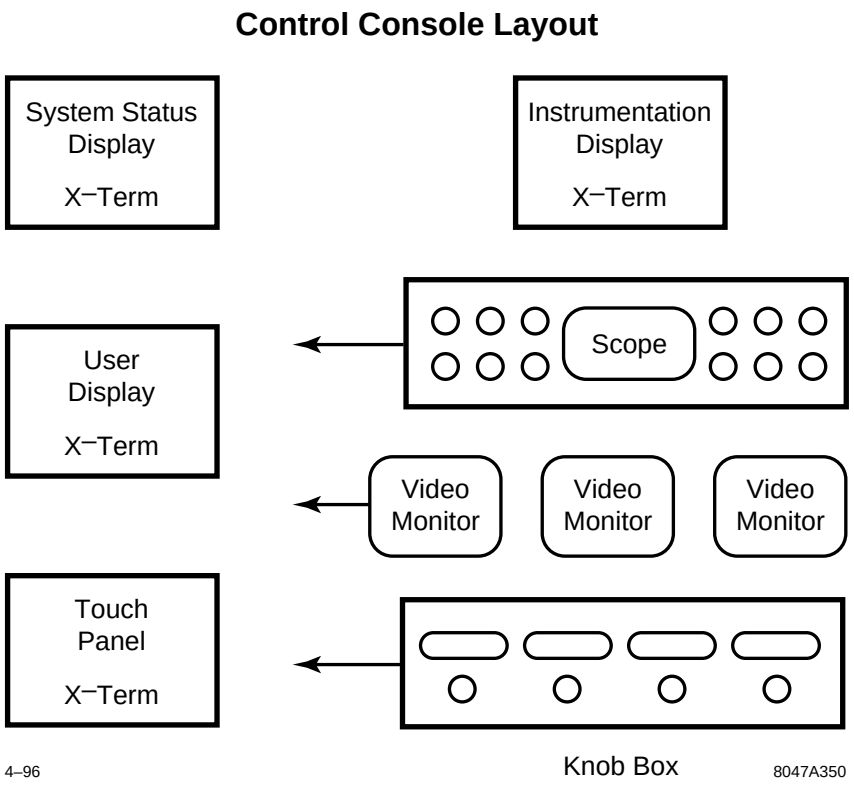


Figure 14-2. *Schematic of operator's console.*

14.4.2 Application and Server Computing Resources

The control system model is a distributed system, and one component of that distributed system will be a set of around 20 computers which will function as servers of various functions. The functions we have in mind are database, network interface, computational, and file servers.

Database servers will include the Control System Database and the Documentation Database (drawings of devices and cable-plant, wire lists, system documentation, operations guidelines, code management, etc.). Network interface servers will handle connections and security issues for computer communication access outside the control system. Computational servers will supply sufficient computing power for accelerator modeling and simulations. File servers will provide storage for a number of functions, including software and firmware development, the archiving of data from the control system of device histories, operation histories, configurations, alarm message and error reporting, computer system and network management monitoring data, etc. Many of these functions will be accessible to the detector data acquisition; conversely, information from the detector will be accessible to control system applications.

Associated with these system will be large disc farms for storage of the large amounts of data associated with such an accelerator complex. Some form of redundancy such as clustering will allow the server complex to gracefully degrade in the event of individual server or other component failure. We expect that this server cluster will be in relatively close proximity to the control room, so that local very high speed links connecting to the consoles in the control room may be utilized.

Network Security Systems

There are requirements to safeguard the Accelerator from remote access and control by unauthorized individuals. Security systems will be implemented at the onset of the program which allow data to be freely available over external networks, while preventing unauthorized remote users from operating equipment. These security systems will provide encrypted remote sessions such that passwords and private data cannot be swept for unauthorized use.

14.4.3 High Speed Networks (FDDI & Ethernet)

This Distributed Control System depends heavily on networks with substantial bandwidth to function properly. Thus it is critical to use network facilities which can handle anticipated loads at low utilization figures. These networks need to be industry standard facilities as well.

Accelerator Control Networks will be subdivided into subnets to isolate the operation of the machine from another networks activities at the laboratory. There will be a control network Firewall router between general laboratory networks and the control networks to allow isolation if required.

In terms of current technology , FDDI and Ethernet make good sense because they are industry standards, they are versatile, and their performance is well understood. At the time of the final design review, it will be necessary to review these specific selections and to use the then current stable dependable technologies.

There are some special considerations for the networks, in that the length of the network runs will generate challenges for any network technology. Fiber implementations will be important, and microwave facilities may be an option to consider.

14.4.4 Front-end Computers

Front-end computers are the Input/Output Controllers (IOC) which actually perform the data acquisition and control in real-time in the accelerator housing. These machines are microprocessor based platform using PCI, and VXI bus backplanes. Each device will have its own Real-time Executive, and will run a standard set of acquisition/control software. Database configurations will be downloaded in order to structure the number and type of devices and processes the controller handles.

These processor crates will normally be operated remotely over the Ethernet from the Control Room, but can be operated locally from a terminal for diagnostic purposes. Long-haul communications will be handled by the FDDI backbones with bridges to distribute Ethernet connections to micro crates, local control X-terms, and other Ethernet devices.

14.4.5 Data Acquisition Crates

Data Acquisition Crates will be a mixture of VME and VXI crates containing the analog and digital input/output cards used to make measurements and define set points. Low-cost VME crates will be used for most of the modules, along with some industrial equipment for the low accuracy systems. VXI will be used for the precision measurements and high-frequency rf and timing modules.

14.4.6 Instrumentation Modules (VXI & GPIB)

There will be a number of modules for high frequency monitoring and control which will emanate from VXI and GPIB controlled instruments. The VXI crates will be controlled over Ethernet through the controller module in the crate. These crates can control locally positioned GPIB instruments as well.

14.4.7 Radio Frequency Control, Phasing & Feedback

The rf requirements are covered in Chapter 8; we will not repeat those here.

14.4.8 Dedicated Control Networks

These networks are dedicated point-to-point networks to connect specific pieces of equipment. This implementation is in use at SLC has been chosen as perhaps the only one available today to get high performance and dedicated functionality.

These links will be found between control crates and large power supplies with self-contained controllers. There will also be a link or links between the control system and the detector data acquisition in order to transmit, for example, tuning information (detector to accelerator) and beam quality information (accelerator to detector). There will be point-to-point links between machine stabilizing feedbacks and feed forwards. This approach (point-to-point links) has serious drawbacks for feedback: it does not scale well and lacks flexibility. Alternatives should be investigated as they become technologically feasible.

Special 1553 links have been employed between elements of the SLC machine protection system for security and speed.

14.4.9 Timing and Beam Rate Control

The purpose of the scheduling system is to provide control of the paths to be followed by beam pulses more or less in real time. Several pulsed beam dumpers and diagnostic stations will be installed throughout the NLC; before and after the damping rings, in the injector and positron system, after the bunch compressor, at several places throughout the linac and on either side of the big bend and collimation systems. The scheduling system will control the firing of these dumper magnets and synchronize the data acquisition on the dumped pulses. It will also be used for the machine protection system (see Chapter 16).

The scheduling system is modeled on the one currently in use at SLC. It will be used to control the injection, extraction and storage time in each of the three rings. Using it, the control system will be able to program the storage time, and therefore the output emittance, of each of the three rings. This feature will be built into the system through a linking, or pointer-based, structure that will be used to track the progress of a bunch or train of bunches from it's inception through to a dump. Another requirement of the scheduling system is to provide control of synchronized data acquisition and sequencing of pulses when a fast pulsed device, such as a kicker magnet or a pulsed phase shifter, is being adjusted. Such acquisitions will be used to quickly scan a beam across a beam size monitor and for certain classes of beam optimization. They are also required for optimizations which involve measuring derivatives such as maximizations

or minimizations or maintaining a specified phase with respect to the rf. The technique to be used is a synchronous detection scheme with a sub-tolerance dither.

Accelerator Beam Rate will be controlled by a dedicated special purpose processor running custom software to support flexible rate control in both accelerators. This Master Pattern Generator will be wired into the machine protection systems to handle both rate limiting and machine protection shutdown. Rate information will be disseminated to appropriate components of the detector data acquisition as well as as to front-end computers, etc., within the control system.

Flexible rate control allows controlling the amount of beam energy transported throughout the machine. This will facilitate the alteration of beam parameters for experiments which require unique beam characteristics or timing.

This flexible beam rate control also allows the accelerator to be rate limited by classes of machine protection problems so that problems can be located and identified with low rate beams with reduced risk of accelerator or equipment damage. Beams will automatically rate limit back to designated rates as problems are resolved. Rate control will be exercised on a pulse-to-pulse basis.

14.4.10 Machine Protection Systems

The purpose of the machine protection system (MPS), which is described in Chapter 16, is to prevent damage to machine system components in the event of a routine failure. The system is not intended to provide comprehensive protection against any possible failure. One of its main functions is to automatically provide a sequence of beam pulses that can be used as effective diagnostic tools during a startup or fault period.

The machine protection system has four logical layers: 1) mechanical, 2) device controllers, 3) power monitoring and 4) beam scheduling and control. The layers provide a graded approach that allows the production of a nominal intensity single bunch beam for diagnostic purposes. All mechanical systems should be capable of withstanding the impact of a single pulse of such a beam. They should also be able to survive another strike in the same location. It may not be possible to develop structures that can stand a single pulse strike from a nominal single bunch at nominal emittance, especially for the higher intensity versions of this design. In this case, an emittance enlarging system must be integrated with the MPS so that proper operation is ensured. Once operation is checked with single bunch beams, the repetition rate may be increased, the emittance brought to nominal and the number of bunches brought to its full value. This sequence must be applied in this order. The only viable way to transport full-intensity beams is to make sure that the transverse deflecting forces acting on the full-power beam cannot change enough in the interval between pulses to target the beam cleanly on a beamline element. One consequence of this is that low-repetition rate, full-train intensity operation, is not possible.

14.4.11 Equipment & Tunnel Access Control

The control system will monitor the status of equipment and the state of Accelerator access, but the actual control of Personal Protection, Machine, protection, and Hazards will be handled by dedicated hardware and Programmable Logic Controllers.

14.4.12 Application Software

Application packages will reside and run on the console processors and on a separate applications processor in the control cluster. This extra applications processor will take on large resource computing loads that would not run well on console machines. Other applications or analysis programs will be off-loaded to user machines via self-describing data files.

Applications packages will include Accelerator diagnostic packages, measurement packages for things like emittance and chromaticity, energy management (LEM), power steering with machine optics models, correlation plots, simulations, machine models, multiknob control, and data archiving.

Measurement packages may be operated remotely at lower priority than control room activities. Analysis of data may be run on remote hosts or the application processor in the cluster. Operation of applications which operate accelerator equipment or change machine configurations or settings will execute from the control room only.

The applications environment will be structured to enhance the ability of the Laboratory to use software developed at other Laboratories or purchased commercially.

Included applications: archiving, correlation plots, steering, LEM, models, emittance, logging.

14.4.13 Software Application Bus

Applications which run on the Console or Cluster Computers will run on a software layer which will isolate them from the complexities of where data comes from, how it is stored, and how it is transported. Hidden facilities will deal with correlation plot data which has to be correlated in time and take into account measurements taken in different parts of the accelerator.

This software structure will make available common measurement and data collection facilities which may be required by application or display processes. Data will be presented and exported in self-describing formats compatible with application packages available at other laboratories [Watson 1995].

Similarly, data files will also be available in Matlab format for local or remote analysis.

14.4.14 Software Development Environment

Software development will be accomplished on workstations similar to those used in the production environment, however, they will be configured to run a parallel but separate control environment so that actual production equipment will not be controlled by accident. With the exception of the separate environments, the software environment will be identical to the production environment. Some special hardware will be developed which will help simulate accelerator operation for software evaluation and testing. Additional software will be required to compile code, to control progressive versions of software, and database facilities to build run-time databases and configure equipment.

Diagnostic systems will include remote diagnostic and debug capability for all networked microprocessor systems (not including embedded systems). Isolated or low device count GPIB instruments will be controlled via Ethernet by GPIB network control boxes placed in the locus of the GPIB instrumentation.

14.4.15 Industry Standards

The Control System will utilize as much commercial equipment and software as practicable. Industry standard equipment and facilities will be utilized to reduce cost and improve maintenance and reliability. To the extent possible, the control system will utilize electronic modules available commercially and utilized in other laboratories to reduce cost and resources involved in hardware development and in writing low-level software drivers.

References

[Humphrey 1992] R. Humphrey, "Lessons From the SLC for Future LC Control Systems," *Proc. of the 1992 Int. Conf. on Acc. and Large Exp. Phys. Control Systems*, KEK, Tsukuba, Japan, KEK Proceedings 92-15 (1992).

[Watson 1995] C. Watson *et al.*, "cdev, a Common Device API," *Proc. of the 1995 Int. Conf. on Acc. and Large Exp. Phys. Control Systems*, Fermilab, Batavia, IL (1995).

Contributors

- Joanne Bogart
- Spencer Clark
- Rusty Humphrey
- Nan Phinney
- Marc Ross
- Hamid Shoaee

15

Instrumentation

Instrumentation performs a critical role in the operation of a linear collider. New acquisition and data processing techniques are required for feedback, tuning procedures, and performance monitoring. For example, many collider systems are initially tuned using complex bootstrap procedures whose convergence rate will depend on the speed and performance of several instrumentation systems. Furthermore, mechanical and electrical tolerances are computed assuming the success of this process.

The next leap in electron-positron accelerator performance will result in part from improvements in instrumentation technology. The latest generation of accelerators, from high-current synchrotron light machines to B-Factories and linear colliders require feedback control loops that are greater both in number and complexity than more conventional machines. As a result, the instrument is no longer a diagnostic tool, intended for use only in cases of sub-standard performance, but a truly integrated accelerator component. This has obvious implications for the instrumentation-system designer, among which is that the system must have the integrity required of other accelerator systems, such as the power conversion and vacuum systems.

Linear colliders represent the most extreme application of this philosophy. The lack of closed, equilibrium conditions that maintain stability in the machine, forces the use of several layers of sophisticated feedback loops. The underlying reason for this requirement is the tolerances that must be applied for the correct transport of low-emittance beams. In some extreme cases, initial bootstrap procedures are required before any beam can be transported through the system. Tight mechanical and rf system tolerances will not only require special systems to address them directly, but will also demand beam-based feedback and tuning procedures. For example, in the X-band linacs and the beam-delivery sections, the magnet alignment is continuously monitored and adjusted using beam-based techniques that rely on high-resolution Beam Position Monitors (BPMs).

Perhaps the most important improvements in instrumentation technology will not come from the harnessing of fundamentally new physical processes to better the performance of beam position or size monitors. Instead, they will come from the integration of existing instrument beam sensors with more powerful controls. Very strong integration with the control system is needed to provide the robust, high data-processing bandwidth needed for higher level control.

An important aspect of the shift in the role of instrumentation will be its use in general optimization systems that will ultimately change the character of the control room operator's task. Traditional applications of instrumentation systems in colliding-beam accelerators have required heavy involvement of the operator. In storage rings, for example, operator technique in optimizing injection and luminosity has proven to be a key factor in long-term performance. In a heavily feedback- and optimization-control-laden system, the operator's task becomes the more complex one of controlling and monitoring the performance of these automated tasks.

Details of the instrumentation design and requirements are distributed through the preceding chapters of this document. Many of the concepts needed for the high-resolution systems have already been tested. For example, the Final Focus Test Beam (FFTB) at SLAC utilizes stripline BPMs with 1-μm resolutions and a beam size monitor that is capable of measuring 40 nm spot sizes. In addition, rf BPMs were installed and measured to have a resolution less than the required 100 nm. Other elements will be tested in the near future. This includes the rf structure BPMs that are needed to align the accelerator structures, a laser wire system similar to those needed to measure the beam emittances in the linacs and final foci, and the PEP-II button BPM system that is similar to those needed in the damping rings.

16

Machine Protection Systems

Contents

16.1 Introduction

One of the most serious operation issues that any future linear collider will face is that of the Machine Protection System (MPS). To produce useful luminosity the beam power and the beam densities must be very high. Unfortunately, these beams will almost certainly damage any material that is intercepted unless extreme care, such as that in the collimation sections, is taken. For example, in the 1 TeV NLC design, the beam power is over 8 MW and a single errant bunch train in the linacs would be sufficient to damage many unprotected accelerator structures. Obviously, this has severe implications on the beam operation during normal running as well as during tuning and commissioning.

The purpose of the MPS is to prevent damage to the collider components in the event of a routine failure or mistake. In addition, it should automatically provide a sequence of beam pulses that can be used as effective diagnostic tools during a startup or a fault period. Furthermore, after a fault, the MPS should be optimized to recover luminosity as quickly as possible in order to minimize the lost time. Of course, the system is *not* intended to provide comprehensive protection against any possible failure; the complexity of the MPS must be balanced against the cost, difficulty, and time for repair of the systems it protects.

The most serious challenge in the NLC is the prevention of `single pulse induced failure' (SPIF). This is component failure that occurs from an aberrant single beam pulse. Because it is impossible to know the precise trajectory of the upcoming pulse, the MPS must: 1) provide pulses that cannot cause SPIF for tuning and diagnostics and 2) insure that the difference between the upcoming pulse and the one that preceded it is within some limit, known as the `maximum allowable interpulse difference' (MAID), during normal operation. These two criteria form the basis of the NLC single pulse induced failure machine protection system.

Multi-pulse or `average power' induced failure is component failure that occurs after a succession of pulses deposit excessive energy on a given component. This type of failure is more familiar from SLC operation and is controlled in a similar fashion, *i.e.*, by using ion chambers, thermocouples, etc., to monitor beam loss.

In the next sections, we will outline the methods that are used to protect against SPIF—as noted the multi-pulse failure mode is protected with a more standard MPS and thus will not be discussed further. At this time, we only have a conceptual description; in the future, we will need to have a detailed solution on an element-by-element basis with greater margins than is outlined in this section. Furthermore, we have only considered protection in the main linac and downstream; we have not considered the MPS issues in the damping rings or bunch compressors although the principles will undoubtedly be similar. Finally, additional detail on the MPS can be found in Sections 7.8, 8.6, 8.7, and 9.2.2.

16.2 Single Pulse Induced Failure

The strategy to protect against SPIF is to use single bunch, nominal intensity, pulses for most diagnostic purposes and at any time the interpulse difference might be outside the MAID. This strategy results from the extreme energy density of the full intensity, multibunch, NLC beam. It is not practical to build mechanical systems that can withstand the nominal NLC beam except in isolated cases such as the collimation region. It is, however, possible to develop structures that can withstand the impact of single bunches, albeit with somewhat increased emittance. Once such structures are realized, the problem becomes one of ensuring that successive pulses are alike. Thus, the single-pulse protection (SPIF) can be subdivided into protection against a diagnostic bunch, a full current, high emittance single-bunch beam, protection against a multibunch bunch train, and the method of transitioning between the diagnostic and normal operating modes.

The protection against the diagnostic bunches is based upon a passive system consisting of thin spoilers which will increase the beam angular divergence so that, by the time the diagnostic beam strikes another component, it will not cause any damage; to prevent damage to the spoilers themselves, the emittances of the diagnostic beam must be increased by a factor of ten from the nominal beam emittance. Thus, the spoilers will allow the beamline to survive the transport of a single diagnostic bunch without regard to the state of the beamline hardware. All preliminary beam-based alignment, tuning, and diagnostics will be performed using a diagnostic bunch. Of course, if beam is being lost during the transport, these operations would need to be performed at low repetition rates to reduce the average power deposition.

After a diagnostic beam can be transported to the beam dumps without difficulty, *i.e.,* after establishing the initial beam-based alignment, the beam trajectory, feedback setpoints, and energy profile, the repetition rate can be increased to the nominal 120–180 Hz. At this point, the beam emittance can be decreased to nominal and additional bunches can be added to the bunch trains. To verify the beam loading compensation is set properly, the train length will be increased in steps. All subsequent tuning must be performed at the high repetition rate; only the diagnostic beam can be transported at low repetition rate.

To prevent the high rate beams from striking accelerator components, a trajectory window of roughly $\pm 200\,\mu$m and an energy window of $\pm 10\%$ are established about the nominal values; this is the MAID. If the beam deviates beyond these limits in any single pulse, the collider is returned to the high-emittance single-bunch diagnostic mode while the source of the problem is diagnosed from data that was taken during the errant pulse. Extensive logic will be used to prevent erroneous MPS faults due to bad BPMs readings.

This system relies on the fact that there are no transverse deflecting fields that can change sufficiently, within a single interpulse period (roughly 8 ms), to deflect the beams from the operational trajectory window into the accelerator structures or beamline elements. In most cases, this is attained by limiting the strength of all fast correctors and limiting the decay time of the quadrupole and bending magnet fields by using solid core magnets and thick conducting vacuum chambers. The few DC magnets whose fields could change too quickly will have to be interlocked directly to the MPS using a Hall probe or similar diagnostic—separate diagnostics, such as precharge monitor, will be needed for strong pulsed kicker magnets like the damping ring kickers.

In addition, the MPS must verify that the rf systems are operational and correctly phased before the beams are launched into the linacs. However, fairly large energy deviations can be tolerated. For example, a 20% energy deviation in combination with 100-μm random quadrupole misalignments, which are well in excess of what we expect, would only cause 1-mm orbit offsets in the linac.

Thus, the MPS system must only verify that 85% of the rf systems are operational and correctly phased. To this end, all modulators will be polled roughly 100-μs before beam time. At the same time, the klystron timing and phase information will be checked. If there are a sufficient number of failures, the beams will be aborted downstream of the damping rings; this verification procedure is described in greater detail in Chapter 8. Finally, there will be sacrificial spoilers and dumps located in the linac diagnostic station chicanes to prevent energy errors larger than 25% from propagating further down the linac; these are described in Chapter 7.

In the next section, we will describe the mechanical protection system and the diagnostic pulse generation. After this, we will discuss issues associated with the transition from the single diagnostic bunch mode to the full bunch train and then we will describe controlling the MAID.

	Be	C	Al	Ti	Cu	Fe
Radiation Length [cm]	35.7	21.7	9.0	3.7	1.4	1.8
dE/dx_{min} [MeV/cm]	3.1	3.6	4.4	7.2	12.8	11.6
Specific Heat [J/(cm^3 °C)]	3.3	1.9	2.5	2.4	3.5	3.8
Melting Point [°C]	1280	3600	660	1800	1080	1530
Stress Limit [°C]	150	2500	140	770	180	135
ΔT [°C]	3240	6530	6060	10330	12100	10520
$\Delta T/T_{melt}$	2.5	1.8	9.2	5.7	11.2	6.9

Table 16-1. *Spoiler material properties and temperature rise due to a single bunch of* 1.3×10^{10} *having emittances of* $\gamma\epsilon_x = 3 \times 10^{-6}$ *m-rad and* $\gamma\epsilon_y = 3 \times 10^{-8}$ *m-rad at end of the 500 GeV linacs; these emittances are the smallest values possible and correspond to an rms beam size of* $\sqrt{\sigma_x \sigma_y} = 3.1\,\mu$*m.*

16.2.1 Diagnostic Pulse Protection

Spoiler Materials

The purpose of the spoiler system is to allow a nominal intensity, single bunch, beam to be transported throughout all systems at low repetition rate without concern for damage. As stated, this will be performed using thin spoilers. If the beam is steered sufficiently far off-axis to intercept an accelerator element, it must first pass through one or more spoilers. These will increase the beam angular divergence so that, by the time the single bunch beam strikes another component, it will not cause any damage.

The length of the spoilers is a trade-off between the increase in the multiple scattering and the heating due to the electromagnetic shower in the spoiler. The projected angular distribution of the beam after a spoiler can be described by a gaussian distribution [Particle Data Book]

$$f(\theta_{x,y})\,d\theta_{x,y} = \frac{1}{\sqrt{2\pi}\theta_0}\exp^{-\theta^2_{x,y}/2\theta_0^2} \qquad \theta_0 = \frac{13.6\,\text{MeV}}{E}\sqrt{t}\big(1 + 0.038\ln(t)\big) \tag{16.1}$$

where t is the length of material in units of the radiation length. Thus, we expect a 0.2 radiation length spoiler to increase the beam angular divergence of a 500 GeV beam by $11.4\,\mu$r; this is hundreds of times larger than the incoming angular divergence.

Table 16-1 lists properties for a number of different materials including the dE/dx_{min}, the specific heat at room temperature, the melting point, and the stress limit. The stress limit is based on the tensile strength, the modulus of elasticity, and the coefficient of thermal expansion for the material. When a beam strikes the material, there is a sudden local temperature rise that may create local thermal stresses. If the temperature rise exceeds the stress limit, micro-fractures can develop in the material. In addition, it has been observed in experiments, that if the local temperature rise exceeds four times the stress limit, the shock wave due to the thermal rise will cause the material at the surface to fail completely or "delaminate" [Walz 1973, Walz 1996].

In the spoilers, we are not actually concerned by micro-fractures or deformations that might develop when the temperature exceeds the stress limit. These will not degrade the performance of the spoilers and would likely be partly re-annealed with further heating. Thus, the allowed temperature rise is limited by either the melting point of the material or four times the stress limit at which point the material will fail catastrophically.

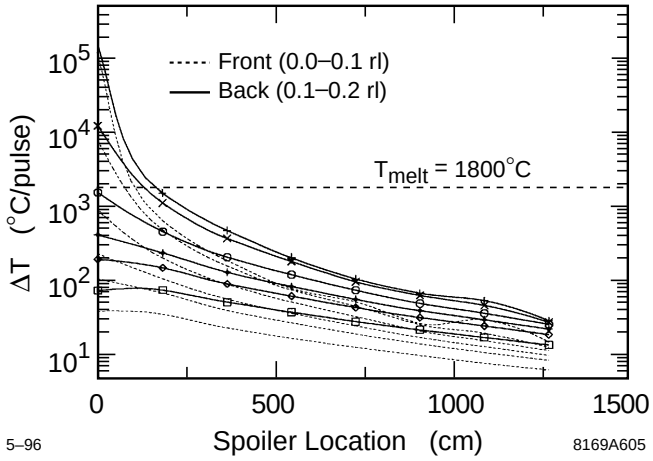


Figure 16-1. *EGS simulations of temperature rise in 0.2 radiation length long Ti spoilers with a single bunch of 1.3×10^{10}. The dotted and solid lines show the temperature rise in the front and back halves of the spoilers for six different incoming beam sizes: $\sigma_r = 0, 3.1, 10, 20, 30,$ and $50\,\mu$m.*

Table 16-1 also lists an estimate of the temperature rise ΔT assuming a single bunch of 1.31×10^{10} particles with a beam size of $\sqrt{\sigma_x \sigma_y} = 3.1\,\mu$m. This beam size corresponds to emittances of $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad at the end of the 500 GeV linac which are the smallest emittances that could be delivered from the damping rings and assumes that there are no emittance dilutions through the end of the linacs. The temperature rise is calculated from a simple analytic model:

$$\Delta T = \frac{N}{2\pi\sigma_x\sigma_y} \frac{dE/dx_{min}}{\text{Spec. Heat}} \tag{16.2}$$

which ignores the effect of the shower buildup and the variation of the specific heat with temperature.

In both the linacs and the beam delivery, the spoilers are constructed from titanium. Although other materials such as beryllium or graphite would be able to handle higher beam densities, titanium is the most practical choice, balancing the spoiler survival against the length of spoiler required and the vacuum and handling properties. Unfortunately, the surface temperature rise due to this low emittance beam is over five times the melting temperature of Ti. Thus, to prevent damage to the spoilers themselves, the emittance of the diagnostic beam must be increased significantly.

These analytic calculations have been supplemented with EGS simulations. In Figure 16-1, the maximum temperature rise is plotted in each of eight linac spoilers separated by 1.8-m for round gaussian beams having sizes of $\sigma_r = 0, 3.1, 10, 20, 30,$ and $50\,\mu$m; it is thought that the round beam case will model a flat beam with similar density although it may slightly overestimate the temperature rise. The dotted and solid lines show the temperature in the front and back halves of the spoilers and the dashed horizontal line shows the melting temperature of 1800°C. Notice that the temperature rise in the back half of the spoilers is roughly twice that in the front half due to the buildup of the electromagnetic shower.

Clearly, all spoilers downstream of the first will survive a single pulse of any incoming beam size. But, the first spoiler will be damaged unless the incident beam has an rms size greater than $\sqrt{\sigma_x\sigma_y} > 10\,\mu$m. Thus, to prevent damage to the spoilers, either both the horizontal and vertical emittances of the diagnostic beam must be increased by a factor of ten or the vertical emittance could be increases by a factor of 100 to yield the required beam size.

Spoiler Placement

In the linacs, the spoilers are primarily needed to protect the accelerator structures since these have the smallest aperture while, in the beam delivery, the spoilers are needed to protect the vacuum chamber and magnets. Unlike the spoilers, it is important to limit the temperature rise of the other accelerator elements to a value below the stress limit listed in Table 16-1. Deformations or micro-fractures in the accelerator structure irises would probably increase the multipactoring and the dark current from the structures and could distort the acceleration field patterns. Similarly, deformations of the magnet poles or coils could lead to large multipole fields or shorted coils.

We can get a first estimate of the requirements by looking at Table 16-1. Here, the temperature rise was calculated assuming a diagnostic pulse with an rms beam size of 3.1 μm. In this case, the temperature rise in Cu is estimated to be over 12000°C while the stress limit is 180°C. Thus, to decrease the temperature to a more reasonable value, the incident beam density must be decreased by a factor of 70. Actually, the requirements are much greater because of the electromagnetic shower. In practice, we need to decrease the beam density of a 500 GeV beam by over a factor of 3000.

Now, to determine the spoiler placement, we need to determine the failure scenarios. Ultimately, we will have to consider the failure modes on an element-by-element basis but at this time we will only consider three global scenarios:

1. First, as a "worst case scenario," we consider a FODO channel where a large deflection arises at the focusing quadrupole; a FODO array describes most of the linac beamline and much of the beam delivery beamline. Such a deflection could arise if one or more poles of the quadrupole magnet becomes shorted or if the magnet mover runs away to an extreme value, typically limited to ± 1 mm. This deflection will offset the beam in the next defocusing magnet which then deflects it further and possibly into an accelerator element.
2. Most other scenarios such as large energy errors or smaller amplitude deflections will cause the beam to oscillate at large amplitudes before being lost into an element.
3. Finally, there are a number of special situations which are not covered by either of these two cases. In particular, failures of the bending magnets or some specialized strong quadrupoles could directly drive the beam into an element with a very large deflection angle.

Unfortunately, there is no way to provide passive protection against this third case. Thus, these components will have to be directly monitored and interlocked to the MPS. In contrast, the second case is straightforward to protect against by placing a number of spoilers along the beamline with apertures smaller than that of all other elements. This leaves the first case which is more difficult to passively protect than the second case because there is relatively little distance in which the beam size can be increased before it intercepts an accelerator element. Of course, we could adopt the solution used for the third case, *i.e.,* direct monitoring, but this is less desirable than a passive system because of the large number of elements that would have to be monitored and the potential reliability problems.

To evaluate the requirements to protect the first case, we can calculate the distance from the defocusing quadrupole to the position at which the beam strikes the accelerator structure or the vacuum chamber ΔL:

$$\Delta L = \frac{L_c}{2} \frac{(r_a/r_{QD} - 1)}{1 + 2\sin\psi/2} \ . \tag{16.3}$$

Here, ψ_c is the phase advance per FODO cell, L_c is the FODO cell length, and r_a and r_{QD} are the aperture of the chamber/structures and the trajectory offset in the defocusing magnet.

Assuming a phase advance of 100° per cell, we find that, to shadow all the downstream elements, spoilers located at the quadrupoles would have to have a radius roughly $\frac{1}{4}$ the minimum radius along the beamline. Although this may

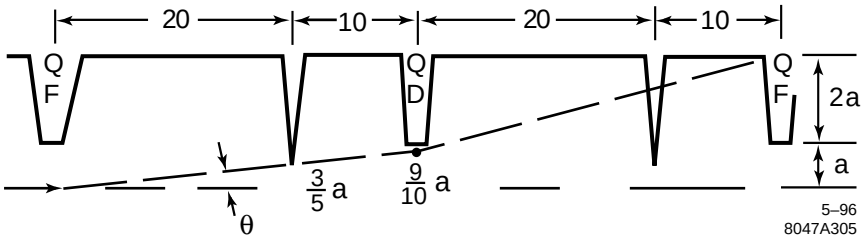


Figure 16-2. *Possible trajectories of a single bunch through a beam line as a result of completely or partially shorted quadrupole legs. Every trajectory encounters a spoiler at least 10 m before reaching the wall or another element. The quadrupoles are assumed to be spaced at 30 m and have a focal length of 20 m in this example.*

be the simplest solution, it implies very small aperture spoilers with correspondingly large wakefields; the transverse wakefield of a single spoiler would be comparable to that of an entire 1.8-m accelerator structure.

This becomes simpler in the beam delivery beamline where the vacuum chamber aperture can be increased to relatively large values between the magnets. In this case, the spoilers can also have relatively large radii which reduces the wakefields. Then, the placement of the spoilers must be chosen so they are located sufficiently far from the magnets, where the vacuum chamber constricts again, so that the spoiled beam will not damage the elements. This is illustrated schematically in Figure 16-2 where 0.25 radiation length Ti spoilers have been placed 10 m from the magnets to increase the beam size to roughly $\sqrt{\sigma_x \sigma_y} \sim 140\ \mu$m. The magnets and vacuum chamber will be further protected with sacrificial absorbers located immediately upstream of the chamber constriction. Both the spoiler placement and the absorbers are discussed further in Chapter 9.

The linac is more difficult to protect and the solution is different for the beginning and end of the linac. At the beginning, the low energy end, the quadrupoles are only separated by a few accelerator structures. In this case, a spoiler with an aperture radius of 1 mm, $\frac{1}{4}$ of the iris radii, would increase the total transverse wakefield seen be the beam an unacceptable amount; as stated, each spoiler would have a transverse wakefield comparable to that of an entire accelerator structure.

Fortunately, at low beam energy the angular divergence due to the spoilers is relatively large, θ_0 scales inversely with the beam energy, and the required spot size to prevent damage to the structures is relatively small. Thus, the spoilers can be placed close to the location that the beam first strikes a structure. This is illustrated in Figure 16-3 which shows the temperature rise in the irises of an accelerator structure due a 30 GeV single bunch with 1.3×10^{10}. The four different curves correspond to initial beam sizes of σ_r = 10, 20, 30, and 50 μm. Notice that, with an incoming 30 μm spot size, the temperature remains less than the stress limit of 180°C. A 30 μm spot size can be attained at 30 GeV by passing through a 0.2 radiation length spoiler 15 cm before striking the structure.

The picture is completed by looking at the failure modes. The magnet movers have a maximum range of ± 1 mm while in typical quadrupole designs, if a pole is shorted, the field on axis is roughly $B_{pole}/5$ and causes a 45° angle deflection of the beam.

Assuming a 100° cells, which is stronger focusing than in most of the linacs, a 1 mm offset of the first quadrupole would cause a 1.5 mm offset in the second quadrupole. In the linac, this beam would then strike the accelerator structure irises, having a 3.9 mm radius, after 60% of the distance to the next quadrupole. Similarly, assuming a single shorted pole of a linac quadrupole with a 7 mm bore radius, the beam is also offset in the next quadrupole by 1.5 mm in both the horizontal and vertical planes and again the beam will strike the structures after 60% of the distance to the next quadrupole.

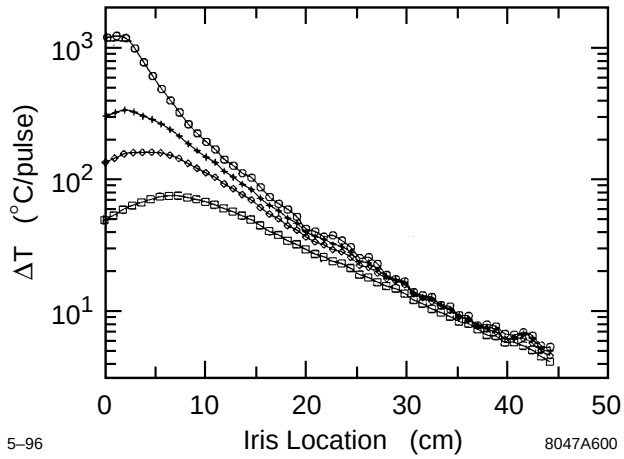


Figure 16-3. *EGS simulations of the temperature rise in the accelerator structure irises due a 30 GeV single bunch of* 1.3×10^{10}*; the different curves correspond to initial beam sizes of* σ_r *= 10, 20, 30, and 50* μ*m.*

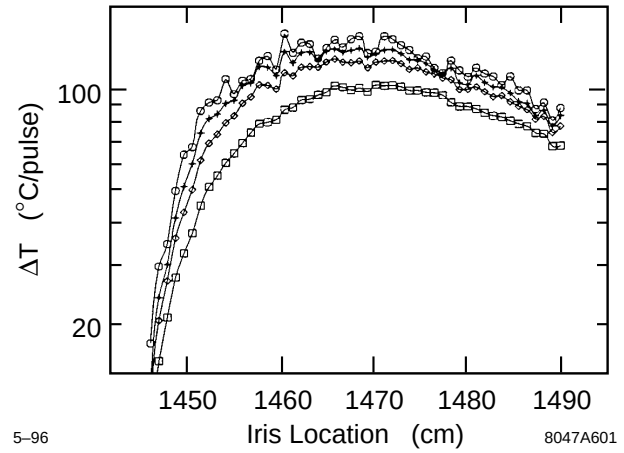


Figure 16-4. *EGS simulations of the temperature rise in the accelerator structure irises at the end of the 500 GeV linac due a single bunch of* 1.3×10^{10} *that has passed through eight spoilers separated by 1.8 m; the four curves correspond to initial beam sizes of* σ_r *= 10, 20, 30, and 50* μ*m.*

Both of these cases, are protected against by placing 0.2 radiation length spoilers with 2 mm radii at the end of every accelerator structure. As discussed in Chapter 7, the transverse wakefield deflection of the spoiler is roughly 10% of that due to an accelerator structure. Provided that the spoilers could accurately mounted to the structure ends, they would be aligned along with the structures as discussed in Chapter 7.

This solution works well until the beam energy is above roughly 400 GeV. At this point, the scattering due to the spoilers is small and a large amount of energy is contained in the electromagnetic shower. Thus, to be effective, the beam must intercept the spoilers a long distance before striking the irises. An example is illustrated in Figure 16-4 which shows the temperature rise in the irises due to a 500 GeV bunch of 1.3×10^{10} that has passed through eight 0.2 radiation length spoilers separated by 1.8 m. The four curves correspond to initial beam sizes of σ_r = 10, 20, 30, and 50 μm.

Unfortunately, in both cases that were considered for the beginning of the linac, a single shorted pole or a 1-mm offset, the beam would only intercept six spoilers before striking an accelerator structure; this would lead to a temperature that is roughly a factor of two above the stress limit.

There are two solutions: first, we could adopt solution (3), direct monitoring of the magnets, or, second, we could use a single spoiler with a 1 mm radius located at the quadrupoles. Although the transverse wakefield of a spoiler with a 1-mm radius is roughly five times greater than that with a 2-mm radius, the single small radius spoiler would replace ten of the larger radius spoilers since there are ten accelerator structures between the quadrupoles at the end of the linac. Thus, the integrated wakefield is the same. Additional benefit is gained by placing these elements before the quadrupole magnets since a spoiler will induce a large energy spread in the beam which leads to a large decrease in the beam density after the beam is deflected by the subsequent quadrupole. Finally, the spoiler could be directly mounted on and pre-aligned to the quadrupoles; thus, they would be aligned to the beam when the quadrupole beam-based alignment is performed. At this time, we have not determined the best option and will reserve that decision for the future.

Diagnostic Pulse Emittance Enlargement

As noted, to prevent destruction of the protection spoilers, the diagnostic beam emittance product $\epsilon_x \epsilon_y$ must be increased by a factor of 100. This enlargement must be turned-on and -off at the full repetition rate of 180 Hz without significantly steering the beam. The logical place to perform the enlargement is either in or just after the damping ring. At this time, we have considered two solutions: first, we could induce a large horizontal and/or vertical dispersion oscillation in the damping ring wigglers using pulsed quadrupoles. These would be located in the dispersion suppressor sections at the end of the arcs. As an example, to increase the vertical emittance by a factor of 100, skew quadrupoles can be used with an integrated strength of 7 kGauss—these change the tunes by less than 0.01 while increasing $\gamma\epsilon_y$ to 2.5×10^{-6} m-rad. Second, we could use a pulsed chicane, at the exit of the damping ring, to direct the beams through a gas filled chamber. Passing the beam through 1-m of Ar at a pressure of 1 Torr will increase the emittance product by 100. Alternately, one could use a thin foil (50 μm) of Be although in this case the power density may be a problem.

16.2.2 Transition between diagnostic pulses and full beams

As discussed, a diagnostic beam, a high-emittance full-current single bunch beam, will be used to recover from a fault or to start up after an off period. To then establish full current operation requires a sequence of steps that include ramping the repetition rate, the emittance, and the number of bunches. Specifically, after the diagnostic pulse can be cleanly transported to the beam dump, the repetition rate must be increased to the nominal full rate. This is necessary since the MAID is only guaranteed for short periods. Next, the beam emittance can be decreased to the nominal and finally the bunch train can be lengthened. It is this last step that is probably the most difficult because of the changes in beam current and beam loading which can have significant dynamical and thermal effects in the damping rings and linacs. A single bunch diagnostic pulse was chosen since it's transport dynamics are, in many ways, similar to that of the full bunch train. Nevertheless, the transition from the diagnostic pulse and full intensity operation must be automatically checked at several steps along the way.

Linac Operations

When transitioning from the single bunch beam to the full bunch train care will be needed to verify the beam loading compensation. For the full bunch trains, the beam loading is roughly 25% in the main linacs. Thus, if the loading compensation is not set, the trailing bunch will have an energy roughly 25% too low.

To verify the loading, we will increase the number of bunches per train in steps. At this time, it seems reasonable to increase the number of bunches from a single bunch to 10 bunches per train. In this case the maximum energy deviation should be less the 3%, even if the loading compensation is not set. At this point, the loading profile can be accurately determined and thus it seems reasonable to increase to a full train in the next step; details of the beam loading compensation and control can be found in Chapter 8.

At each stage during the ramp the collider should operate for enough time to allow the beam-based feedback systems to stabilize. The time required should be less than a second (100 pulses).

In addition, active feedforward will be required to control the rf power delivered to the accelerator structures. Without changing the temporal profile of the rf power to the structures, they will cool significantly during full bunch train operation and will heat up after an MPS fault when operating with a single bunch. This arises because the full train absorbs roughly 25% of the rf power. As described in Section 8.7, the solution to this problem is to adjust the klystron phase profile so that the additional power from the rf pulse compressors is sent to the loads and is not delivered to the structures.

Damping Ring Operation

The damping ring rf system has been designed to accept full current trains at the maximum repetition frequency but it is heavily beam loaded. Thus, changes in the total current must be accommodated with the appropriate beam scheduling information which includes the repetition frequency and a measure of the incoming charge which is required by the rf feedforward algorithm.

There are two sequences that can be used to ramp the rings from zero to full current. The first is used after an extended off period or for initial checkout at reduced current. The second is used for ramping to full current quickly as required, for example, following a spurious MPS ion chamber trip. In either case, the number of bunches per train will be increased in steps from 1 bunch per train (bpt) to 10 bpt and then to 90 bpt.

In addition to the feedforward algorithm to adjust the rf systems for changes in the beam current, feedforward will be needed for thermal regulation. The damping rings emit an enormous amount of power as synchrotron radiation and higher-order modes which is removed by the cooling systems. Thus, the feedforward system, likely consisting of both flow control and heating elements, will be required to prevent significant thermal fluctuations during changes in the stored beam current.

Recovery procedure after extended off period For single bunch operation, beam loading of the rf cavities in the damping ring may be neglected. However, the cavity loading angle must be adjusted to at least partially minimize the reflected power during extended periods of low current operation. Because the reflected power depends only weakly on loading angle with 1 bpt, the tuners positions are moved to minimize reflected power at 10 bpt. This allows for efficient ramping of the current from 1 to 10 bpt. The injection procedure is given below:

1. With the beam off and the extraction kicker deactivated, the tuners are adjusted to $\phi_{z,10}{}^{*}$, which is the tuning angle for which the reflected power is minimized at 10 bpt.

2. Since feedforward for reduced beam loading is not required, the tuner loops are left closed and the ring is sequentially filled with 4 trains of 1 bpt.
3. The extraction kicker is activated and the ring is operated for as long as desired with single bunches using single-turn extraction and injection.
4. The repetition frequency is then increased, if desired.
5. The number of bunches per train is then increased to 10 bpt. The principle is the same as that for full-current (90 bpt) injection, with the exception that the voltage and phase offsets for feed-forward are adjusted for 10 bpt using Eqs. 4.64 and 4.65. (If the current of the incoming beam is known, then the offsets are automatically adjusted by the feedforward controller.)
6. To resume operation at 90 bpt, the beam is turned off and the tuner setpoints are adjusted to ${\phi_{z,90}}^*$ and the rapid recovery procedure is followed. (Here ${\phi_{z,90}}^*$ is the tuning angle for which the reflected power is minimized at full current—Section 4.5.3.)

Rapid recovery procedure The fastest possible recovery from no beam to full current operation requires a minimum of 12 cycles (or 66.7 ms at 180 Hz); the actual process will be slower to allow the feedback systems sufficient time to stabilize between changes. The procedure is as follows:

1. With the tuners fixed at ${\phi_{z,90}}^*$, the extraction kicker deactivated, and feedforward deactivated, inject 4 trains of 1 bpt.
2. Activate the extraction kicker and feed-forward. Begin single-turn extraction and injection. Inject 4 trains of 10 bpt.
3. Inject 4 trains of 90 bpt. (The presence of the 10 bpt trains in the ring is taken in into account in the feedforward algorithm.)
4. Close tuner loops and proceed with nominal operation.

16.2.3 Controlling the Interpulse Difference (MAID)

When operating with more than one bunch or when operating with the nominal beam emittances, the MPS must guarantee that the trajectory of the upcoming pulse is within the MAID. In order to do this, every device that can change the beam energy or trajectory by more than the MAID in the interval between pulses must be checked before allowing the permit to the scheduler. For transverse deflecting magnetic fields, the best way to do this is to actually prevent those fields from making large changes during the interpulse period. This is practical for all but the strongest magnets since the typical magnet L/R decay times can be made ~ 200 ms. High bandwidth magnets, such as linac fast feedback correctors, must not have enough strength to exceed the MAID. One consequence of this is that low repetition rate, full train intensity operation, is not possible.

Fast, powerful, pulsed dumper magnet and extraction kicker systems must provide an electronic warning of their behavior prior to each pulse. This warning must be timely enough to stop beam extraction from the damping ring (in the case of the linac) or to fire protection dumper magnets.

The linac klystrons must also provide a summary status signal that can be used for MPS. Since the NLC linac MAID requires the energy difference to be less than 10completion of these tests for each klystron is not required. The

klystron pre-pulse checking system must be most effective for warnings about common mode failures that affect many klystrons. On a pulse-to-pulse basis, a klystron's amplitude, phase or trigger timing can vary outside nominal limits and both must be checked. The modulator switch high voltage will have a charging step in it's cycle that takes place about 100 μs before beam time. A comparison of that amplitude against the expected value will be used as an input to the MPS permit. The RF drive will also be tested for amplitude at about that time. The phase is more difficult to measure and will require that a phase measurement be made using an independent reference system.

To verify the beam loading compensation when going to the longer bunch trains, the train length will be increased in steps. Presently, we believe that the loading could be checked using only one intermediate train length, namely, going from one bunch per train to 10 bunches per train and then to a full 90 bunches per train; this needs further verification and increasing the steps has minimal implications for the rest of the systems. In addition, to prevent thermal changes of the accelerator structures when changing from low repetition rate to high rate and from the single- to the multibunch modes, the klystron phases are varied so that the addition power which normally would accelerate the missing bunches is dumped into loads rather than into the accelerator structures; this is discussed further in Chapter 8.

References

[Particle Data Book] Particle Data Group, "Particle Properties Data Book", *Phys. Rev. D* , **50**, 1173 (1994).

[Walz 1973] D. Walz, D. Busick, T. Constant, K. Crook, D. Fryberger, G. Gilbert, J. Jasberg, L. Keller, J. Murray, E. Seppi, and R. Vetterlein, "Tests and Description of Beam Containment Devices and Instrumentation–A New Dimension in Safety Problems", SLAC–PUB–1223 (1973).

[Walz 1996] D. Walz, "Justification for Temperature Rise and Thermal Stress Limits", NLC–Note–22 (1996).

Contributors

- Vinod Bharadwaj
- John Irwin
- Michiko Minty
- Ralph Nelson
- Tor Raubenheimer
- Sayed Rokni
- Marc Ross
- Dieter Walz

17

NLC Reliability Considerations

Contents

17.1 Goals

The SLC operates with an overall accelerator availability of about 80% [Erickson 1995]. The NLC will be nearly ten times as large and consume approximately six times the power of the SLC. Simple scaling of the SLC fault rates to the NLC results in an NLC which is effectively never operational. It is important that the issues of NLC availability be addressed from the onset of the design and engineering process so that the required component and system reliabilities are achieved. The goals of this chapter are threefold:

1. Establish an availability/reliability specification for the NLC on a machine basis (e^- injector, damping ring, main linac, etc.) and on a system basis (power supplies, magnets, klystrons, etc.). These specifications are arbitrary by nature but are to be compared, as far as possible, with the operational experience of existing accelerator complexes. An availability target of 85% for the full NLC has been adopted.
2. Develop a formal solution to the problem of how availability/reliability is to be accomplished. This requires shifting the responsibility for availability/reliability from a separate and detached upper-level oversight management team to those who are responsible for system development, engineering, implementation, and maintenance. In order to succeed, the concepts of reliability and availability need to be integral to the systems development and must receive necessary resources through a bottoms-up approach with top-down support and review.
3. Identify where reliability engineering effort should be initiated because of discrepancies between performance requirements and known behavior, in those areas where information is lacking, and where exorbitant costs are projected.

17.2 Reliability and Availability

Reliability is the probability that an item or system will perform the necessary function without failure for a given period of time. Reliability, $R(t)$, is characterized by the mean time to failure, $MTTF$. For a system of N_s identical components, the $MTTF$ of the system is taken to be $MTTF_i/N_s$ where $MTTF_i$ is the mean time to failure of an individual component. For the case of constant failure rate λ,

$$\lambda = (MTTF)^{-1} \tag{17.1}$$

and

$$R(t) = e^{-\lambda t} \tag{17.2}$$

A more reliable system thus lasts longer between repairs than a less reliable system.

Availability is the probability that a repairable system will be available for use when required. Availability for the NLC is defined as A,

$$A = 1 - MTTR/MTTF \tag{17.3}$$

wherein $MTTR$ is the mean time to recover which is the average repair time plus accelerator operations recovery time. In general, system availability is enhanced by high reliability and short repair and recovery times. The definition 17.3 is adopted for the NLC accelerator as representative of a complex system in which additional components continue to fail during the time in which recovery is being made for a previous fault. To show how availability of the NLC can vary with respect to that of the SLC, Eq. 17.3 is plotted in Figure 17-1 over a range of $MTTF$, normalized by the $MTTF$ of the SLC ($MTTF/MTTF_{\rm slc}$), for several values of $MTTR$, again normalized by the $MTTR$ of the SLC ($MTTR/MTTR_{\rm slc}$). In Figure 17-1 it is seen that the availability is increased by reducing the $MTTR$ or

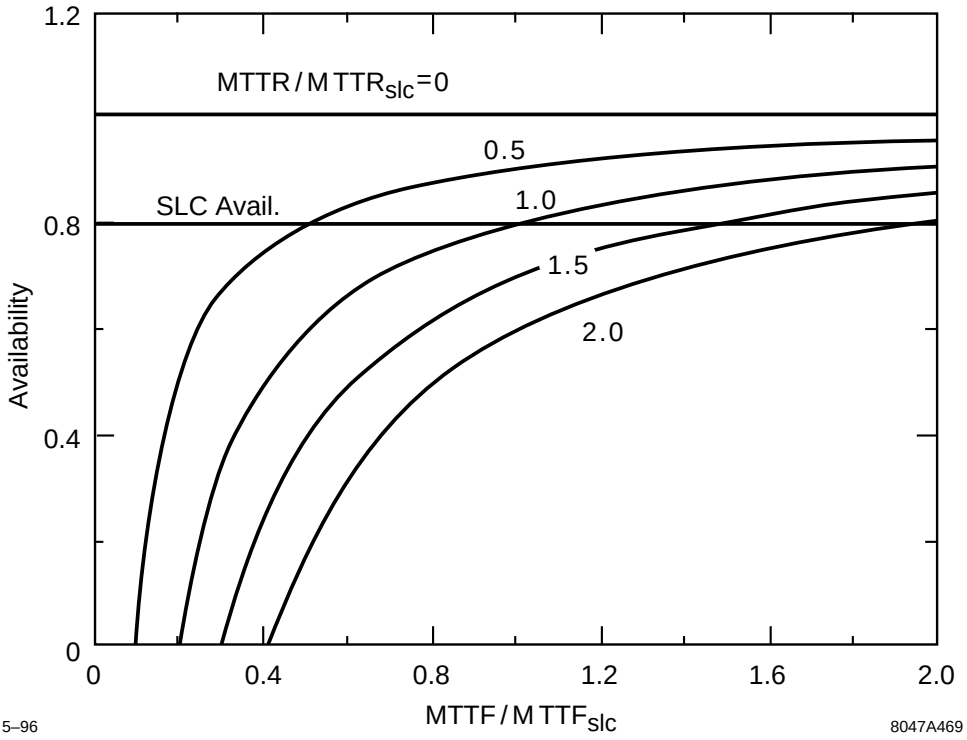


Figure 17-1. *Availability vs.* $MTTF/MTTF_{\rm slc}$ *for various values of* $MTTR/MTTR_{\rm slc}$.

alternatively increasing the $MTTF$ for a given $MTTR$. If the $MTTF$ in the NLC equals that of the SLC (albeit the increased number of NLC components) while the $MTTR$ increases by a factor of two due to say travel time, the availability drops to 60% compared with the value of about 80% for the SLC. Also from Figure 17-1 it is seen that if the $MTTR$ exceeds the $MTTF$, availability drops to zero. This has been the experience with SLC operations.

When discussing system performance, reliability is most often used as a figure of merit. For the machines which comprise the systems, availability is the appropriate figure of merit. As noted above, availability can be enhanced through high reliability. Fortunately, availability can also be improved through reduced repair and recovery times and through component redundancy.

17.3 Target NLC Availability

A target of 85% NLC availability over a scheduled running cycle of 6500 hours is assumed. This is a running period of nine months on and three months off in a calendar year. Numerous short maintenance and repair periods erode the time allocated for machine operations. One-shift-per-week maintenance is a 5% cost to operation. In a strict accounting view, one shift per week of scheduled maintenance during the nine-month cycle leaves only 10% of the time to be allocated to unscheduled outage, from all causes. Experience at SLAC indicates that with the exception of some of the utility installations, very few of the accelerator components have a preventive maintenance program which require scheduled outage during a nine-month cycle. Most, if not all, scheduled maintenance tasks can be accomplished during an annual three-month down. The bulk of the eight-hour scheduled outages taken during a running cycle are used to

accomplish remedial repair tasks which have accumulated during the period since the previous outage. These are most appropriately charged to unscheduled downtime.

It appears to be a straightforward task to design an NLC which has minimum maintenance requiring scheduled outages. The incremental cost to accomplish this is minimal given that so little of the present accelerator systems (both at SLAC and at other accelerator laboratories) have scheduled maintenance requirements. It is important, however, to identify those components which presently require periodic outages for maintenance and to reduce such requirements through judicious design and configuration modifications.

17.4 NLC Machine Availability and System Reliability

A proposed NLC availability specification has been developed for the NLC machines and systems for 85% availability over 6500 hours per year of scheduled operations. To develop this specification, the NLC has been divided into 12 machines: e^- source and linac, e^- damping ring and first compressor, e^- booster linac and second-stage compression, e^- main linac, e^- final focus and dumpline, and an identical breakout for the positron complex with the addition of the e^+ source and linac and e^+ pre-damping ring. Similarly, the NLC has also been divided into eight categories of systems: power supplies, magnets, klystrons, modulators, etc. When divided in the same fashion, the SLC consists of six distinct machines, each of similar complexity as an NLC counterpart; the SLC has the same eight categories of systems but with fewer components per system. The overall product of NLC machine availabilities is 85%; the overall product of the NLC system availabilities is 85%. A mean time to recover ($MTTR_s$) of one hour has been chosen for the systems. Equal weighting for each of the machines has been assumed except for the cases of the main linacs which are each given three times the weighting of the other machines. Table 17-1 lists the proposed availability specification for the various NLC machines and the assumed weighting factors. In Table 17-1 the listed availability is simply A_m,

$$A_m = 0.85^{w_m/16} \quad , \tag{17.4}$$

wherein w_m is the weight factor for a given machine, 16 is the sum of the 12 weight factors, and 0.85 is the target availability for the full NLC. Table 17-2 lists the proposed availability specifications for NLC systems. For the noted assumed $MTTR_s$, the required $MTTF_s$ for the system as a whole is given by

$$MTTF_s = MTTR_s / (1 - A_s) \tag{17.5}$$

where A_s is the listed system availability. For Tables 17-1 and 17-2, the allowed unscheduled outage is based on an assumed 6500 hours per cycle of scheduled operating time. Given the $MTTF_s$ for a system, the corresponding required $MTTF_i$ for an individual component is noted in Eq. 17.5.

Each of the 12 machines must be available 99% of the time (97% for the main linacs) in order to achieve the 85% availability goal. For a scheduled operating cycle of 6500 hours this allows for 66 hours of outage per machine per cycle (195 hours for each of the main linacs) The subtotal outage for the e^+ machines is greater than that of the e^- machines because of the added complexity of a positron production system and pre-damping ring.

A preliminary specification of NLC component reliability has been developed. The minimum $MTTF_i$ of the components which is needed to achieve the system availability specification is given by

$$MTTF_i = N_s MTTR_s / (1 - A_s) \tag{17.6}$$

wherein N_s is the number of identical components in a system, $MTTR_s$ is the mean time to recover of the particular system, and A_s is the specified availability for the system. As an example, for N_s = 1500, $MTTR_s$ = 1 hour, and A_s = 0.995, the required $MTTF_i$ = 300,000 hours.

Scheduled Operating Hours: 6500			
	Weight	Availability	Unscheduled Outage (hours)
e^- Inj, Source and Linac	1	0.99	66
e^- DR and Compressor 1	1	0.99	66
e^- Booster Linac and Comp. 2	1	0.99	66
e^- Main Linac	3	0.97	195
e^- Final Focus and Dumpline	1	0.99	66
Subtotal e^- machines:	**7**	**1**	**458**
e^- Inj, Source and Linac	1	0.99	66
e^+ Source and Linac	1	0.99	66
e^+ Pre-damping Ring	1	0.99	66
e^+ DR and Compressor 1	1	0.99	66
e^+ Booster Linac and Comp. 2	1	0.99	66
e^+ Main Linac	3	0.97	66
e^+ Final Focus and Dumpline	1	0.99	66
Subtotal e^+ machines:	**9**	**1**	**589**
Totals:	**16**	**0.85**	**1047**

Table 17-1. *Availability specifications for the NLC machines.*

NLC Systems	Availability	$MTTR_s$ (hours)	$MTTF_s$ (hours)	Unscheduled Outage (hours)
Power Supplies	0.975	1	40	163
Magnets	0.975	1	40	163
RF Systems	0.950	1	20	325
Motors	0.975	1	40	163
BPMs	0.990	1	100	65
Controls	0.985	1	67	98
Utilities	0.995	12	2400	33
Miscellaneous	0.995	1	200	33
Totals:	**0.85**			**1040**

Table 17-2. *Availability specification for the NLC systems.*

SLC machine and system availabilities and component $MTTF_i$ have been compiled for the 1992, 1993, and 1994/1995 SLC operating cycles. Operating experience of SLAC systems compares favorably with experience at Fermilab, CERN, KEK Photon Factory, Cornell, APS, and AGS. Table 17-3 lists the accelerator availabilities for physics of these various laboratories; the running cycles are noted. In general it was found that the same sorts of problems exist at all the labs. When the lengths of the running cycles are considered along with the sizes of the various machines and the peculiarities of the various accounting methods, the performance of the different accelerators are quite similar. Some labs do better with certain technologies than others but there are no clear differences on the whole. Because of the apparent similarities between the labs, it has been decided to base NLC technology expectations on SLAC experience, since the details of the SLAC data are more readily available at SLAC. It is important however to make comparisons with the other labs on a case-by-case basis when anomalies or uncertainties occur. On average, the six SLC machines (injector, two damping rings with compressor systems, e^+ source, linac, and arcs and final focus) each had an availability of approximately 97%.

Table 17-4 lists a preliminary parts count for the NLC. This information was taken from the NLC ZDR WBS [NLC WBS 1996]. For comparison purposes, Table 17-5 lists a parts count for the SLC. The data in Table 17-5 was gathered by counting entries in the SLC control system database. Initial counts of the numbers of NLC components indicate that there is about a factor of ten more components of all types in the NLC compared to a similar count of SLC components. Attention must be paid to improving the performance of NLC systems over that which is being achieved in existing systems of similar complexity.

17.5 A Formal Solution

Achievement of the specified NLC availability comes through the integration of the system and machine availability/reliability specifications into the component, system, and machine-functional specifications at the onset of the engineering design phase. Performance specifications of individual components will include the specification of reliability. The design review process must include attention to the availability/reliability requirements. A precision supply that never works is no better than an out-of-tolerance supply that never fails. Within a machine the availability budget must be respected. This task is best done at the engineering level but must be managed in the same fashion and at the same time that the more familiar performance criteria are managed.

Reliability engineering is a recognized discipline which plays an important role in all technologically-oriented industries (*e.g.,* semiconductor, aeronautics and astronautics, automotive, telecommunications, and power industries). There are a number of professional societies dedicated to developing the techniques and methodologies of reliability (*e.g.,* IEEE Reliability Society, Society of Automotive Engineers, Society of Reliability Engineers, Society of Logistic Engineers, American Institute of Aeronautics and Astronautics [RS IEEE, SAE, SRE, SLE, AIAA]). There are numerous annual meetings of these societies wherein tutorials on these methods are given in addition to the familiar conference presentations of topical issues (the Annual Reliability and Maintainability Symposium [ARMS 1996], for example). There are a large number of textbooks and courses on availability [Lewis 1996, O'Connor 1985]. It is important to take advantage of the tools developed and to apply them to the issues of NLC reliability. It is also necessary to understand the lessons learned in areas other than accelerators and to apply these lessons to the problems facing NLC construction. In many cases, the detailed solutions of how reliability in a Boeing 777 is achieved are not directly applicable to the NLC, but the thought processes going into developing a Boeing 777 are identical to what is required to successfully meet the NLC reliability goals.

For the NLC CDR, it is important that the issues associated with component reliability and system availability be fully integrated into the component and system engineering. Segregation of the discussion of availability into a separate chapter (in the CDR) will not fulfill the need to infuse the requirement for reliability beginning at the most basic levels of NLC design. If availability is to be achieved for a system which is nearly ten times larger than what has previously

Laboratory	Availability	Reference
ANL (APS) 95	68.30%	Argonne National Lab., Private Communication, Site Visit – R. Gerig, D. Ciarlette
CERN (SPS) 94	69.30%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 93	72.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 92	74.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 91	72.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 90	74.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 89	71.20%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
Fermi 91	72.64%	Fermi Accelerator System Tally Sheets, Site Visit – R. Mau
Fermi 92	65.86%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
Fermi 93–94	63.71%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
Fermi 93–94	63.71%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
SLAC (SLC) 92	81.00%	1992 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (SLC) 93	84.53%	1993 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (SLC) 95	80.87%	1994/95 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 92	87.01%	1992 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 93	93.25%	1993 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 94	93.33%	1994 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC SSRL 94	97.04%	SSRL, Private Communication, Site Visit – E. Guerra
SLAC SSRL 95	96.60%	SSRL, Private Communication, Site Visit – E. Guerra
AGS, FY95Q3	86.30%	Brookhaven National Lab, FY 95 3rd Qtr. Report – F. Weng
AGS, FY94Q4	86.70%	Brookhaven National Lab, FY 94 4th Qtr. Report – F. Weng
Cornell 91–92	74.10%	CESR Reliability Summary FY 1992–FY 1994 – D. Rice
Cornell 92–93	77.90%	CESR Reliability Summary FY 1993–FY 1994 – D. Rice
Cornell 93–94	84.00%	CESR Reliability Summary FY 1994–FY 1994 – D. Rice
KEK Photon Factory Linac 10/92–9/93	98.70%	KEK Operations Report FY 1992–FY 1993
KEK Photon Factory Linac 10/91–9/92	98.40%	KEK Operations Report FY 1991–FY 1992
KEK Photon Factory Linac 10/90–9/91	97.70%	KEK Operations Report FY 1990–FY 1991

Table 17-3. *Availabilities of several accelerator laboratories.*

	Pwr sup	Magnets	Klystrons	Modulators	Motors	BPMs	**Sys. Total**
e^- Inj, Source and Linac	245	229	16	16	0	381	**887**
e^- DR and Compressor 1	817	709	5	5	300	555	**2391**
e^- Booster Linac and Comp. 2	452	482	116	116	1077	291	**2534**
e^- Main Linac	736	756	2264	1132	14643	5300	**24831**
e^- Final Focus and Dumpline	871	1466	1	1	1344	472	**4155**
e^- Inj, Source and Linac	244	229	40	40	0	381	**934**
e^+ Source and Linac	236	241	32	32	0	81	**622**
e^+ Pre-damping Ring	700	700	2	2	300	300	**2004**
e^+ DR and Compressor 1	817	709	5	5	300	555	**2391**
e^+ Booster Linac and Comp. 2	452	482	116	116	1077	291	**2534**
e^+ Main Linac	736	756	2264	1132	14643	5300	**24831**
e^+ Final Focus and Dumpline	871	1466	1	1	1344	472	**4155**
NLC Total	**7177**	**8225**	**4862**	**2598**	**35028**	**14379**	**72269**

Table 17-4. *Preliminary NLC parts count for several systems.*

	Pwr sup	Magnets	Klystrons	Modulators	Motors	BPMs	**Sys. Total**
e^- Inj, Source and Linac	249	247	16	16	10	37	**575**
e^- and e^+ DRs and Compressors	40	456	5	5	6	199	**711**
e^+ Source and Linac	30	452	2	2	5	204	**695**
e^- Main Linac	608	608	242	242	22	283	**2005**
SLC Arcs	119	1000	0	0	912	978	**3009**
SLC Final Focus	192	192	0	0	23	59	**466**
SLC Total	**1238**	**2955**	**265**	**265**	**978**	**1760**	**7461**

Table 17-5. *SLC parts count for several systems.*

been achieved by the accelerator community, reliability must be fully accepted by the engineering and fully supported by the management.

Availability of the systems is based on the reliability of the individual components in concert with component configurations which include considerations of system repairability and redundancy. The solutions are specific to the particular systems; redundancy in the rf systems is a straightforward cost-effective solution, whereas component reliability combined with ease of changeability appears to be the proper solution for many of the magnet power supply applications.

17.6 Three Examples: Klystrons, Power Supplies, and Motors

In the main linacs, the expected $MTTF_i$ of the klystrons is 20,000 hours [Caryotakis 1995] and the $MTTF_i$ of the thyratrons is 10,000 hours [Wait 1996] Given an estimated count of 4000 klystrons and 2000 thyratrons in the NLC, approximately 1300 of each will fail and need replacement every cycle; this is a combined failure rate of one klystron or modulator every 2.5 hours. In order to operate the machines, on-line redundancy is required. By necessity, the

repair rate must be equal to or faster than the failure rate. Therefore, the availability for the rf system is simply A_{rf}

$$A_{rf} = 1 - e^{-1}/n! \tag{17.7}$$

where n is the number of redundant rf modules available for use when needed. For $n = 6$, $A_{rf} = 0.9995$. Present plans call for 3% redundancy in the number of rf modules which is quite sufficient. The rf systems are an operating cost issue but not so much one of availability. It is important to work to extend the $MTTF_i$ of the klystrons and thyratrons so as to reduce the cost of these consumables. It is worth noting, that effort must go into developing reliable waveguide valves to permit changing to klystrons during accelerator operations and to design the modulators such that the thyratrons can be easily changed.

There are approximately 750 quadrupoles per main linacs. The power supplies for these magnets are expected to be in the power range of a few kilowatts each. For the pair of linacs, the $MTTF_i$ of the power supplies is 300,000 hours to give system availability of 0.995, assuming the nominal one-hour $MTTR_s$. Should the $MTTR_s$ increase to two hours due to travel time or complexity of changing, the $MTTF_i$ increases to 600,000 hours. Rack-mounted power supplies in this power range used at SLAC have an $MTTF_i$ of about 300,000 hours [Donaldson 1996] and an $MTTR_i$ of about 1.5 hours. Whereas the present performance of similar power supplies meet the NLC goals, care must be taken to keep the $MTTR_s$ of less than one hour.

There are approximately 35,000 motors in the quadrupole and structure mover systems of the two NLC linacs. Since a stuck mover is a "soft" failure that contributes to emittance growth but does not stop the machine dead, it has been decided to allow 1% of the motors to fail each month before stopping to fix the accumulated failures. A failure rate of 1% per month corresponds to a $MTTF_i$ of 8.3 years. Motor manufacturers claim $MTTF_i$s of five to seven years for 100% duty factor usage and seven-to-ten-year $MTTF_i$ for 50% duty factor usage [Parker 1996, Warner 1996]. SLC experience has been quite good with motors. However, it will be important to design the movers with motor replaceability in mind since 1% per month failure rate is 3500 failures per year and the $MTTR_i$ needs to be small (on average 350 motors need to be replaced each month during a "short" machine access).

17.7 Summary

Simple scaling of the SLC fault rates to the NLC results in an NLC which is not operational. Reliability and availability need to be fully integrated into the functional requirements of the NLC. Reliability and availability must be explicit at the component, system, and machine levels in the CDR as a natural and normal part of the accelerator design. Real consideration and effort must be dedicated to defining and solving the reliability issues. The solutions to these issues necessarily arise from the engineering teams charged with building the systems. There exist significant engineering disciplines dedicated to addressing the issues, but care needs to be taken such that the correct solutions are properly applied to the relevant problems.

References

[AIAA] American Institute of Aeronautics and Astronautics, 370 L'Enfant Promenade, SW, Washington, DC 20024-2518 USA.

[ARMS 1996] *1996 Proceedings* and *1996 Tutorial Notes*, Annual Reliability and Maintainability Symposium, c/o Evans Associates, 804 Vickers Ave., Durham NC 27701-3143 USA.

[Caryotakis 1995] G. Caryotakis, personal communication (1995).

[Donaldson 1996] A.R. Donaldson, personal communication (1996).

[Erickson 1995] R. Erickson, C.W. Allen, T.K. Inman, W. Linebarger, M. Stanek, SLAC–PUB–95–6895, in "SLAC Accelerator Operations Report: 1992–1995", *Proc. of the 1995 Part. Acc. Conf.*, Dallas TX (1995).

[Lewis 1996] E.E. Lewis, *Introduction to Reliability Engineering* 2nd ed., (John Wiley & Sons, Inc., New York, 1996).

[NLC WBS 1996] NLC Project WBS and Outline, Version 3–20–96 (1996).

[O'Connor 1985] P.D.T. O'Connor, *Practical Reliability Engineering*, 2nd ed., (John Wiley & Sons, Inc., New York, 1985).

[Parker 1996] Parker Hannifin Corp., Compumotor Div., 5500 Business Park Dr., Rohnert Park CA 94928 USA, personal communication (1996).

[RS IEEE] Reliability Society IEEE, Institute of Electrical and Electronic Engineers, PO Box 1331, Piscataway NJ 08855-1331 USA.

[SAE] Society of Automotive Engineers Inc., 400 Commonwealth Dr., Warrendale PA 15096-001 USA.

[SLE] Society of Logistic Engineers, 8100 Professional Place, Suite 211, Hyattsville MD 20785-2225 USA.

[SRE] Society of Reliability Engineers, 2002 Harpoon Dr., Stafford VA 22554 USA.

[Wait 1996] G.D. Wait and M.J. Barnes, "Thyratron Lifetimes, A Brief Review", *Proc. of the 1995 Second Modulator-Klystron Workshop*, SLAC (1995); SLAC Report 481 (1996).

[Warner 1996] Warner Electric, Linear Motion and Electrical Div., Superior Electric, 383 Middle St., Bristol CT 06010 USA, personal communication (1996).

Contributors

- C.W. Corvin
- A.R. Donaldson
- J.W. Humphrey
- W.A. Linebarger
- M.H. Munro
- A. Saab
- J.C. Sheppard

18

NLC Conventional Facilities

Contents

18.1 Introduction

The NLC conventional facilities integrate the geotechnical and topological conditions of a generic site with a sub surface linac housing and klystron gallery structure. Conventional above-grade infrastructure and support facilities generally found in high-energy physics laboratories are included along with other necessary items associated with such a very large project. A non-specific site is assumed as no site has yet been selected. Many detailed specifics regarding the conventional facilities are not as yet addressed. The goal for this report has been to quantify the general ideas to a degree adequate to define the starting point for a detailed bottoms-up conceptual design report effort. Figures 18-1 through 18-5 show details of the NLC facility and are found at the end of this chapter.

18.2 Site

The overall NLC site is approximately 32 km in length and two km in width with four distinct functional areas; detectors, campus, injectors and linac. The site is assumed to be located such that reasonable access to adequate power, water, transportation and housing is available. The center of the site would be the location of the interaction point and the detector facilities. This central location might also be the location of the campus area, however the best campus area will very likely be selected with local access and transportation considerations being the driving factors. Opposite ends of the site would each have an injector area with a damping ring. At just one of the injectors a positron target and a pre-damping ring would be added. Between the detector area and each of the two injector areas would be a linac housing and a klystron gallery.

The linac areas with their associated utility support structures would make up the largest single portion of the overall site conventional facilities

The NLC site geology and topology should be such that both tunneling and cut-and-cover methods of construction are practical and utilized. A site having a high proportion of competent rock would be ideal and is the preferred site considered for the NLC. Excellent work with respect to geology, topography and site selection was done for the SSC project and has been used to advantage for the NLC. Attachment C to the SSC Conceptual Design Report, dated March 1986, documents that earlier work.

The NLC is estimated to require 300 MW of electric power to be supplied from two 230-kV transmission lines running parallel to the NLC. Three main substations are planned, one near each end of the site and one in the center. The primary distribution voltage from the main substations along the linac would be 34.5 kV. Two hundred small unit substations, adjacent to or in utility clusters along the klystron gallery, would supply the appropriate utilization equipment distribution power. The site water-cooling system would consist of approximately eight forty-MW cooling towers distributing water to utility cluster heat exchangers that transfer heat from ten-megohm low-conductivity water systems. Total site water demand would be about 3,000 gallons per minute and would come from a combination of existing offsite sources and those developed onsite. Conventional site facilities would include roads and parking for about 1400 vehicles as well as fire protection, water and waste treatment, communications, cable plant, and construction-related mobilization and site preparation. The primary roadway to the site would be four 12-ft lanes to support heavy assembly deliveries. The secondary linac roadway would have two 12-ft lanes with shoulders adequate to stabilize the roadway base.

For the major infrastructure construction phase of the project a concrete batch plant is assumed to be provided on site with a rail siding extension to bring in the associated bulk material needed for concrete. This rail extension would be of use later for locating modular installation and maintenance clusters associated with NLC operations. These clusters

would be 40 and 20-ft transportainers loaded with tools, stored components, field offices, and mini shops. The modular clusters would be relocated off-shift to follow and support on-shift operations, installation and maintenance.

18.3 Campus

The campus area is planned to house and support about 2500 persons and provide the needed facilities to build, operate and maintain the NLC. It consists of 16 buildings, the largest being a four-story main laboratory building of 350,000 square feet. This building would include a 1000-seat auditorium, a cafeteria, conference rooms, offices and light electronics laboratories as well as a central computer facility. This building would house the central administrative area for the laboratory.

Six heavy-fabrication buildings with high bays, office mezzanines, light and medium cranes, would be needed in the industrial area of the campus together with three smaller shops buildings. The shop buildings would be used for various machining and assembly operations. These industrial buildings would total roughly 210,000 square feet and have an adjacent paved outdoor staging area for larger assembly and preparation. Two conventional warehouses will be needed during and after construction for storing the various materials and assemblies that will make up the NLC. These structures would total about 80,000 square feet in area. The campus area would also have a vehicle service area, water and sewage treatment facilities and an emergency rescue and services facility.

18.4 Injectors

Each electron injector facility will include an injector with a short linac section, a 715-ft-circumference damping ring, and a compressor section with a second linac section. The positron injector will also have a pre-damping ring and a positron target. These facilities will be rigid-frame box structures of cast-in-place concrete. They will be positioned at a depth to align with the arcs at the ends of the main linac beam tunnels. The roof spans in the damping rings are substantially greater than for the bored tunnels. A hard rock site, mined with a horizontal cutter wheel machine, would be ideal for the damping-ring enclosures. Tunnel sections between the injector and damping-ring enclosures may be bored. Structures for the supporting injector utility clusters would be somewhat similar to those used for the main klystron gallery and linac. Should the campus area be located adjacent to one of the injector areas, then the supporting utility, power supply and instrumentation structures could be installed at the surface. The injectors will have local control rooms that slave to the main NLC control room.

18.5 Linac

The linac housing tunnel is planned to be bored with a tunnel boring machine at an average depth of forty ft. Its total length is a maximum of about twenty miles, less sections where the beam passes through cast buildings and enclosures such as the detector area and the damping rings. The finished inside diameter is planned to be 12 ft, with a concrete or shotcrete liner and a concrete floor or invert. A hard rock site might possibly require less concrete depending on rock strata and water migration. Utility chases from the linac housing to the klystron gallery and utility cluster alcove above will be installed at 156-ft intervals along the linac to provide water cooling and other utilities. The minimum earth cover between the ceiling of the linac beam housing and the floor of the klystron gallery above is 20 ft for adequate personnel radiation shielding. The klystron gallery above will enclose klystrons, modulators,

power supplies, instrumentation and controls. It will extend over most but not all of the beam housing tunnel below, as klystrons and related rf penetrations are not required along the entire length of the beam housing tunnel. A minimum of twelve miles of klystron gallery is planned for the rf systems. It is planned to be 30 ft in width. The utility cluster alcoves, at 156-ft intervals, will enclose motors, pumps, cooling and power system panels and controls.

The klystron gallery and utility cluster alcoves would be constructed using cut-and-cover methods and be below the surface at a depth sufficient to distribute surface traffic loads without the need for an excessively thick roof structure. Besides the utility chase housing penetrations, the housing will have 24-in inch diameter rf waveguide penetrations at six-m intervals along the beam housing. The utility chases will contain all utilities, including low conductivity water and cables, but would exclude rf waveguide. Housing and gallery entrances for vehicles and personnel are planned to be spaced at one-km intervals with adequate access ventilation and lock-up baffles to restrict air movement during linac operations.

The final-focus areas will extend from the detector building to the interaction point switch wyes at the ends of the linac where the beam is turned gradually to allow for two IPs and two detectors. The wyes are approximately two km from the interaction points. The two interaction points will be 40-m apart in the detector building making a single building feasible to house two detectors. A total of eight muon spoilers will be distributed in the four final-focus sections. Sizable structures to support and install these will be required. Their nominal dimensions fill the 12-ft inside diameter of the bored linac tunnel for three m along the beam z axis.

18.6 Detectors

The detector building facility will require about 100,000 square feet of total area. It will have a high bay area with a heavy bridge crane common to the entire facility. Two pit areas will match the two detector bore elevations to the final-focus beam line elevations. The pit will have light concrete sidewalls with earth tiebacks for support. The two pits will be structurally isolated from each other to segregate respective motion and vibration. Mezzanines will be included for computer, control room and office facilities. Areas for essential power conversion, low conductivity water, and cryogenic equipment will also be included both within and about the detector building. Shop areas will be provided for welding, metal fabrication, and instrument testing and calibration.

The two detectors will weigh in the order of 50,000 tons each and will require a precision steel and concrete floor pinned to bedrock for support. The detectors are assumed to be constructed in sub-assemblies that are skid, jacked and hoisted into place. The design of the detector building is driven by the design of the detectors themselves as the orientation and sequencing necessary for detector assembly must be consistent with the building layout. The detector control room will operate both detectors along with the instrumentation and data acquisition necessary for the relevant experimentation. Detector operation and control will be both local and remote to the NLC main control room. Kitchen, shower and rest facilities will be included to support personnel around the clock. Table 18-1 lists attributes of the interaction building, as well as those of the balance of the NLC conventional facilities.

General Site			
Area:	16,000 acres	2 km $\times$ 32 km	64 square km
Orientation:	North-South	Adjacent power R/W	Two circuits @ 230 kV
Population:	2500 persons	Peak occupancy	400 vehicles
Power:	300 MW	3 230-kV substations	@ 100 MW each
Water:	3000 gpm	Base + peak blowdown	@ 300 MW
Cooling:	320 MW	8 cooling towers	@ 40 MW each

Site Components		
Campus:	800,000 sq. ft.	Total square feet
Laboratory Bldg.	350k sq. ft.	One 4-story office-lab
Fabrication Bldgs.	181k sq. ft.	2 hvy., 4 med., w/ cranes
Shop Bldgs.	37k sq. ft.	Three light machine shops
Warehouses	80k sq. ft.	Two @ 40k sq. ft. each
Miscellaneous Bldgs.	152k sq. ft.	Small light buildings
Structures:		
Klystron Gallery	18.4 km $\times$ 9.25 m	1,809,000 sq. ft.
Beam Housing	31 km $\times$ 3.7 m ID	14 ft. TBM bore @ -40 ft. el.
I&C, Utility Alcoves	650, two per sector	400 sq. ft. each
Utility Chases	650, two per sector	6-ft dia. $\times$ 25 ft
rf Penetrations	2318 at 26-ft intervals	2-ft dia. $\times$ 25 ft
Housing & KG Access	34 with labyrinths	Two tiered ramps @ 1 km
Damping Rings	Two e^+ and One e^-	715-ft cir., 50k sq. ft. ea.
Injectors	Two housings	23k sq. ft. each
Position Target	One dual housing	28k sq. ft. for 2 targets
Final Focus	Four @ 2 km each	w/muon spoilers & dumps
Interaction Bldg.	One bldg., pit & crane	100k sq. ft. for 2 detectors
Utilities:		
Electrical	34.5-kV ring bus & ckts.	200 double-end unit subs
Communications	Phone, computer, radio	w/ fibre net, SCADA, video
Site H_2O + Waste	Water-waste, processing	Fire, reclaimed water lagoon
Roads & Parking	54 km roads, 1400 spaces	34 km 2-lane, 20 km 4-lane
Alcove Utilities	650 system clusters	Accel, klys, mag LCW; CA, HVAC
Cable Plant	Wire, tray, racks	337,750 cables; 6,825 racks
Alignment Network	Surface, X-fer, tunnel	Laser trackerbot, barcode index

Table 18-1. *Next Linear Collider zero-order design: Facility criteria.*

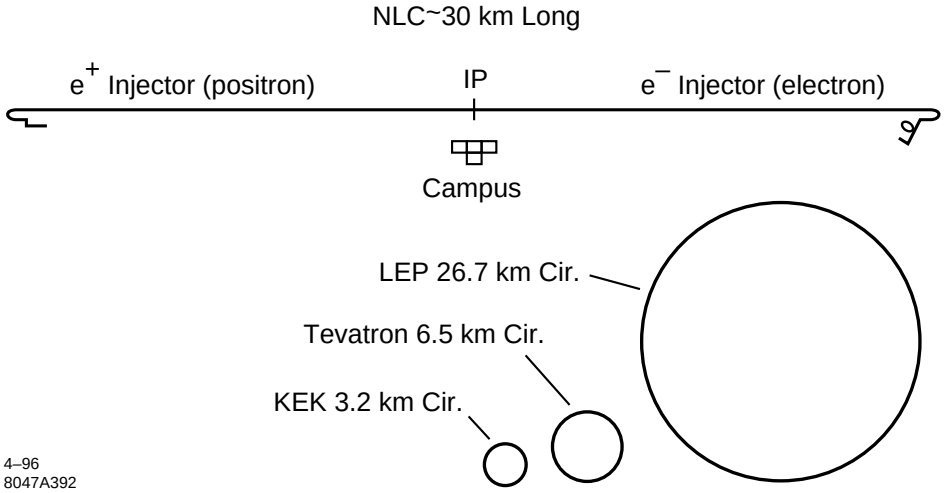


Figure 18-1. *NLC site relative proportions.*

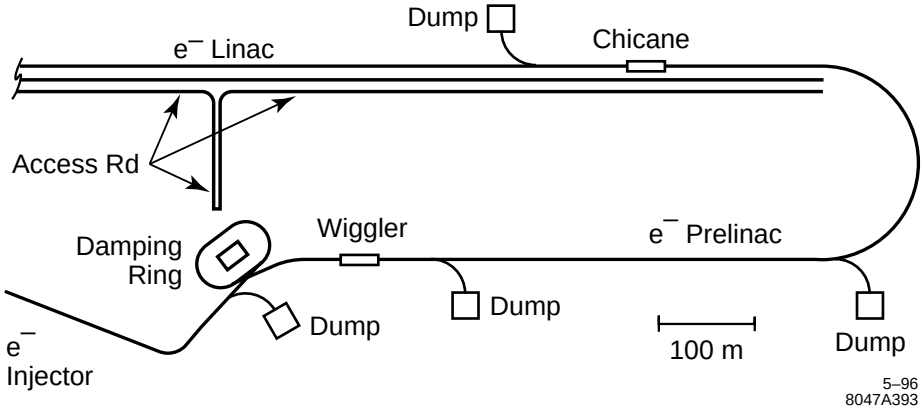


Figure 18-2. *NLC e^- injector and damping ring.*

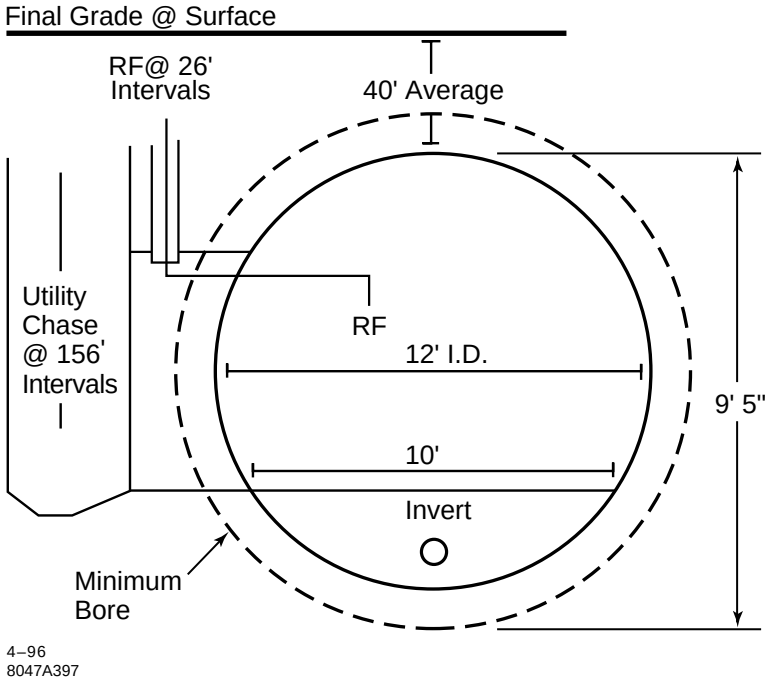


Figure 18-3. *NLC beam tunnel housing.*

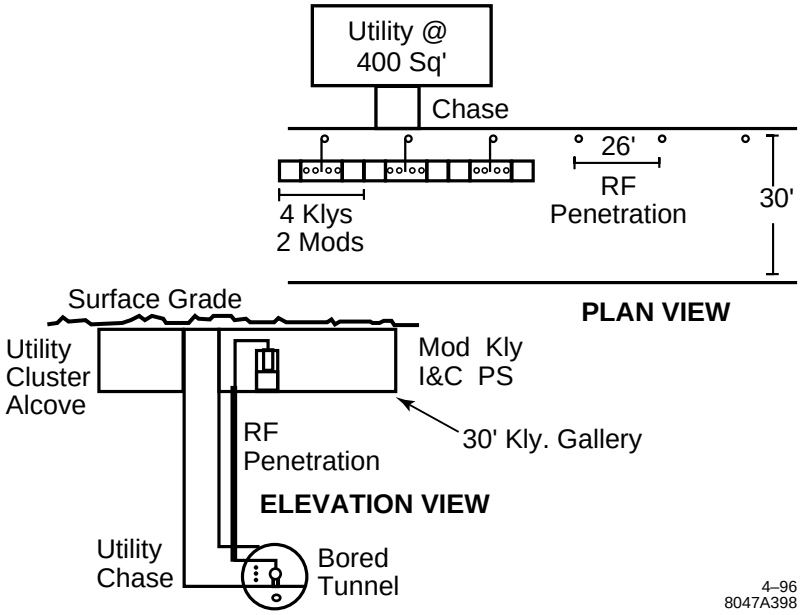


Figure 18-4. *NLC half-sector linac layout.*

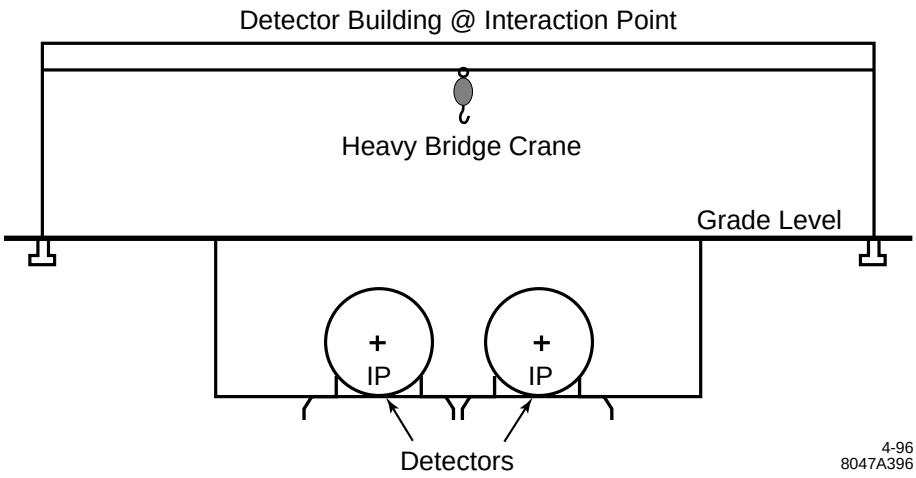


Figure 18-5. *NLC interaction point.*

Contributors

- C. Adolphsen
- C. Corvin
- T. Elioff
- R. Fuller
- D. Hopkins
- P. Kaul
- R. Ruland
- P. Rodriguez
- B. Schmidt
- S. Virostek

A

An RF Power Source Upgrade to the NLC Based on the Relativistic-Klystron Two-Beam-Accelerator Concept

Contents

A.1 Introduction

There are a number of possible routes to upgrade the center-of-mass energy to 1.5 TeV. The most straightforward is to use the same accelerating gradient as the 1-TeV design, 85 MV/m, and increase the length of the X-band linear accelerators by 50%. This would require extending the 10-GeV trombone arm to incorporate the extra length. In addition, magnets in the final focus and collimation regions would have to be rearranged, but the lengths of these regions have already been sized for the 1.5-TeV upgrade.

Unfortunately, the present power sources which would be used to attain the 1-TeV-c.m. energy, *i.e.*, the 50-MW or 75-MW klystrons and SLED-II pulse compression system, are too inefficient for the 1.5-TeV design; they would require an AC power in excess of 350 MW. There are a number of possible future power sources that might be utilized for the 1.5-TeV NLC such as the sheet beam and cluster klystrons, the Binary Pulse Compression system, and the Two-Beam Accelerator. In this appendix, a more detailed description of a Two-Beam Accelerator source, which has been designed at LBNL and LLNL, is described.

As an rf power source candidate for linear colliders, two-beam accelerators (TBA) [Sessler 1982, Sessler 1987] have the inherent advantage of very high efficiency for power conversion from drive beam to rf. In addition, induction-linac-based TBAs have favorable scalings with high frequencies ($\geq$11.4 GHz) and high accelerating gradients ($\geq$100 MV/m). Conversion of high-current electron beam power to rf power has been demonstrated at the gigawatt level at 34 GHz in free-electron laser experiments [Orzechowski 1986], and at several hundred megawatt levels at 11.4 GHz in relativistic klystron experiments [Allen 1989]. Reacceleration experiments [Westenskow 1994] have successfully demonstrated bunched beam transport through two reacceleration induction cells and three traveling-wave extraction cavities for a total rf output of over 200 MW. The phase and amplitude were shown to be stable over a significant portion of the beam pulse.

The technical challenges for making TBAs into realizable power sources lie in the dynamics of the drive-beam which must propagate over long distances. In particular, the beam breakup (BBU) instability through a long, multicavity, relativistic klystron (RK-TBA) is known to be severe. While BBU suppression techniques have been successfully demonstrated for a few cavities [Haimson 1992, Houck 1992a], a scenario with acceptable BBU control over many traveling-wave cavities must be constructed. Similarly, the longitudinal stability of the rf bunches over a multicavity TBA must be demonstrated. For rf phase and amplitude stability, the induction machine must produce and maintain a beam with constant current and energy over the duration of the pulse. In addition to technical feasibility, a case for economic attractiveness is no less essential for the viability of the TBA. Cost and overall system efficiency are essential elements of an acceptable linear-collider power source.

With these general considerations in mind, a systems study, including physics and engineering designs, as well as bottom-up costing, was conducted for a point design using the RK-TBA concept as a power source for the upgraded Next Linear Collider (NLC) [Siemann 1993]. We refer to this design as the TBNLC. In the following we present the results of the TBNLC system study and our current estimate of total system efficiency of 50%. We also describe an experimental program to reduce the risk of implementing the RK-TBA concept to a large collider. Demonstrations of the key concepts of the TBNLC design are expected to be completed by 2002 at which time the TBNLC could be seriously considered for a 1.5-TeV NLC upgrade. A schematic of a portion of the low-energy e^- side of the TBNLC is shown in Figure A-1.

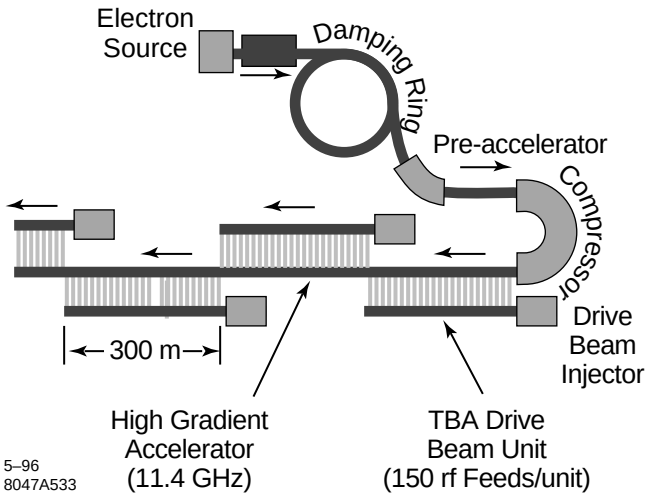


Figure A-1. *Low-energy portion of the e^- side of a linear e^+e^- collider using an RK-TBA design for the rf power source. Our design requires 76 TBA units to provide 1.5 TeV of energy gain.*

A.2 A Design for an RK-TBA-Based rf Power Source

A.2.1 Power Source Requirements

The objective of the TBNLC system study was to construct a conceptual design of the power source for the NLC. Our philosophy was to stay within the design constraints of the NLC and to have a power source system that matches the high-gradient structures being considered for the NLC. However, since the parameters of the NLC were not finalized at the time of the study, we made some choices for power source requirements so that the design and costing studies could be concrete. The design goals for the study were as follows:

Rf frequency	11.424 GHz
Repetition rate	120 Hz
Peak power/structure	360 MW
Distance between extraction structures	2 m
Pulse length (flattop)	200 ns
Pulse rise time	125 ns
Center-of-mass energy	1 TeV

The peak power of 360 MW at 11.424 GHz corresponds to an unloaded gradient of 100 MV/m in the NLCTA high-gradient structures presently tested at SLAC. This power is generated from an extraction cavity in the RK every 2 m, to match the 1.8-m accelerator sections, and to provide adequate spacing in between the HGSs for input and output couplers, focusing, etc.. In the study, we costed the power source for a total of 15 km of this high-gradient structure, 7.5 km for each arm of the collider. There is a total of 7500 rf extraction cavities, with a total peak power of 2.7 TW. The unloaded energy gain from each arm of the collider is 675 GeV. Assuming a loaded gradient that is 75% of the unloaded, final energies of the electron and positron bunches are 506 GeV each.

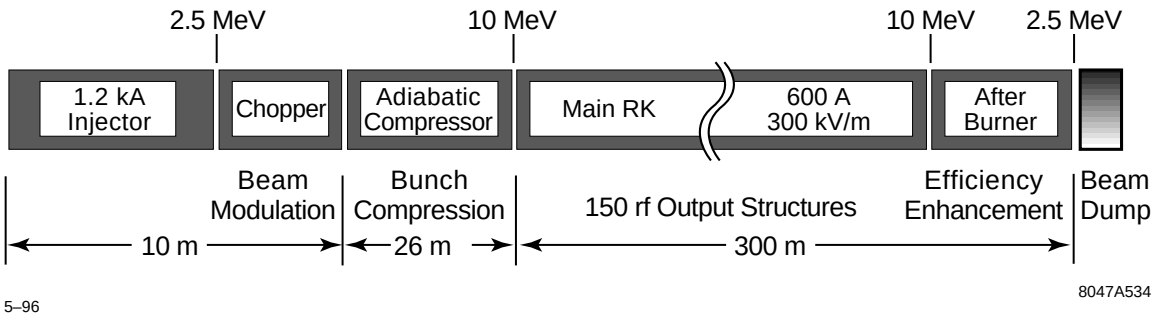


Figure A-2. *Schematic showing major components of the RK-TBA unit.*

The pulse length was determined by the fill time of the structure plus the length of the multi-bunch particle train. To maintain constant particle energy from bunch-to-bunch, the required electric field should grow linearly in the first 100 ns, and remain constant over the remainder of the beam-on time. The flattop is specified as 200 ns, approximately twice the flattop pulse length of the present parameter list (∼100 ns flattop), which allows us to accelerate pulse trains of 143 bunches and leads to a corresponding increase in luminosity. This is possible for our TBA design at only modest cost increase, and is a natural extension of the NLC parameter set to take advantage of the RK-TBA structure. Because of the intrinsically high-efficiency of the RK-TBA, the average power required stays at about 200 MW.

The required electric field at the front of the 100-ns rise is roughly 25% of the asymptotic value. Hence, we specify our input pulse to rise linearly (in field) from zero to its full value in 125 ns. The power increases quadratically over the rise time. This pulse shape can be generated in an RK by a corresponding linear rise in the current waveform of the drive beam, and a linear rise in the reacceleration voltage as well. Further refinement of the pulse shape to match dispersion in the high-gradient structure and to improve energy flatness can be achieved by appropriate shaping of the drive beam current and voltage, but were not considered in the TBNLC study.

A.2.2 TBNLC Architecture

To provide rf power for a 1.5-TeV high-gradient linear accelerator, the TBNLC design requires 76 independent RK-TBA units. Each RK-TBA unit is about 300-m long, and has 150 extraction cavities (Figure A-2). To replenish the 360 MW generated from the extraction cavity every 2 m, the 10-MeV drive beam with average current of 600 A is reaccelerated at 300 kV/m. The design current could be raised slightly to compensate for rf losses in the induction cells (*e.g.,* 625 A for 4% loss). The main body of the RK consists of identical 2-m modules each of which has six 100 kV induction cells and one extraction cavity. Both the drive beam current and reacceleration voltage have a rise time of 125 ns and a 200-ns flattop, with a falltime that is comparable to the rise time.

The front end of each RK consists of an injector in which a 1.5-kA beam of electrons is generated and accelerated to 2.5 MeV. A 5.7-GHz chopper is placed at this point to generate a bunched beam at 11.4 GHz. See Section A.5.3 and Figure A-20 for additional information on the modulation of the beam. Chopping reduces the DC current from 1.5 kA to 600 A. This bunched beam at 2.5 MeV is then accelerated to 10 MeV in an adiabatic compressor section. Bunching cavities in this section further reduce the length of the bunches, and prepare the beam with the right phase space for injection into the main RK. At the end of the RK, there is an afterburner section in which rf extraction continues, but reacceleration is absent. The afterburner allows us to further extract rf power from the bunched beam, and thus enhances the overall efficiency of the RK. At the end of the afterburner, the spent beam (at 2 to 3 MeV) is collected

at a beam dump. The adiabatic compressor section is 26-m long, while the main RK has 138 extraction cavities over 276 m. The afterburner consists of 12 rf cavities, making a total of 150 extraction cavities.

The overall length of a RK-TBA unit is determined by a balance of two opposing considerations. The longer the RK-TBA unit, the higher the overall efficiency, as the overhead losses from the front and the back ends become a smaller fraction of the total power. On the other hand, the control of beam instabilities and beam degradation effects become increasingly more difficult as the overall length is increased. Longitudinal and transverse beam dynamics simulations indicate that the stability of the bunched beam can be maintained in the proposed RK-TBA unit.

The efficiency for conversion of power in the drive beam to rf power is easily evaluated for the proposed RK system. Allowing for some rf loss in the induction cavities, the overall efficiency is estimated to be 90%. The 10% loss is shared among the beam loss on the chopper (3.7%), beam dump (2.8%), and rf into induction cavities (3.6%). The efficiency for wall plug to drive beam is 55%. Hence the total efficiency from wall plug to rf is 50%.

The rf power requirement of 360 MW/2 m determines the product of the accelerating gradient and beam current in the drive beam. Our particular choice of 300 kV/m and 600 A is again based on a balance of two considerations. With higher current and lower gradient, the volume of magnetic material required is reduced and the efficiency is increased. However, beam transport becomes more difficult with increasing current.

The linear rise in the drive beam current assures that the extracted rf field has the right waveform for beam loading in the HGS, while the linear rise in the voltage of the reacceleration cells assures that the entire drive beam stays at 10 MeV from head-to-tail. This is important for both beam transport and for rf phase stability. Hence, the rising portion of the current and voltage generate useful rf power, but the fall portion is wasted. The corresponding loss in efficiency due to finite falltime is included when we consider the efficiency of the pulse power system.

A.2.3 Main RK

The main RK consists of identical 2-m modules. The key elements of a module are the permanent quadrupole magnets for focusing of the drive beam, the induction cells for reacceleration, and one rf extraction cavity. Design considerations for each of these major components is discussed below. The key issues in the design are related to the physics of rf extraction, reacceleration, and transport. However, attention was paid to making the module compact and efficient. The resulting unit has a diameter of 35 cm, which is quite small compared to existing induction accelerators. A schematic of the 2-m module is shown in Figure A-3.

Pipe size

In the design of induction accelerators, the size of the beam tube is usually determined by BBU considerations, since the transverse impedance from the acceleration gap is inversely proportional to the square of the pipe radius. In our RK design, the low-frequency BBU arising from the induction gaps is ameliorated by the low-beam current of 600 A and Landau damping associated with the energy spread inherent in the rf buckets. With proper attention to the gap geometry, we were able to design a pipe with a radius of 2.5 cm, and predict acceptable BBU growth.

Permanent magnet focusing

The small pipe radius allows us to take full advantage of permanent magnets, which have significant cost advantages when the focusing systems are small. Another major reason for using permanent magnets is the associated efficiency, as they eliminate the need for power supplies.

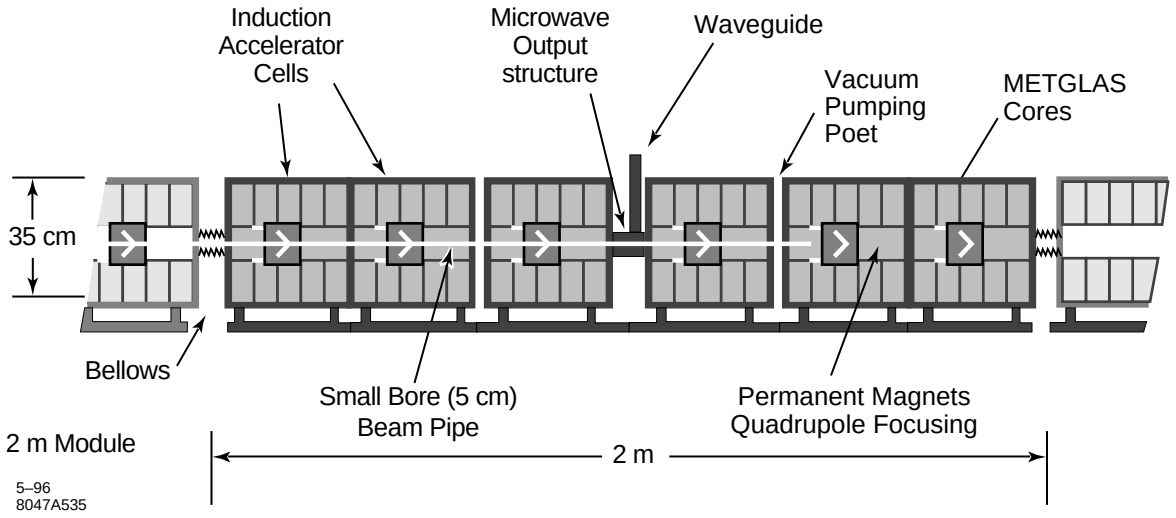


Figure A-3. *Schematic of 2-m RK Module*

Several basic considerations govern the design of the FODO lattice. The foremost requirement is associated with a technique concocted to minimize the high-frequency BBU growth arising from the HEM_{11} mode in the rf extraction cavities. To suppress the instability growth, we demand that the betatron period be equal to the distance between adjacent extraction cavities,*i.e.,* 2 m. In addition, the phase advance per lattice period must be less than 90° to ensure beam stability. Finally, the focusing strength must be sufficient to keep the beam envelope small enough to stay within the beam tube.

The zeroth-order design equation is given by the thin lens approximation

$$\sigma_0 = \frac{\eta B L^2}{R[B\rho]} \tag{A.1}$$

where σ^0 is the phase advance per lattice period, and B is the quadrupole field strength at pole-tip with radial position R. L is the half-lattice period, and η the occupancy factor for the quadrupole. The rigidity of the electron beam is given by

$$[B\rho] = \frac{\beta\gamma mc}{e} \tag{A.2}$$

where $\beta\gamma mc$ is the momentum of the relativistic beam, and $[B\rho]$ has the value of 0.035 T-m for 10-MeV electrons.

A first-order correction for deviations from the thin-lens approximation is obtained by replacing the phase advance and occupancy factor with scaled variables given by [Barnard 1991]

$$\hat{\sigma}_0 = \sqrt{2(1-\cos\sigma_0)} \quad \text{and} \quad \hat{\eta} = \eta\sqrt{2\left(1-\frac{2\eta}{3}\right)} \tag{A.3}$$

In our design, we use σ_0 of 60°. There are 12 half-lattice periods in 2 m. The physical occupancy factor is 0.48. The modified thin lens formula then gives a B-field at pole tip of 800 Gauss. This estimate is within a few % of the actual B-field required to give a 2-m betatron period, when we include realistic spatial profile of the magnetic field, as well as the energy variations of the drive beam as it is accelerated over the 2-m module.

A preliminary quadrupole design was constructed by Klaus Halbach [Halbach priv, Halbach 1980]. It consists of four rectangular blocks of ferrites with residual field B_r of 3.2 kg. The dimensions of the blocks are 1.1 cm $\times$ 3.22 cm $\times$ 8 cm. This design has as the first nonzero harmonic $n = 10$, with 8.6% strength of the quadrupole field at aperture radius. Finite differential permeability effects have not been included in this preliminary design.

A basic requirement of the transport channel is that the focusing strength must be sufficiently strong to keep the beam from hitting the beam tube. The beam envelope in the focusing channel can be estimated from the equilibrium formula

$$\frac{\varepsilon^2}{a^2} = \sigma_0^2 \left(\frac{a}{L}\right)^2 \tag{A.4}$$

where $\varepsilon = \varepsilon_n/\beta\gamma$ and a are respectively the unnormalized edge emittance and the edge radius of the beam envelope. With a normalized edge emittance of 600 mm-mr, (which is achievable with a good gun design), we obtain an edge radius of about 2 mm for the beam.

Steering and Focusing Corrections

Steering and focusing corrections are achieved in the design by means of low-field ($\sim$10 Gauss) DC coils located in the region immediately outside of the permanent magnets, at radial positions of about 4 cm from axis. These correction coils can be manufactured inexpensively on printed circuit boards.

Since the beam apertures are small, particularly around the extraction cavities, beam centroid displacements must be kept to a minimum. Quadrupole misalignments, when combined with energy variations from head to tail, can lead to beam offsets that change over the length of the pulse, *i.e.*, the corkscrew phenomenon [Chen 1990]. Our strategy for minimizing beam displacements and associated corkscrew is to impose misalignment tolerances that are not excessively tight, design for energy flatness of $\leq$1%, and to correct alignment errors with closely spaced dipole-steering coils. Detailed steering algorithms have not been worked out yet, but for the purpose of costing, we provide three sets of steering coils and two beam position monitors per betatron period.

A key ingredient for the suppression of high-frequency BBU is that the betatron period be equal to the distance between adjacent cavities. To ensure that this requirement is obeyed, a feedback system with two correction quadrupoles per betatron period is incorporated into the design. A possible monitor for the betatron motion is to use the rf output from the extraction cavities at the dipole frequency. Whether the signal sensitivity is adequate for this purpose needs further study.

Induction Cores

The design of induction cells for reacceleration is based on the relation

$$V \cdot (\Delta t) = A \cdot F_p \cdot (\Delta B) \quad , \tag{A.5}$$

where the core size (with magnetic flux over area A and F_p is the packing factor, or the fraction of the total core that is occupied by magnetic material) is determined by the required voltage V and pulse length Δt. The allowed flux swing ΔB is different for different magnetic material. Our design uses METGLAS®[1] with ΔB of $\sim$1.3 T. Flux swings for METGLAS are several times larger than ferrites, for example, which are commonly used in short-pulse induction machines such as the Advanced Test Accelerator (ATA) [Kapetanakos 1985] and the Experimental Test Accelerator II (ETA II) [Clark 1988]. The cores are correspondingly more compact. Furthermore, METGLAS is quite inexpensive, particularly when bought in large quantities.

[1] METGLAS® is a registered trade name of Allied-Signal.

To determine the transverse dimensions of the core, (with outer radius R_o, and inner radius R_i), we recast Eq. A.5 in terms of the accelerating gradient G and core occupancy factor η_c (fraction of the axial length occupied by induction cores) as follows:

$$R_o - R_i = G\Delta t/(\Delta B \eta_c F_p) \tag{A.6}$$

where F_p has a value of 0.65 to 0.75 for typical METGLAS cores. Since the accelerating gradient and pulse length are fixed by overall system requirements, our design philosophy was to maximize axial core occupancy. The proposed design has η_c of 0.75 and core thickness $R_o - R_i$ of 10 cm.

Compact cores reduce cost as well as energy loss. Empirical measurements of core losses can be parameterized to give the following phenomenological formula for the core current (in Amps):

$$I_c = 360(R_o + R_i) \cdot (\Delta B/\Delta t) \tag{A.7}$$

where the radial dimensions are in meters, ΔB in tesla, and Δt in microseconds. The fraction of energy lost in the core ($I_c/I_{\rm total}$) is proportional to the size of the core. Our design parameters give a core current of about 117 A. With a 600-A beam, the core efficiency is 80%. See Section A.4.2, "Induction Accelerator Module", for results of core testing.

The magnetic material in a 2-m module is packaged into six independent 100-kV induction cells. Each cell in turn consists of five 20-kV cores. Packaging into small induction cores provides a natural match to the low-voltage pulse power system. Figure A-4 is an illustration of the proposed induction cell showing the five cores.

Pulsed Power

The desired voltage pulse from each core has a 125-ns linear rise, followed by a 200-ns flattop at 20 kV, followed by a fast fall. This voltage waveform is generated with pulse forming networks (PFNs) with tapered impedance to match the induction core. Power input to the PFNs consists of a DC power supply and a Command Resonant Charging unit. The entire pulse power train is at low voltage, and no step-up transformers are needed. Hence, high efficiency and low cost is possible.

Acceleration Gap

Each 100-kV induction cell has an induction gap, which is one of the more critical components of the design. Ongoing design efforts evolve around the following key issues:

- **Reduction of high voltage breakdown risks.** To have an induction gap that is safe from breakdown, the gradient across the insulator must be sufficiently low (<30 kV/cm), sufficiently low fields on the metal surfaces (<100 kV/cm), proper shielding of the insulators from secondary electrons and X-rays generated by the beam, and proper design of the triple-point (the interface between insulator, metal, and vacuum).
- **Suppression of low-frequency BBU.** Dipole modes associated with the gap (at a few GHz) can lead to severe problems if not carefully damped. The required transverse impedance was achieved with heavy de$Q$ing by placing microwave-absorbing material at critical locations around the gap.
- **Reduction of transverse and longitudinal impedances at high frequencies.** While the high-frequency BBU (HEM_{11} mode at ~14 GHz) is generated primarily in the rf extraction cavities, and the major BBU suppression activities center around them, one must be careful to ensure that the transverse impedance contribution from the induction cavities are indeed negligible. In the present designs, the transverse impedance around 14 GHz can be

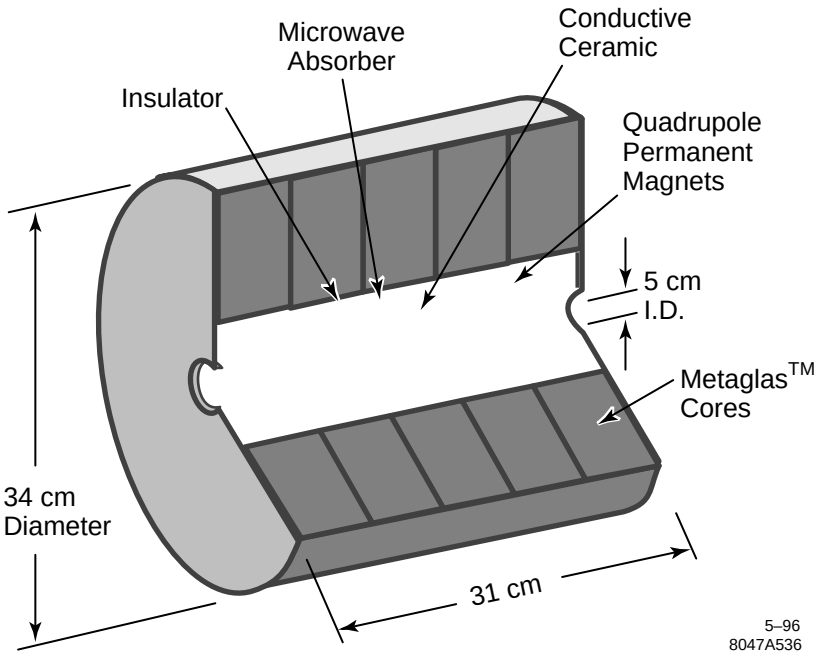


Figure A-4. *Proposed RK induction cell design.*

made to be lower than half an ohm, and its effect on the high frequency BBU is insignificant. The longitudinal impedance at 11.424 GHz must be low to minimize the microwave power loss. Our design goal is to maintain the induction gap loss to less than 4%, which requires that the longitudinal impedance be less than 2.4 Ω. Present designs are approaching the required impedance.

RF Extraction Cavities

Design of the rf extraction cavity is a major ongoing activity in theory and simulations. Present designs evolve around traveling-wave structures with three cells of eight-mm inner radius. The rf output is extracted through two separate ports in the third cell, with 180 MW each transported through separate waveguides, and fed directly into the two input couplers of the high-gradient structure. An illustration of a proposed design is shown in Figure A-5. Detailed design of the cavities centers around five key issues:

- **Required output power.** The extraction cavity must have the right effective impedance of about 540 Ω to extract 360 MW from a highly bunched beam (70° microbunch length) whose first harmonic current at 11.424 GHz is about 1.15 kA.
- **Minimal breakdown risks.** Our present designs use traveling-wave structures with three cells to reduce the surface fields. Surface-field gradients of less than 70 MV/m appear possible.
- **Inductive detuning.** In addition to generating the right amount of power, the rf cavities provide continuous longitudinal bunching for the drive beam. This is accomplished by inductively-detuned traveling-wave-structures.
- **Dipole de$Q$ing.** The rf cavities must have low-transverse impedances for the suppression of BBU. Fortunately, there is a natural de$Q$ing mechanism, as the required extraction cavities have inner radii of around 8 mm, and the

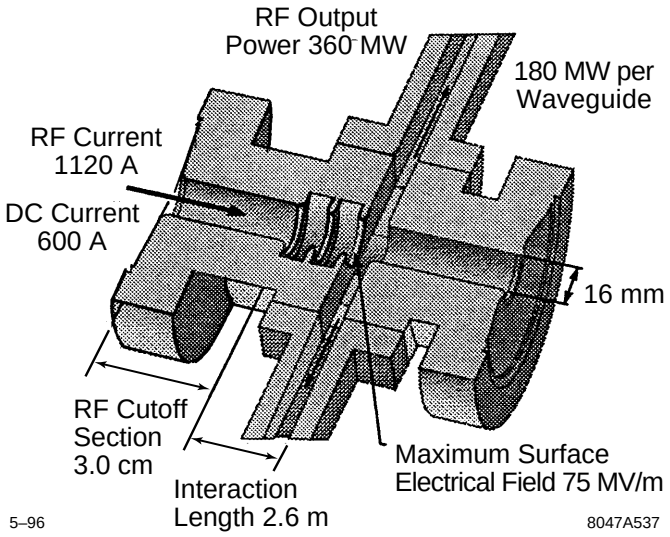


Figure A-5. *Illustration of an rf extraction structure.*

dipole modes in the cavity couple to the TE_{11} mode in the pipe. Simulations to date suggest that the resulting impedances are sufficiently low for BBU suppression provided that the Betatron Node Scheme is used.

- **Transverse focusing.** The rf cavities have transverse fields associated with the fundamental mode. They can degrade the beam envelope if they are sufficiently strong and/or not properly corrected. Evaluation of these transverse forces from theory and simulations is ongoing.

A.2.4 Front End and Back End

Each RK unit has a front end that consists of an injector, a chopper, and an adiabatic compressor section. At the end of each RK unit, there is an afterburner, followed by a final beam-dump. These five elements are described below.

Pre-chopper

The injector is basically a short-induction accelerator to generate an unbunched 300-ns electron beam at 2.5 MeV. This section may consist of an induction injector (at 1–1.5 MeV), followed by a short-induction accelerator section that takes the beam to 2.5 MeV. The required current from the injector is 1–1.5 kA, depending on the efficiency of the chopper which is required to produce at its exit a bunched beam with a DC current of 600 A.

The current waveform required consists of a linear rise over 100 ns, followed by 200 ns of flattop. The beam energy should be constant over the entire 300 ns, including the front portion with the rising current. To generate the linearly rising current, the voltage pulse in the 1-MeV injector will also, by the Child-Langmuir Law, have a (nonlinear) 100-ns rise time. The PFN in the subsequent short accelerator section must be arranged to produce a higher voltage at the beam front, thereby compensating for the low-energy at the injector exit.

Since the electron beam will be transported over 300 m of narrow pipes, caution must be taken to minimize head-to-tail energy variations, as well as transverse beam offsets and beam temperature. The injector design should produce a

bright beam with a normalized edge emittance of 600 mm-mr or less, and transverse displacement of 200 microns or less.

Chopper

The purpose of the chopper is to generate cleanly-separated microbunches at 11.4 GHz. The basic scheme follows the design of Haimson Research Corporation's Choppertron [Haimson 1989], a 11.4-GHz microwave generator that has been deployed at LLNL. A subharmonic dipole deflecting cavity at 5.7 GHz causes the electron beam to oscillate about a limiting aperture, leading to chopped bunches at twice the oscillating frequency. A similar device can serve as the TBNLC front-end chopper, except that much more effective heat dissipation is required for the 120-Hz operation of the NLC upgrade.

Since a substantial amount of energy is lost on the chopper, there is incentive to make it more efficient. One idea is to precede the subharmonic deflecting cavity with conventional bunching cavities at 11.4 GHz. The role of the chopper then is primarily to clean up the particles with the wrong phase. This requires straightforward phasing of the 5.7-GHz input power to the deflecting cavity relative to the incoming prebunched beam.

Adiabatic Compressor

The exiting beam from the chopper has microbunch lengths equivalent to 180° or greater in longitudinal phase space. In the adiabatic compressor region, the microbunch lengths are further reduced to 70°, which is the needed bunch length for long-distance propagation in the main RK. This is accomplished by a number of idler cavities that are more inductive than the rf extraction cavities in the main RK.

In addition to microbunch sharpening, this section also serves to provide the energy transition from 2.5 MeV to 10 MeV. The induction core structures are very similar to the main RK. However, the voltage waveform should be flat over the 300 ns of beam-on time, so that the head-to-tail energy flatness required for phase stability could be maintained at 10 MeV.

The quadrupole magnets are weaker at the lower energies, and continues to increase with increasing energy, so that the betatron wavelength is kept fixed at 2 m. Structurally, the adiabatic compressor section looks very similar to the main RK, except that no power is extracted from the idler rf cavities.

Afterburner

At the end of the RK is an afterburner section, the primary purpose of which is to increase overall system efficiency by extracting more power out of the bunched beam at the end of the main RK. This section has a number of rf extraction cavities, permanent magnets for focusing, but no reacceleration cells. The spacing of the rf cavities is changed as the average energy of the beam continues to decrease. The impedance of the cavities is also modified to compensate for changes in the rf bucket. The TBNLC design has 12 cavities in the afterburner section, each generating 360 MW, and together covering a total of 12 m.

Beam dump

As the spent electron beam reaches the final beam dump, its average energy is less than 3 MeV. The design of this component is straightforward.

A.2.5 The Key Ideas

A few key ideas underlie the whole design of the relativistic klystron for the TBNLC and are summarized as follows:

- **Betatron Node Scheme for high frequency BBU control.** The most severe BBU instability is associated with the HEM_{11} transverse mode in the rf extraction cavities. Strong suppression of this mode is achieved by introducing a Betatron Node Scheme [Li 1994] in which adjacent rf extraction cavities are placed exactly one betatron period apart. This scheme minimizes beam centroid displacement which excites the transverse mode, and alters the nature of the instability from exponential to a slow secular growth.
- **Landau damping for low-frequency BBU suppression.** Quite apart from the high-frequency dipole mode associated with the rf cavities, there is a low-frequency (a few GHz) dipole mode associated with the induction reacceleration gaps. This BBU instability is ameliorated by Landau damping due to a large energy spread inherent in the rf buckets of the bunched drive beam. Combining this feature with dipole mode suppression measures in the induction gap design, the calculated low-frequency BBU growth is minimal.
- **Inductively detuned cavities for longitudinal beam stability.** To maintain tight rf bunches over long distances with multiple extraction cavities, the rf output structures are inductively detuned. While the concept of inductive detuning is not new, the theoretical framework has to be developed and implemented in simulation codes for traveling wave output structures. Cavity parameters required for stable beam propagation through multiple structures were determined in the TBNLC system study.
- **Small low-cost induction cells.** The induction cells in the TBNLC design are about one foot in diameter, much smaller than most of the existing induction machines. This design was possible because of three key features:
  - Use of METGLAS for the induction cores. This is a low-cost magnetic material that can accommodate a large flux swing.
  - Use of low-field permanent magnet quadrupoles. The relatively high drive beam energy of 10 MeV (compared to klystron beams) is a natural match to a strong focusing transport system with low-cost ferrites. An additional advantage of permanent magnets is, of course, the elimination of power supplies.
  - Narrow beam pipes—the design pipe diameter of 5 cm is much smaller than other existing high current machines. This is possible because of the relatively low current (by induction machine standards) of 600 A, and the low-frequency BBU suppression features described earlier. Large beam pipes in usual induction machines are dictated by BBU considerations.
- **Low-voltage pulse power system for efficiency and cost.** The induction cell design matches naturally to a 20-kV system with PFNs triggered by ceramic thyratrons and powered by DC power supplies and command resonant charging systems. This system does not require step-up transformers and eliminates losses associated with a high-voltage system.

A.3 TBNLC Physics Studies

The TBNLC system study addressed the issues of longitudinal and transverse stability of the drive beam and the related areas of rf extraction cavity and induction cavity designs.

A.3.1 Longitudinal Beam Dynamics

Longitudinal stability of the drive beam is among the important issues for the demonstration of technical feasibility of the RK-TBA concept. In the TBNLC design, the drive beam is required to stay bunched longitudinally over 150 extraction cavities. Space charge effects cause initially tight bunches to expand. The debunching process is further aggravated by the energy spreads along the bunches as they interact with rf fields. In usual traveling wave extraction cavities, rf waves are in synchronism with the drive beam and debunching becomes very severe after a few cavities. Beam debunching, if uncompensated, will result in reduced power extraction in subsequent cavities. To overcome this problem, we employ a scheme in which the drive beam is not synchronous with the operating wave mode; more specifically, the phase velocity of the rf field is larger than the speed of light so that the bunches always lag behind the wave. This is what we call inductive detuning and can effectively bunch the beam. The concept of inductive detuning is well known for standing wave structures (SWSs), *e.g.,* the penultimate cavity in a klystron, in which the frequency of the cavity is detuned from the resonant frequency. For the traveling wave structures (TWSs) of our RK, the approach we take is to keep the frequency ω of the operating mode unchanged, but reduce its wave number k, such that the wave field advances faster than the beam. In this case particles at the front of the bunch lose more energy and slow, while particles at the tail of the bunch lose less energy and catch up. This mechanism causes a continuous sharpening of the bunch, thus counteracting the debunching forces. The resulting longitudinal phase space continues to rotate in a stable rf bucket with a relatively stable bunch length. Thus both constant power extraction and beam bunching can be achieved simultaneously.

Parameters for the detuned rf output structures were estimated from analytical expressions developed during the TBNLC system study. Numerical simulations were then performed using the relativistic klystron code RKS [Ryne 1990]. This code solves self-consistently the single-particle equations of motion for the beam and the coupled circuit equations that govern the cavity excitation, and it includes the calculation of the space charge effect. It assumes a single-dominant mode and cylindrical symmetry of its fields inside the cavity.

Figure A-6 presents the output power from each of 150 TWSs in the main RK section for both a successful inductive detuning case and its corresponding synchronism case. The parameters for the inductive detuning case are given in Table A-1. For the synchronism case, the level of the extracted power $P_{\rm out}$ declines sharply due to the space charge debunching effect as the drive beam traverses the RK. In contrast, when the TWSs are properly detuned, the rf bucket can remain stable and output power can be sustained at the desired level ($\sim$360 MW) for the 150 extraction cavities.

The synchronism case in Figure A-6 consists of conventional three-cell TWSs operating at the $2\pi/3$ mode. In the inductively detuned case the operating detuning angle is 30°. The cavity is therefore operating in a $\pi/2$ mode (*i.e.,* phase advance of the field across one cell is 90°). The longitudinal dimension of each cell is the same in the two cases while the transverse dimensions are varied. URMEL and MAFIA codes were used for detailed cavity design.

A key feature for RK design is the cavity filling time, *i.e.,* the time it takes for the rf field in a cavity to reach equilibrium state. In Figure A-7, we present the time dependencies of output power at the 50th, 100th and 150th extraction cavities for the inductive detuning case in Figure A-6. It shows that after about 15 ns, the fields in all the cavities reach their equilibrium states. This indicates that the erosion on the beam head due to the cavity filling process is not serious. The short fill time is a result of low Q and high V_g.

Further RKS code simulation studies were conducted to examine the sensitivity of the inductive detuning scheme to the important parameters of bunch length and detuning angle. For 30° of detuning, the scheme was insensitive to variation in bunch lengths from 50° to 90°, although performance deteriorated for lengths above 90°. Similarly, the scheme was insensitive to increasing the detuning angle from 30° to 35°, but performance deteriorated for angles below 30°.

The issue of radial defocusing, or emittance growth, due to the interaction of the beam with the rf fields in the extraction structures is an active area of study. The underlying theory of the RKS code involves power balance. Detailed rf

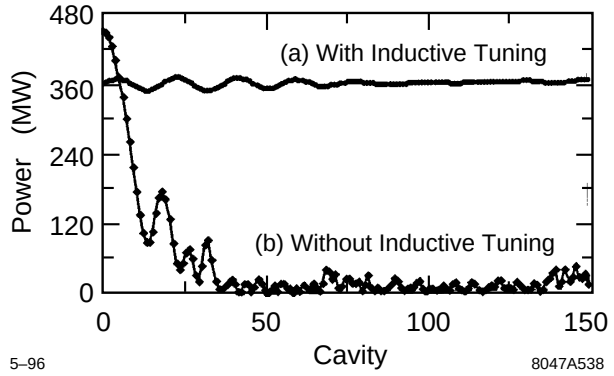


Figure A-6. *Power extraction from 150 cavities in one unit of RK-TBA: (a) With inductive detuning ($\Delta = 30^{\circ}$, phase velocity 1.33 c); (b) No inductive detuning ($\Delta = 0^{\circ}$, phase velocity 1.0 c).*

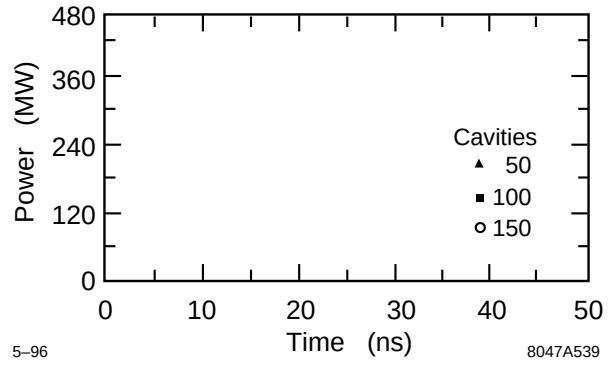


Figure A-7. *Time evolutions of output power at 50th, 100th and 150th TWSs for the nonsynchronism case in Figure A-6.*

field information is less stringently modeled. This approach works satisfactorily for standing wave and synchronous traveling wave structures. However, for detuned traveling wave structures, the RKS code does not adequately account for radial forces on the beam limiting its usefulness for extended structures to 1-D simulations.

A.3.2 Transverse Beam Dynamics

The excitation of transverse beam instabilities due to higher order modes has been identified as a major issue in the design of a long multicavity RK-TBA. The narrow aperture and high average current of the RK accentuates the problem. There are two separate structural components that contribute to the transverse instability. The induction module with an aperture of 5.0 cm has a trapped dipole mode around 3 GHz and the output structure with a 1.6-cm aperture has a transverse mode near 14 GHz. Both modes can interact strongly with the beam. The effect of these structures on transverse instability can be studied separately due to the difference in resonant frequencies.

Drive frequency	11.424 GHz
Forward traveling mode	TM_{01}
Number of cavities	3
Phase shift per cavity	90°
Wave length	2.626 cm
Phase velocity	1.33 c
Group velocity	0.28 c
Shunt impedance per cell $(R/Q)^*$	27.0 (Ω)
Eigenfrequency for the first 2 cells	11.424 GHz
Eigenfrequency for the 3rd cell	11.666 GHz
Wall-dissipation quality factor	7000
External quality factor for the 3rd cell	6.5
Aperture inner radius	8 mm
Aperture outer radius	12.5 mm
Iris thickness	2.5 mm
Longitudinal dimension of each cell	8.754 mm
Beam energy	10 MeV
Beam current (peak)	600 A
Bunch length	0.51 cm
Beam radius (rms)	2.5 mm

*Traveling-wave shunt reactance is twice the standing-wave shunt reactance.

Table A-1. *Parameters related to the inductive detuning case.*

Low-Frequency BBU

There are three induction modules per meter through the RK. Although considerable damping can be accomplished by the insertion of absorbing material, design constraints imposed by maximum surface electrical fields and vacuum requirements preclude complete damping of higher order modes. Modeling the module's gap as a cylindrical resonator with an impedance boundary condition on the outer wall [Brigg 1985] and assuming an optimum design, the transverse impedance, $Z_{\perp 1}$, will be on the order of 4,000 Ω/m.

Analytical theory [Panofsky 1968] indicates that the transverse instability will grow exponentially along the length of the accelerator. Experience with ATA and ETA II indicate a growth of 4 to 5 e-folds from noise is tolerable in the transverse instability. For a 10-MeV e-beam, using a betatron wavelength of 2 m, a current of 600 A, and gap spacing of m/3, a theoretical estimate of the maximum tolerable transverse impedance of the module, $Z_{\perp 1}$, is about 573 Ω/m. Thus, additional measures are needed to suppress the transverse instability.

The most promising technique for suppressing the transverse instability is Landau damping. To maintain longitudinal equilibrium, inductively detuned rf extraction structures are used. The resulting rf buckets have an intrinsic spread in beam energy over the microbunch on the order of $\pm 7.5\%$. Hose instability theory [Lee 1978, Houck 1993] can be used to estimate a maximum $Z_{\perp 1}$ of about 5,178 Ω/m for total suppression of the low-frequency BBU growth. Theoretically, we can build a suitable induction module to avoid BBU.

The OMICE code [Houck 1992b] was used to model the growth of the transverse instability. In Figure A-8, relative growth in the displacement of the beam's centroid from axis as a function of distance along the axis is shown for several different energy spreads. The finite number of beam slices used in the OMICE code did not permit the level of phase mixing available in a physical beam. Thus, the simulations presented provide a conservative upper bound.

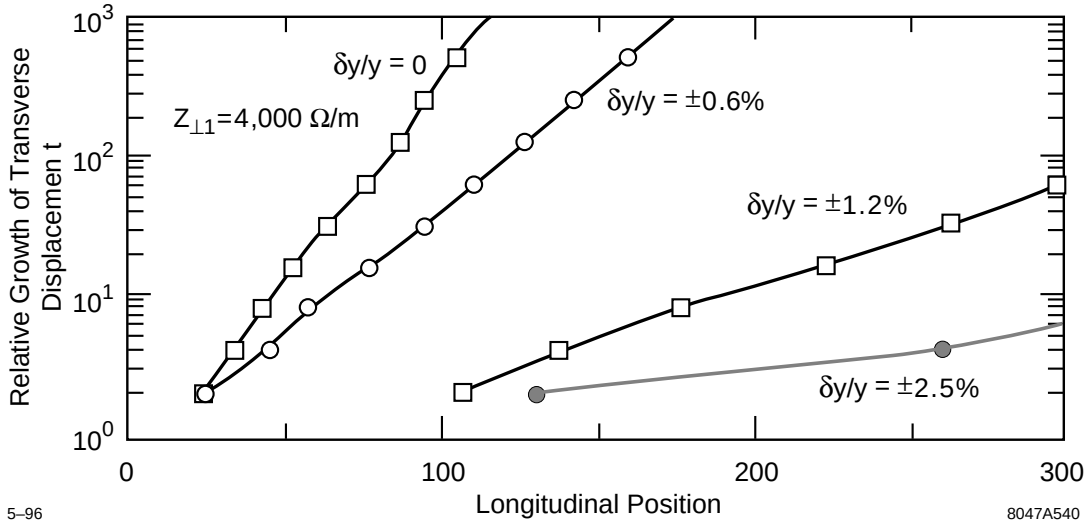


Figure A-8. *Effect of energy spread on the low-frequency instability.*

Table A-2 lists additional input parameters used in the OMICE code. The seed perturbation for exciting the instability was a sinusoidal oscillation of the centroid's transverse displacement at the instability frequency. The conclusion from the numerical modeling is that the instability growth will be acceptable for $Z_{\perp 1}$ less than or equal to 5,400 Ω/m and energy spreads of $\pm 2.5\%$ or greater.

High-Frequency BBU

There is one three-cell, traveling-wave, output structure (TWS) every 2 m through the RK. In the context of the TBNLC design, we limit the growth of the transverse beam displacement to 4 to 5 e-folds over the 300-m system, or 150 TWSs. To keep BBU to this level, we need to minimize and/or suppress excitation of the higher order (transverse) modes in the TWSs.

The contribution of the output structures to the transverse instability of the beam is greater than that of the induction cell gaps for several reasons. The output structure has a smaller aperture to obtain the desired longitudinal shunt impedance. Damping of the higher order modes must not affect the fundamental mode used for power extraction. Also, the three cells comprising the TWS are electromagnetically coupled. Within a TWS, the regenerative BBU mechanism will increase the interaction of the higher order modes with the beam.

The basic design of the output structure has a transverse shunt impedance ($Z_\perp/Q$) of about $3\,\Omega$ per cell. This is significantly lower than for the induction cell gaps. However, the $Q_{\rm wall}$ of the cells is high, on the order of several thousand. The third cell has two output ports that remove energy from the higher order modes. This produces an effective $Q_{\rm ext}$ of about 15. In addition, fields in the first and third cell couple strongly to the TE_{11} mode of the connecting pipe, thus leading to a drastic reduction of Q in both of these cavities. The growth in the transverse instability is also reduced by Landau damping due to the energy spread on the beam. However, even with these instability reduction features, the high-frequency BBU growth is still too high.

The Betatron Node Scheme [Li 1994] was used to limit the growth in the transverse instability to an acceptable level. Figure A-9 shows the large difference in BBU growth as we vary the focusing field from the optimum for the Betatron Node Scheme. This is especially true for the case with a high-Q value for the first cavity. A list of the pertinent

BBU frequency	3.0 GHz
$Z_{\perp 1}$ (Ω/m)	5,400
$Z_{\perp}/Q$ (Ω)	42.972
Cell Q	2
Gap width	2.5 cm
Average beam energy	10 MeV
$\Delta\gamma(\gamma_{\max} - \gamma_{\min})$	3
Current (DC component)	0–600 A in 100 ns
	600 A flattop for 200 ns
Focusing system	Quadrupole - 2-m period
Time step	1/600 ns

Table A-2. *Input parameters for simulated induction cell gaps.*

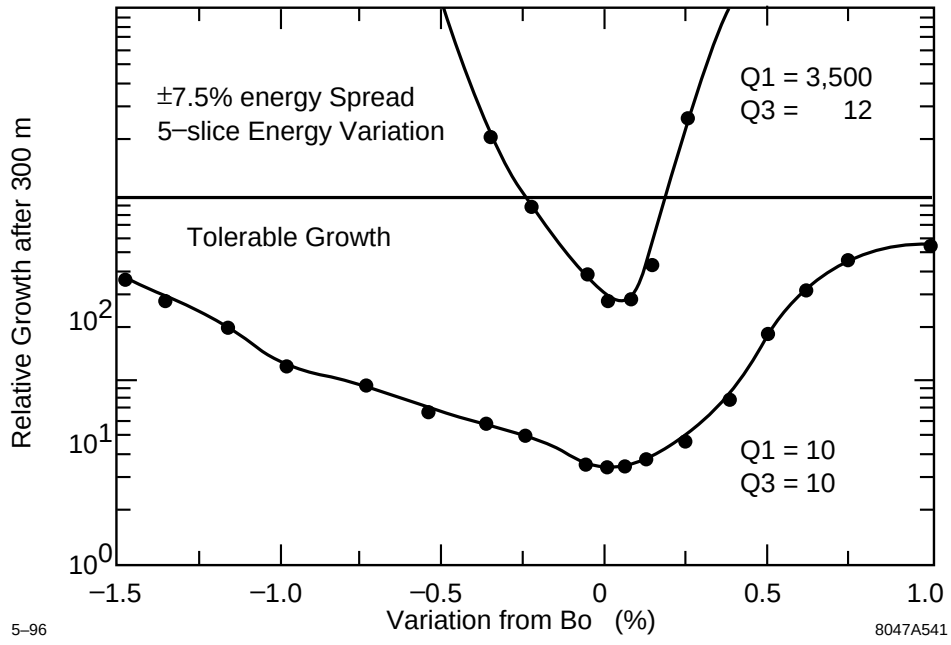


Figure A-9. *Relative growth after 150 TWSs vs. variation from the optimum focusing field. Comparison of an aggressively lowered Q to a high-Q for the first cell.*

Transverse mode	HEM_{11}
Frequency of mode ω_o	14.1 GHz
Number of cells/TWS	3
Electrical length of each cell L_c	8.754 mm
Phase advance/cell ϕ	120°
Group velocity v_g	0.25c
$Q_{\rm wall}$ (1st and 2nd cells)	3,500
Typical $Q_{\rm damped}$	15
$Z_\perp/Q$ for each cell	3 Ω
Current (DC component)	0–600 A in 100 ns 600-A flattop
Pulse length	300 ns
Average beam energy	10 MeV
Quadrupole field	0.0812 T

Table A-3. *Base parameters for simulated traveling wave structures.*

variables, and their base values, is given in Table A-3. The Betatron Node Scheme works well in principle; the key issue is the robustness of the scheme to deviations in the betatron wavelength resulting from errors in focusing field and/or beam energy, as well as sensitivities to cavity parameters.

The main effort of the numerical studies was to delineate the acceptable parameter space. The OMICE code was used to model the time-dependent transverse behavior of a 300-ns pulse over the 300-m length of the RK. An initial constant offset of the beam centroid was used as the seed perturbation for exciting the instability. Base parameters were varied individually to characterize the system's sensitivity to different design parameters such as effective impedance, Q values of the cells, quadrupole focusing field strengths, and variation in average beam energy.

The conclusion of this parameter study was that the Betatron Node Scheme is a promising means for controlling the high-frequency instability caused by these structures. Requirements on energy flatness and field accuracy are quite acceptable ($\pm$1%) provided that the output structures are indeed strongly de$Q$ed. Feedback schemes may further relax the requirements on field and energy accuracy.

A.3.3 RF Extraction Cavity Design

The most basic requirement for the cavity design is that it generate 360 MW of rf power for a drive beam current of 600-A DC and 1–1.5-kA rf current. The cavities must also be inductively detuned to maintain longitudinal beam stability over long distances. The surface fields of the cavity must be sufficiently low to avoid breakdown. Finally, the cavity must have low transverse shunt impedance in order to minimize the high-frequency BBU. The required cavity parameters have been discussed in Sections A.3.1 and A.3.2. In this section, we present electromagnetic calculations to determine the structure of the extraction cavities. We chose in this design effort one specific path to meet the general requirements. The procedure adopted is by no means unique, and we anticipate further optimizations and more detailed calculations in the future.

The numerical tools we have used are URMEL and SUPERFISH for 2-D frequency domain calculations, ABCI and TBCI for 2-D time-domain calculations, and MAFIA for 3-D frequency as well as time-domain calculations. The design procedure is carried out in several steps, starting with the simplest approximations, and adding more realistic features with each successive iteration. The successive approximations are summarized as follows:

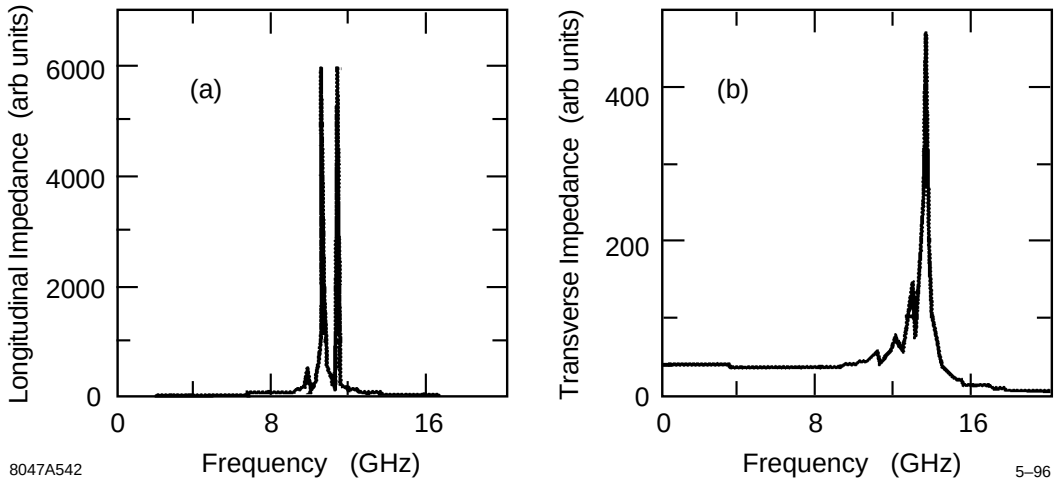


Figure A-10. *Longitudinal and transverse impedances of a three-cell structure with beam pipes.*

- Step 1. We construct a three-cell, disk-loaded, synchronous traveling-wave structure. Desired power extraction determines $(R/Q)/v_g$. URMEL is used to study an infinitely periodic structure, the inner radius a and outer radius b of the structure is varied, and the set of solutions with $v_p = c$ is obtained following the procedure of Thompson *et al.*, [Thompson 1993].
- Step 2. The inductively-detuned structure is constructed by a variation of Step 1. We want the resonant frequency to remain unchanged, but the wavelength increased by a factor of 1.33. The R/Q and v_g for the new configuration are determined with URMEL. The geometry is adjusted to ensure that the R/Q and v_g provide the right power extraction.
- Step 3. Finite beam pipes are then included to model the finite cell structure. The effect of the modified geometry on the field configuration and cavity parameters is then studied.
- Step 4. The 3-D aspects of the output ports are studied using MAFIA. The geometry of the output is varied to achieve the value of Q and ω for proper matching.

The relevant dipole cavity parameters for BBU considerations and the field enhancement factor for assessing the surface field are also determined.

Longitudinal and transverse impedances for the proposed extraction structure are shown in Figure A-10. The field pattern from a MAFIA run is shown in Figure A-11. The pertinent parameters of the structure are given in Table A-3. The determination of the external Q of the 3-D cavity with output structure is based on the Kroll-Yu method [Kroll 1990]. A $Q_{\rm ext} \approx 80$ was calculated for various waveguide iris apertures. The theory in Section A.3.1 was used to estimate the Q of individual cells. Initial estimates suggest that the transverse mode Q for the first and last cavities is as low as 10.

A.3.4 Induction Cavity Design

The interaction of the induction cell gap with the beam is a critical issue for the RK. There are three major criteria that the cell gap must meet: hold-off of the applied 100-kV voltage, low transverse impedance for BBU minimization, and

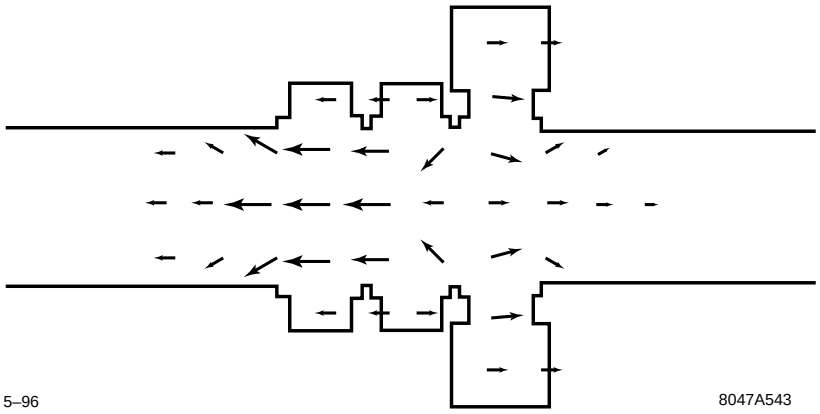


Figure A-11. *Field configuration for the $\pi/2$ mode of the 3-D cavity with output structure.*

Material	ε'	ε''	$\mu\prime$	μ''
Ferrite[a]	13.0	0.0	1.0	9.4
TDK[b]	25	3.3	1.9	2.1
Carbon-loaded ceramic[c]	5.9	1.07	1.0	0.0

[a] Ferrite is modeled after the ferrite used in the ATA induction cells [Deford 1990a], measured at 1 GHz.

[b] TDK Electronics Co., LTD., Absorber IR-B006, measured at 2.5 GHz.

[c] University of Maryland ceramic [Calame 1991], carbon concentration 0.34% (by weight), measured at 9.9 GHz.

Table A-4. *Electrical properties of simulated microwave absorbers.*

low longitudinal impedance at the beam modulation frequency and harmonics to minimize power loss. The gap should be as narrow as possible consistent with the maximum surface electrical fields. Larger apertures reduce the transverse impedances, but increase the core volume. An aperture of 5 cm was chosen as a base design. This allows sufficient room for focusing magnets while meeting the desired core volume.

Microwave absorbing materials can be used to damp resonant modes. These materials have a complex permittivity and/or permeability. Table A-4 lists nominal values of the permittivities and permeabilities for the absorbing materials used in the numerical modeling. The permittivity and permeability are expressed as $\varepsilon = \varepsilon_o(\varepsilon' + i\varepsilon'')$ and $\mu = \mu_o(\mu\prime + i\mu\prime\prime)$, respectively. Here ε_o and μ_o are the free space permittivity and permeability.

Several geometries were considered for the gap design. Several variations to each geometry was made to study the effect on the gap impedance. Simulated material properties of the absorbers and insulator were varied over a nominal range as well as their location and size. POISSON was used to adjust the geometry for specific designs to achieve acceptable surface electrical fields. The induction cell design code AMOS [Deford 1989, Deford 1990b], was then used to determine the rf characteristics of the gap. AMOS simulations include power loss to absorbing material. All geometries considered met the transverse impedance requirement of $Z_{\perp 1} \leq 5,400\Omega/\mathrm{m}$ for transverse stability by aggressively damping all resonant modes.

A conventional design similar to the ETA II induction cell gap was chosen as the base design for engineering and costing purposes. Advantages of the conventional design include demonstrated performance on several induction accelerators, a ceramic insulator hidden from the electron beam thus lowering the susceptibility for arcing, and relatively low-surface electrical fields. The disadvantages are a large inner radius for the core and a low-gap capacitance. A

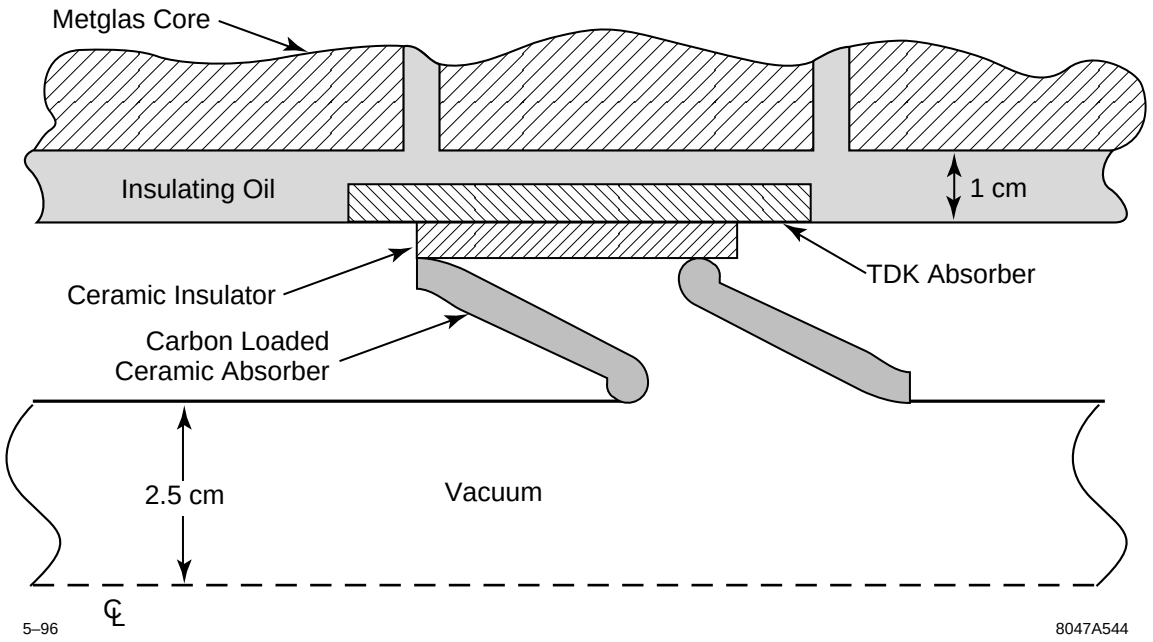


Figure A-12. *Schematic of a conventional design induction cell gap.*

low-gap capacitance is desired for fast rise times, but tends to be related to a high-longitudinal impedance which is undesirable. Figure A-12 displays a full-scale schematic of this design. Longitudinal and transverse impedances are shown in Figure A-13.

The most difficult design problem for all geometries studied was achieving a low-longitudinal impedance. The goal is for a power loss to the induction cells of $< 4\%$ of 360 MW per two-m module. This requires a $Z_{L0} < 2.4\Omega$ at 11.424 GHz. While impedance decreases rapidly above the cutoff frequency, the lowest Z_{L0} achieved thus far is about 10Ω at 11.4 GHz. Determining the longitudinal impedance accurately at 11.424 GHz is difficult. Analytical theory is available [Chattopadhyay 1990], but it is not clear that the quantitative accuracy is adequate. Numerically, the parameters of the test charge used, fineness of the grid, and the boundary conditions significantly impact the results. A cold test model needs to be fabricated and tested in the laboratory for a definitive answer.

A.4 TBNLC Engineering Design

A.4.1 TBNLC Systems Study Approach

The objective of the TBNLC system study was to assess the technical feasibility as well as the cost and efficiency of a TBA-based system. To achieve this goal, we felt that it was essential to be reasonably concrete. For this reason, our study specifically addressed a TBA power source for the NLC Upgrade considered by SLAC. The present thinking about the NLC is a 14.16-km active length linac with a center-of-mass energy of 500 GeV powered by X-band klystrons and rf pulse compression systems. This machine should be upgradeable to at least 1-TeV-c.m. Our study considered

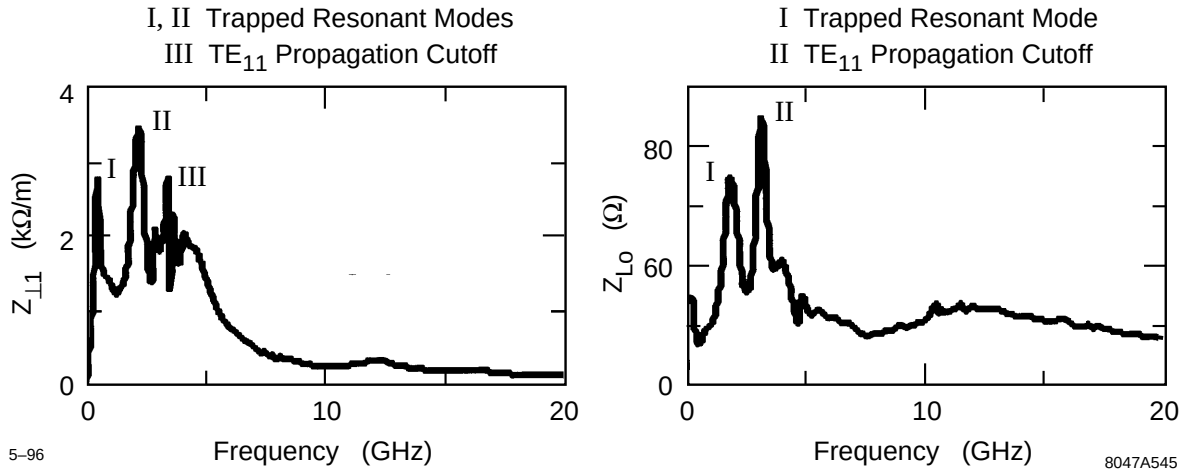


Figure A-13. *Longitudinal and transverse impedance of the induction cell gap.*

the TBA as a power source candidate for the 1-TeV version of the NLC. However, our modular architecture is directly applicable for a 1.5-TeV, ∼22-km NLC upgrade.

The engineering and costing efforts in the study drew heavily from recent work in Heavy Ion Fusion (HIF) [Hogan 1992]. The U.S. effort in HIF is based on the induction accelerator approach. Research activities in the past few years have centered around the Induction Linac Scaled Experiments [Fessenden 1992] (ILSE). Substantial engineering activities were committed to the ILSE CDR, and a 2-MV heavy-ion injector at full-driver scale was constructed and successfully operated in 1993. An engineering and costing study was performed recently for a recirculator version of the induction fusion driver.

A first engineering and costing exercise for the full TBNLC system was performed. The electrical design included all components starting from the AC power distribution system, to the DC power supplies, command resonant charging system, PFNs, and induction cores. Racks and installation, as well as instrumentation and control, were included in this exercise. The mechanical design and costing included details of the induction cells, rf structures, vacuum system, alignment, and utilities. Important aspects of the engineering design are summarized below.

A.4.2 Electrical Systems

A significant factor in the total efficiency of the TBNLC rf power source is the conversion efficiency of wall plug power into induction beam power. Figure A-14 is a schematic of an equivalent circuit of an induction accelerator cell. I_C is the core magnetizing current, I_B is the beam current, and I_N is any compensating network current. For optimum efficiency, I_N is minimized by designing the impedance of the PFN to match the nonlinear impedance of the induction core. The efficiency of an induction accelerator can approach 100% if the beam current is much greater than the current required to magnetize the transmission line (or autotransformer) which forms the induction cell. For example, the ATA induction cell required less than 1 kA of magnetizing current while accelerating a 10-kA beam by 250 keV for 70 ns; this corresponds to an induction cell efficiency of 91%.

The efficiency of an RK-TBA induction accelerator will depend on a number of factors. Beam transport dynamics will determine the size of the beam pipe and accelerating gradient. The rf power requirement will determine the pulse

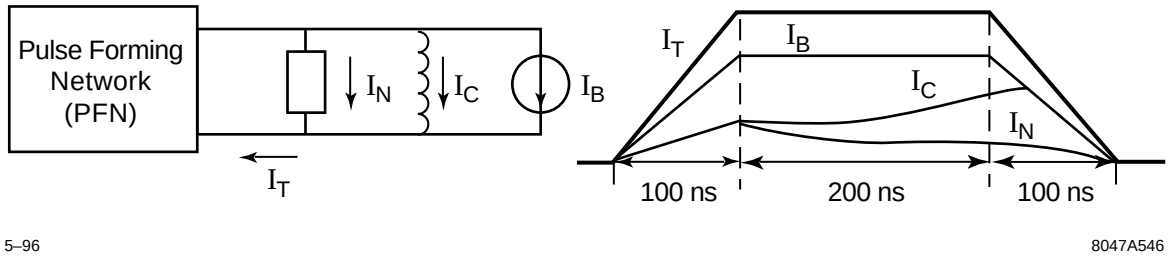


Figure A-14. *Equivalent circuit of an induction cell.*

duration, beam current, and repetition rate. Once these factors are set, the outer radius and flux swing of the core can be calculated from Equ. A.6.

Induction Accelerator Module

The linear induction accelerator can simply be described as multiple transmission line transformers driven in parallel by a pulse modulator which add energy to the electron beam in series as it passes through them. In order for the acceleration process to be efficient, the transformer or induction losses should be small compared to the energy imparted to the beam. Simply put, the transmission line current should be small compared to the beam current; therefore, the effective impedance of the transmission line, $Z_{\rm eff}$, should be greater than the beam impedance $Z_b = V_{\rm drive}/I_{\rm beam}$. In previously constructed induction accelerators, the high impedance is achieved by loading the transmission line with a ferri- or ferromagnetic material which is appropriate for the pulse duration. This material increases the impedance by $\sqrt{\mu/\varepsilon}$ and the electrical length by $\sqrt{\mu\varepsilon}$.

Typically, for induction accelerators with pulse durations of less than 100 ns, Ni-Zn ferrites have been used for the magnetic material due to their relative low saturation losses at very short pulse durations. However, the flux swing for Ni-Zn ferrite is only about one-fifth of that for Ni-Fe or METGLAS, and these materials are preferred for pulse durations of several hundred nanoseconds to microseconds. The optimization process in selecting the type of material and quantity is based on pulse duration, impedance, magnetizing losses, and economics. For the TBNLC design, the parameters are as follows:

$$I_{\rm beam} = 600\,{\rm A}\ , \quad V = 300\,{\rm kV}\ , \quad T = 300\,{\rm ns} \quad .$$

To satisfy the requirement that the effective impedance, $Z_{\rm eff}$, is greater than the beam impedance, $Z_{\rm beam}$ ~500 Ω, we can use the equation for a coaxial transmission line where $Z = \sqrt{\mu/\varepsilon}60ln(r_o/r_i)$. For ferrites, $200 < Z_{\rm eff} < 1000$, and for ferromagnetic materials $500 < Z_{\rm eff} < 5000$ depending on the type of material used and the magnetization rate. To satisfy the pulse duration requirements, the transmission line length should be about a single transit time. Since transit time is proportional to $\sqrt{\mu\varepsilon}$, for ferrites and ferromagnetic materials $T_{\rm eff} \simeq 100\ T_{\rm vac}$. The design of the transmission line (induction cell) must also ensure that no portion of the line is driven into saturation. This requirement dictates the outer radius of the line once the inner radius is determined by beam transport physics.

One of the most important considerations in the selection of a magnetic material for an induction cell is the actual losses during the magnetization pulse. Ferromagnetic materials such as nickel-iron have been available in thin ribbon cores since the late 50s when the first induction accelerator, the Astron, was constructed at Livermore by N.C. Christofilos. These iron-based alloys are very competitive in magnetic properties to the amorphous materials manufactured by Allied Signal and referred to as METGLAS. The time-dependent losses in ferromagnetic ribbons are typically explained in terms of a saturation wave which encircles the tape and proceeds toward the center at a rate

proportional to the applied voltage. The magnetic intensity required to change the state of magnetization is

$$H_a = H_c + (d^2/4\rho)(\Delta B/2B_s)(\Delta B/\Delta t) \quad , \tag{A.8}$$

where H_c is the DC anisotropy coercive field, d is the ribbon thickness and ρ is the resistivity. The energy density deposited in the tape during saturation is given by

$$E_L = H_c\Delta B + (d^2/4\rho)(\Delta B^2/2B_s)(\Delta B/\Delta t) \quad . \tag{A.9}$$

At high-magnetization rates, one can see that the losses are proportional to the thickness of ribbon squared and inversely to the resistivity. Since the ΔB of the iron alloys is similar to that of METGLAS, but the resistivity of METGLAS is three times that of the iron alloys, the ribbon thickness must be thinner in order for the losses to be comparable for METGLAS. The Ni-Fe alloys are manufactured by a rolling technique which yields a more uniform cross section while METGLAS alloys are manufactured by rapidly quenching a mixture of silicon and iron which remains amorphous. The quenching process leaves the ribbon surface rough compared to the rolled ribbon, hence it is more difficult to insulate and can yield an uneven cross section in the core winding process. Currently, the unannealed 20-μm METGLAS is wound into cores with 2.5-μm mylar between layers yielding a packing factor of 0.7 to 0.8 for a finished core. The iron alloys can be rolled into thinner ribbon than the METGLAS, and have produced lower overall losses. However, the manufacturing and winding process for METGLAS usually results in less expensive cores.

The induction cell will be driven directly by the modulator, *i.e.,* without step-up transformers. The modulator design is described below. To obtain 300 kV, 15 induction cells will be required. Since a standard width for the METGLAS is two inches, the TBNLC induction cell design was based on this width. The METGLAS alloy used for the design is 2714 AS. An example of a core test for the 2714 AS is shown in Figure A-15. The minimum amount of core material is realized when the flux swing approaches saturation. Minimizing the amount of magnetic material, however, does not lead to the most cost-effective design. A flux swing approaching saturation requires a very nonlinear drive current. This leads to a PFN which is difficult to design to achieve a flat voltage pulse. Furthermore, since the losses per unit volume are nearly proportional to ΔB^2, a better design is achieved by allowing a flux swing of less than one half of the saturation value. In this case, a ΔB of 1.0 T results in a dB/dt of 3.3 T/ μs and losses per unit volume of about 150 J/m^3. The inner radius of the induction cell is dictated by beam transport physics and is 4.5 cm for our TBNLC design.

Equation A.6 can be used to determine the required outer radius of the core. With ΔB = 1.0 T, Δt = 300 ns, G = 300 kV/ m, R_i = 4.5 cm, core occupancy $\eta_c = 0.75$, and a packing factor $F_p = 0.75$, we find that R_o = 20 cm and the volume of METGLAS is 4.5 10^{-3} m^3. The losses in joules are $E_L = (150\,\mathrm{J/\,m^3})$ (4.5 10^{-3} m^3) or about 0.68 J per core. The drive voltage for the induction core is 20 kV and the average current dissipated can be estimated from the losses, $E_L = VI_C\Delta t$, or from Eq. A.7. For a 300-ns pulse, the average magnetizing current or drive current of the core I_C is 117 A. The magnetizing current actually has a nonlinear component, as seen in Figure A-15, which reflects the characteristics of the B-H loop. The total drive current (I_T) from the PFN is the sum of the beam current (I_B), the core current (I_C), and the compensation network current (I_N), which for our design is approximately $I_T = I_B + I_C + I_N = (600 + 117 + 33)$ A $= 750$, yielding a core efficiency of 600/750 = 80%.

Line Modulator

The modest repetition rate (120 Hz) and current rise time (100 ns) envisioned for the NLC permits the use of a simple and cost-effective thyratron-driven modulator. For the TBNLC, each induction cell is comprised of five cores individually driven at 20 kV. Driving at this voltage level avoids any step-up transformers and can be generated directly by a thyratron with a 40-kV charging voltage on the PFN. After a preliminary search, the English Electric Valve (EEV) CX1525A appears to be an excellent thyratron for our application. It is a two-gap, deuterium-filled thyratron with a

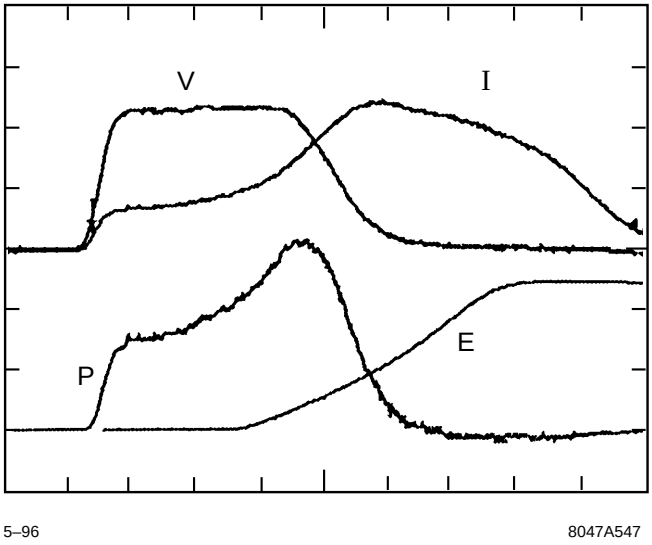


Figure A-15. *Oscilloscope traces for pulsed METGLAS 2714 AS core with cross section of 12.7 cm^2, volume of 450 cm^3, and packing factor = 0.7. Time scale = 100 ns/div. V is applied voltage (1 kV/div), I is current dissipated in core (50 A/div), P is power (VI, 50 kW/div), and E is energy (time integrated power, 20 mJ/div, 250 ns offset). Total energy dissipated = 47 mJ, or 150 J/m^3.*

voltage hold-off of 50 kV and capable of delivering 15 kA at 120-Hz repetition rate. Such a device would allow us to drive 15 cells for a total of 300 kV. A simplified schematic of the line modulator is shown in Figure A-16.

As shown in Figure A-15, the current drive to the cores is nonlinear, but a constant amplitude pulse can be generated, within bounds, simply by tapering the impedance of the PFN stages. The PFN will consist of many coupled L-C stages with impedances adjusted to temporally match the induction core impedance. The PFN charging current flows through the induction core, resetting the core prior to the next acceleration cycle. The combination of 2714AS METGLAS cores and the thyratron-driven modulators will result in a conversion efficiency of wall-plug power to cell power of 62%.

Command Resonant Charging and Core Reset

During the energy delivery cycle to the beam, the magnetic core requires a large magnetic intensity to swing 1.0 T in 300 ns. After the pulse is over, the magnetic intensity is reduced to zero, but the flux density has a remnant field, B_r. In order for the core to be useful on the next forward or energy delivery cycle, a reversed magnetic intensity or reset must be applied to the core. The reverse magnetic intensity, however, is much smaller than the forward one, since the rate of demagnetization can be much lower or $\Delta I_{\rm reset} \ll \Delta I_{\rm forward}$. By applying the reset current, the flux density is returned to $-B_r$, and the core is ready for the next forward or acceleration pulse. The PFN must be recharged after each forward cycle. The charging process occurs when the command resonant charging (CRC) switch is closed and delivers a half-sine current pulse from the large energy storage capacitor, C_E, to the PFN capacitors. It can be seen from Figure A-16 that the charging current for the PFN flows through the induction cells in the proper direction to reset them. By arranging the impedance and charge time, the proper reset current can be delivered to the cores at the same time that the PFN is being recharged. The reverse magnetic intensity required for reset is about 25 A/m. This is equivalent to about 24 A per core or 360 A for a complete induction module. Since one CRC charges six line modulators, a total of 2.2 kA are required. At the charging voltage of 40 kV, the reset/recharge impedance is 40 kV/2.2 kA= 18.2 Ω. The total

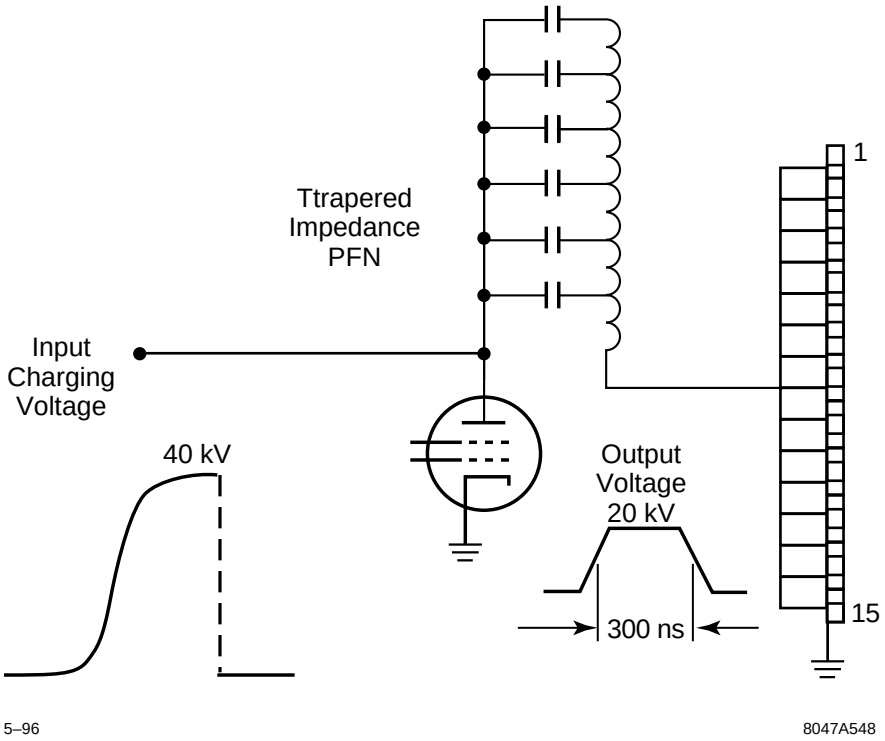


Figure A-16. *Simplified schematic of the line modulator. The PFN will have the same temporal impedance as the nonlinear magnetic core of the induction cells.*

capacitance (C_T) of six line modulators is 0.66 μF. The resonant impedance, $Z_r = \pi\sqrt{L_T/C_T}$, requires a charging inductor $L_T = 225\ \mu H$. Hence, the reset/recharge period $T_R = \pi\sqrt{L_T C_T} = 38\ \mu s$.

Efficiency of Induction Accelerator Components

The power conditioning system for the low-energy accelerator has energy losses associated with each major component from the utilities feeding the DC power supply to the induction cell coupling to the electron beam. These losses will be described in reverse order beginning with the induction core to beam coupling as shown on Figure A-17.

The acceleration voltage pulse is shown in Figure A-16. It has a rise time and a falltime of 100 ns with 200-ns flattop. There is no useful energy generated during the falltime, hence, this is lost. The energy during the rise and fall of the pulse is:

$$E_r = \int_o^{t_1} V \cdot I \ \ dt \quad , \tag{A.10}$$

where $V = V_o(10^7 t)$ and $I = I_o(10^7 t)$. By taking the integral, we find that $E_r = 10^{14} V_o I_o (t^3/3) = 0.4$ J, while the total energy during the pulse is $E_t = 3.2$ J . The induction core to beam efficiency is then $\varepsilon_b = 2.8/3.2 = 87.5\%$.

The magnetic materials losses have been discussed above. The magnetization current for each induction core was estimated to be 117 A. Another 33 A of current was added for compensation network at the core for a total current drive of 750 A. Hence, the efficiency of the induction accelerator cell is $\varepsilon_c = 600/750 = 80\%$.

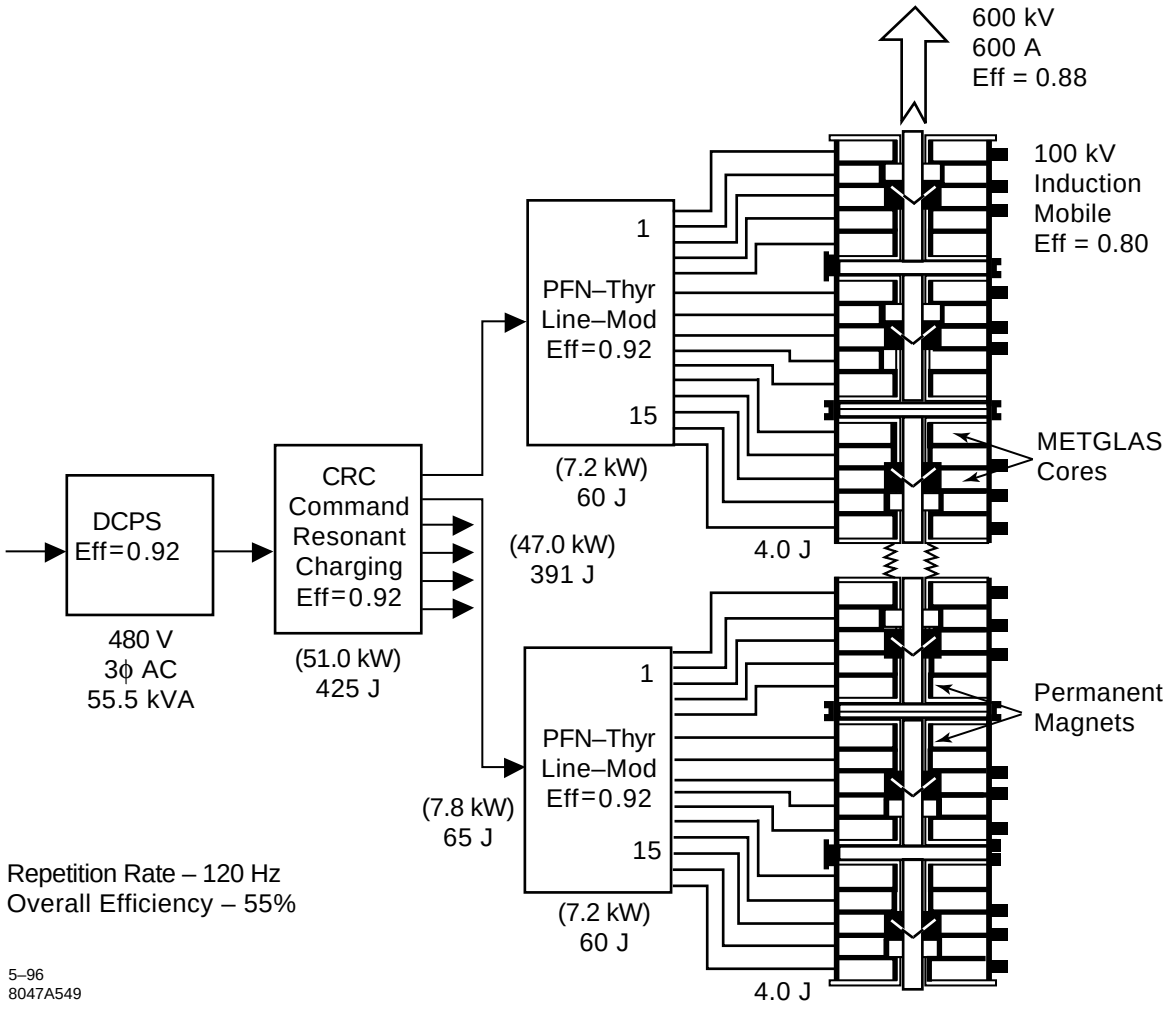


Figure A-17. *Schematic of the power conditioning system for a two-m RK module section.*

As previously discussed, the modulator consists of an impedance-tapered PFN matched to the drive required by the magnetic material, METGLAS 2714 AS. The efficiency is calculated by including the thyratron dissipation, filament, grid bias, keep-alive, and the pulse-shaping resistors losses. The modulator losses add up to nearly one kilowatt. At 120-Hz operation, the total power input has been calculated at 10.6 kW which yields an efficiency ε_m = 92%.

The line modulator is charged by a solid-state CRC system. The CRC generates a $1-\cos\omega t$-voltage waveform which charges the PFN. The current in the CRC is a half-sinusoid which flows through the induction cells in the resetting direction while the PFN is being charged. The losses for the CRC include the inductor mode-damping resistors, the solid-state switch losses, the trigger generator, and the de$Q$ing or regulation system if the DC power supply has insufficient regulation. The overall efficiency of the CRC is calculated to be $\varepsilon_{ch} = 92\%$.

The 20-kV DC power supply will be a conventional 60-Hz, 3-phase, full-wave rectifier with filter. It is projected that the power factor for the power supply will be 0.9 for an efficiency $\epsilon_{\rm ps} = 92\%$.

The overall efficiency of the induction accelerator from the utilities to the electron beam which drives the RK is the product of the individual component efficiencies or $\varepsilon_o = 55\%$.

A.4.3 Mechanical Systems

Module Design and Fabrication

The TBA consists of essentially 50 identical RKs stacked end-to-end to form the 15-km overall length. Each of the 300-m-long RKs contain 150 identical extraction and reacceleration sections. These 2-m-long sections are referred to as RK modules. A schematic of a 2-m RK module is shown in Figure A-3. An RK module consists of six identical induction modules and one extraction cavity. Each 300-m RK contains 900 modules and the complete TBNLC will contain 45,000 identical modules.

In the design of the module a large effort was made to keep the geometry of individual parts simple and amenable to mass production. This basically requires designing so that a minimum amount of material has to be removed during part fabrication. Drilling, tapping, and machining of small slots are kept to a minimum as these operations are relatively slow operations. All metal parts are made of 304 stainless steel alloy. Advances in modern stainless steel alloys enable significantly greater cutting speeds. Metal stamping and molding of plastic parts are used to further reduce part costs.

The module assembly relies on brazing and welding. The beam pipe assembly uses six brazed joints. Use of simple fixtures to align the stacked parts and brazing in large batches keeps the time and cost per assembly very low as compared to other methods of assemblies. The final housing assembly is done with an inner and outer weld at each end plate. Automatic welding machines make this a quick and reliable operation.

Core Winding

An individual core winding is made by winding approximately 4700 turns of METGLAS ribbon onto a winding mandrel. The ribbon is 20-μm thick and 51-mm wide. A thin layer of insulation will be dipped or sprayed onto the METGLAS during the winding process to provide adequate resistance to eddy currents between turns. The average voltage between turns is 5. This insulation will replace the thin mylar ribbon presently used to provide core interturn insulation. The mylar ribbon overhangs the core edges and presents a serious problem to edge-cooling the cores. A technique being developed at LBNL uses a thin layer of epoxy, dipped or sprayed on, and then dusted with alumina grit to form a tough interlayer insulation. The epoxy additionally binds the METGLAS layers together to form a rigid self-supporting core.

Core Cooling

Heat generation within the cores will be about 300 W per core. Cooling of the core windings is accomplished by circulating oil through the module housing. Oil flows in at one end of the module housing and flows out at the other end on the opposite side. This causes the oil to flow transversely in the gaps between the cores. Spacers around the outside diameter of the cores prevent oil from flowing around the outside of the cores. Only moderate oil flow rates are required and as a result there is a low-pressure drop between supply and return manifolds. For a flow rate of 27°C oil at 1.25 gpm through the module, the maximum temperature for all the cores is about 60° C on the down-stream side. Quadrupling of the oil flow rate will drop the core temperature by about 13° C. Core temperatures in this range

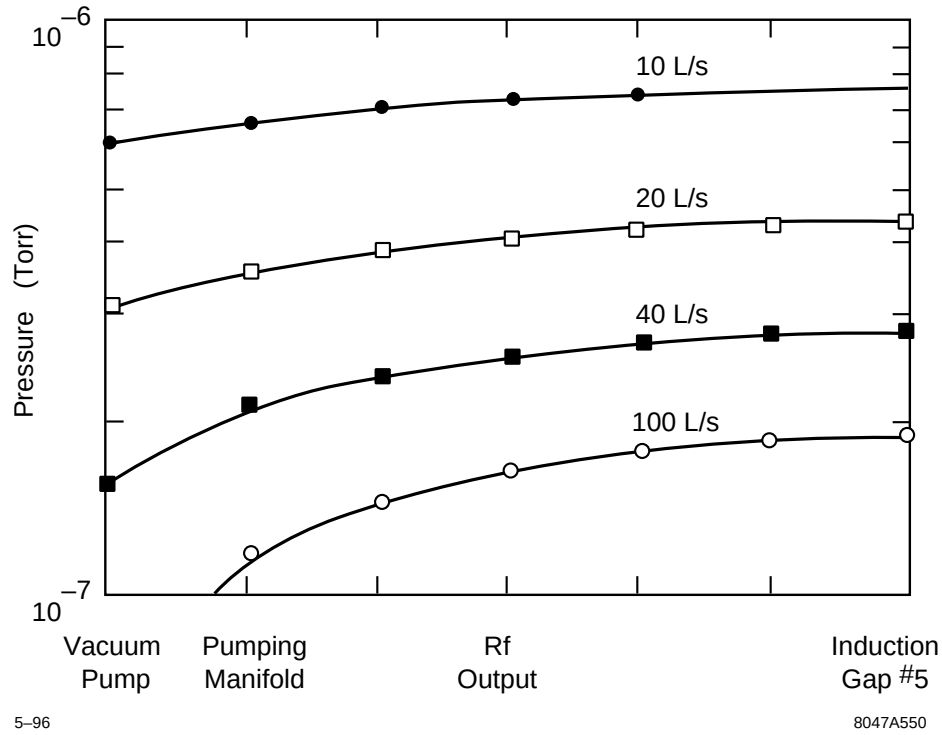


Figure A-18. *Gas pressure distribution within the 2-m module.*

are acceptable; magnetic properties are not affected, thermal stresses are minimal since the core remains at a nearly uniform temperature, and temperature limits on the plastic materials within the module are not approached.

Development of an interlayer insulation to replace the presently used mylar tape is critical for the cooling of the cores. Presently used mylar tape overhangs the edges and largely blocks the transfer of heat from the METGLAS to the oil. As described in Section A.4.3, "Core Winding", alternative interlayer insulations are being developed.

Vacuum System

The accelerator will be pumped in two places in each 2-m module. Pumping will occur at the start of each module on each side of the extraction cavity. There is a 2-cm gap between modules at these points which permits good pumping conductance to a standard mounting flange. The two ports are manifolded together to a single ion vacuum pump. The total gas load per cell block is about 8×10^{-7} Torr-ℓ/s. For this geometry, the pressure between the pump and the rf output structure entrance is calculated for several pumping speeds and results are shown in Figure A-18. In this figure, Induction Gap #5 refers to the most distant induction cell from a pumping manifold. If a 20-ℓ/ s pump is chosen, the pressure at the rf output structure entrance will be about 5×10^{-7} Torr.

An inline gate valve between every fourth 2-m module will enable isolation of small sections of the accelerator. In each of these four 2-m modules a right-angle gate valve will be used to enable initial pump down with a portable turbomolecular pumping station.

Moderate pressure levels within the linac will enable some of the more costly cleaning methods used on ultrahigh vacuum systems to be avoided. This will be an unbaked system with a base pressure in the 10^{-7}-Torr range. Systems of this type are generally cleaned by chemical or glow discharge cleaning techniques and perhaps a bake at low temperatures.

Assembly and Alignment

Alignment requirements for the TBA can be achieved using standard fabrication, assembly, and alignment techniques. The centerlines of the 2-m modules will be aligned to within 0.1 mm of their nearest neighbor's centerlines. A gradual accumulation of position errors between successive 2-m modules will occur, and will result in the accelerator not being straight. This is acceptable as long as these bends are not extreme and occur gradually over many half-lattice periods.

Six induction cells will be supported and aligned accurately with respect to one another on one strongback support beam. As shown in Figure A-4, each cell has two quadrupole assemblies. The permanent magnets will be sorted by field strength to optimize paring in the quadrupole assemblies. The magnets will then be bonded into the assemblies that are held in place by the cell end plates. Quadrupole correction coils will be used to minimize the field errors, instead of attempting to adjust the magnetic center of individual quadrupole assemblies. These procedures will require the permanent magnets to be positioned with a tolerance of 0.1 mm within the 2-m modules. The required tolerance for each step of the assembly will be 0.05 mm. This will ensure that the quadrupole assemblies are well centered with respect to the outside diameter of the cell housing end plates.

The support mounts for all the induction cells on each strongback beam will be positioned and aligned accurately during fabrication of the beam. Mounts for the modules will be welded, machined, and then ground to the required precision. As a result, within each 2-m module, the 12 magnetic quadrupoles will be aligned to the required precision when the modules are assembled onto the beam. There will be no provision made for adjusting the position of magnets within a 2-m module with respect to each other on the beam.

There will be an X and Y set of fiducials at each end of the beam. After the installation of the 2-m modules onto the beam line, a precise measuring fixture will be used to adjust the position of each fiducial accurately with respect to the center of the last quadrupole at each end of the beam. It is desirable to make the fiducials reflect the position of the end quadrupoles rather than some average position for each 2-m section. This will enable the last quadrupole on one beam to be aligned well to the first quadrupole on the next beam.

As successive 2-m modules are installed, they will be aligned using the fiducials to a straight-line reference system. Each module is supported and articulated using a six-strut system. This enables the section to be moved easily and with precision in all six degrees of motion. The straight-line reference system will be able to accurately detect the position of each fiducial. The 2-m module is then articulated to bring the fiducials into a straight line.

A stretched wire can be used with optical sensors to detect the position of the wire. A laser-based system can be used although a pipe under vacuum is necessary for the light. A series of overlapping straight reference lines will be needed to form the entire 300-m length.

A.5 RTA Test Facility

We are constructing a prototype RK-TBA rf power source to verify the analysis of the TBNLC system study. This prototype, called the RTA, will be located at LBNL. All major components of the TBNLC rf power source will be tested. However, due to fiscal constraints, the prototype will have only 8 rf output structures, with a possible upgrade

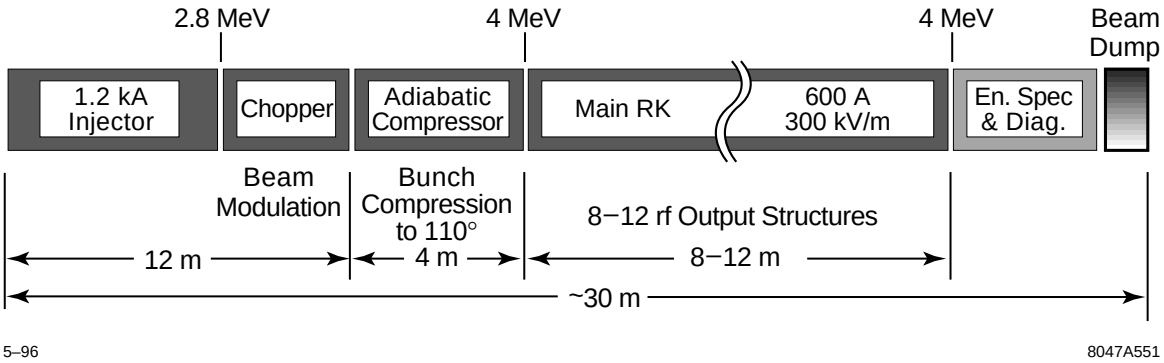


Figure A-19. *Schematic of the RTA showing major components.*

to 12, instead of the 150 envisioned for the TBNLC. A schematic of the RTA is shown in Figure A-19. Table A-5 lists pertinent parameters for the RTA and TBNLC power source.

The more important issues to be addressed by the RTA are efficiency, longitudinal dynamics, beam stability, emittance preservation, and rf power quality. Efficiency can be separated into the conversion efficiency of wall plug power into beam power and beam power into rf power. The conversion of wall plug power into beam power is described in Section A.5.1 and can be fully measured in the RTA.

High conversion efficiency of beam-to-rf power can be obtained in a system with a large number of output structures. For the TBNLC rf power source, the number of output structures will be limited by beam stability and transport issues. The direct study of beam dynamics issues involving the beam transport through many tens of output structures will not be possible with the prototype. However, the reduced beam energy in the extraction section of the prototype permits the observation of almost an entire synchrotron period. This will be sufficient to allow the beam to approach a steady state condition that can then be extrapolated to a full-scale system with high confidence. The verification of computer simulations used to model the beam dynamics in the TBNLC system study will be a high priority. Beam dynamics issues related to transverse modulation, misalignment of magnetic focusing systems, and adiabatic compression, *e.g.*, emittance growth and corkscrew motion, can be adequately studied.

A.5.1 Induction Cores and Pulsed-Power System

For our prototype, we will use two different pulsed-power systems. The adiabatic compressor and extraction sections will use induction modules and a pulsed-power system very similar to those described in Section A.4.2 and the TBNLC system study. The primary difference is that each induction module is comprised of three individually driven cores. The cores are still driven at 20 kV by a thyratron charging a multistage PFN. This part of the experiment will be used to verify the efficiency, technical aspects, and cost of the induction modules in the TBNLC.

As described below, the injector is comprised of modified versions of existing equipment. Here, the issue is generating the required volt-seconds within the geometrical constraints of these components. The magnetic material will be 2605 SC METGLAS to maximize flux swing, and the cores in each module will be driven as a single unit to maximize the available cross-sectional area. Driving multiple cores will require the use of a step-up transformer to deliver the required 80–100 kV per cell.

Parameter	RTA	TBNLC
Pulse		
Duration	200 ns	300 ns
Rise Time	50 ns	100 ns
Current:		
Pre-modulation	1,200 A	1,200 A
Extraction section	600 A DC	600 A DC
	1,100 A rf	1,150 A rf
Beam energy:		
Injector	1 MeV	1 MeV
Modulator	2.8 MeV	2.5 MeV
Extraction	4.0 MeV	10.0 MeV
Bunch compression	240°-110°	240°-70°
Extraction section		
PPM quadrupoles:		
Betatron period	1 m	2 m
Lattice period	20 cm	33.3 cm
Phase advance	72°	60°
Occupancy	0.5	0.48
Pole-tip field	870 G	812 G
Beam diameter	8 mm	4 mm
Rf power:		
Frequency	11.4 GHz	11.4 GHz
Power/structure	180 MW	360 MW
Structures	SW & TW	3 cell TW
Output spacing	1 m	2 m

Table A-5. *Comparison between RTA and the TBNLC.*

A.5.2 Injector: Gun and Accelerator Sections

Two main goals of the injector design are minimizing electrical field stresses in the gun and realizing the lowest possible emittance growth. The gun and accelerator section comprising the injector, will be modifications of the SNOWTRON induction injector and induction cells from ETA II. Modifications include replacement of the original ferrite cores with METGLAS and a new mechanical alignment system. The present electrode package will be used during initial testing. However, a new electrode package and larger dispenser cathode are required to produce the desired 1.2-kA, 1-MeV beam. The solenoidal field configuration must be optimized for the injector to control the beam radius while minimizing emittance growth. The design goal for the experiment is for a radius < 5 mm and $\epsilon_N < 250\pi$-mm-mr at the chopper entrance.

Alignment of the focusing solenoids is critical to avoid corkscrew motion and emittance growth in the injector. A stretched wire alignment scheme [Griffith 1990] will be used to determine the offset of the solenoid's magnetic axis from a reference mechanical axis for each cell. From past experience, the resolution of this alignment scheme to offset errors is approximately ± 0.05 mm. For offset errors of more than 0.5 mm, the solenoids will be repositioned in the cells. Offset errors of less than 0.5 mm will be recorded, but the solenoids will not be repositioned. A 0.5 mm

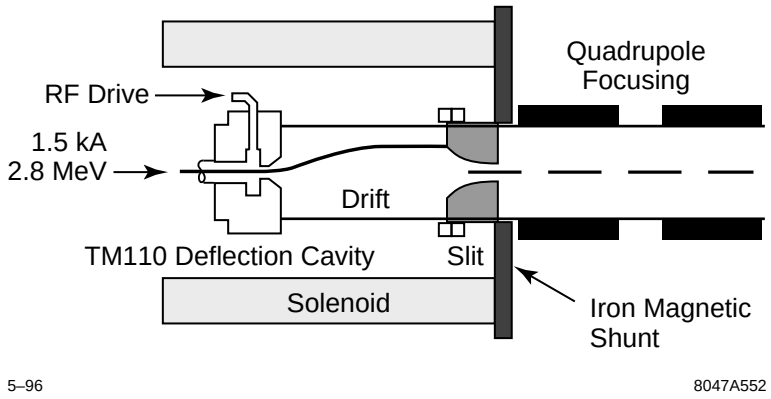


Figure A-20. *Schematic of the beam modulator (chopper).*

represents the maximum one induction cell can be mechanically offset from an adjacent cell to achieve magnetic alignment. Each cell also contains a steering (sine/cosine) coil to correct for tilt errors. The required tilt correction will also be determined. The entire alignment will be performed in a precision mill with tolerance on the order of tenths of a mil. A fiduciary will be placed on the outer case of the cell to permit alignment of the magnetic axis when the cells are mounted on the strongback. With this procedure, we expect solenoid offset errors of less than ± 0.08 mm and negligible tilt errors.

Experience operating the ETA II accelerator has shown that careful alignment of the solenoids is not sufficient to reduce the amplitude of the corkscrew motion [Allen 1991] to 0.5 mm desired for the RTA injector. Individual adjustments for the induction cells will permit improved solenoid alignment in the RTA. However, we anticipate using a time-independent steering algorithm [Chen 1992] developed for ETA II to control steering coils on the solenoids. This algorithm corrects for the Fourier component at the cyclotron wavelength of the field error, and led to an order of magnitude reduction in the corkscrew amplitude of the ETA II beam.

A.5.3 Chopper: Beam Modulation

A transverse chopping technique will be used to modulate the beam. The modulator section of the Choppertron, a 11.4-GHz rf generator, will be refurbished for this purpose. A schematic of the modulator is shown in Figure A-20. The solenoidal field immersed incoming electron beam is deflected in the horizontal plane by a 5.7-GHz TM_{110} deflection cavity causing the beam to describe semihelical trajectories along the drift space. The beam scans in a vertical plane across an on-axis aperture placed a quarter betatron wavelength after the deflection cavity. Thus, the 5.7-GHz spatially-modulated DC beam incident on the aperture becomes a phase-coherent, amplitude-modulated beam at 11.4 GHz.

The desired bunch length, peak current, and energy for the drive beam in the extraction section of the prototype is respectively 110°, 600 A, and 4 MeV. Such a train of bunches could be generated by directly chopping the unmodulated beam. However, considerations of efficiency (70% of the beam would be lost) and practical feasibility (we expect approximately 1.2 kA of peak current from the injector) require that we do not fully modulate the beam by chopping. Our intent is to chop the beam at an energy of 2.8 MeV into bunches of approximately 240°. An adiabatic compressor section will be used to further bunch and accelerate the beam.

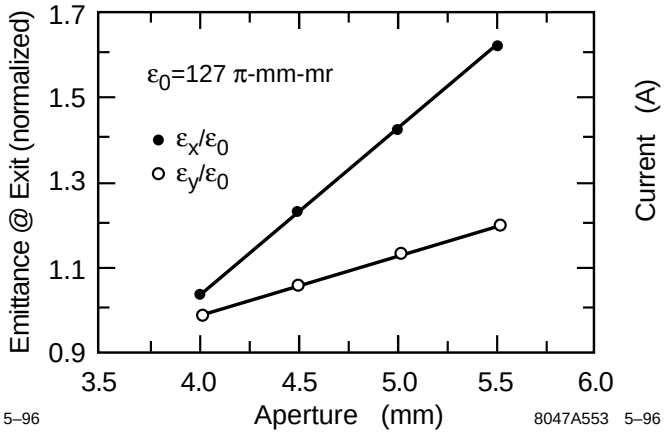


Figure A-21. *Emittance growth as a function of aperture size*

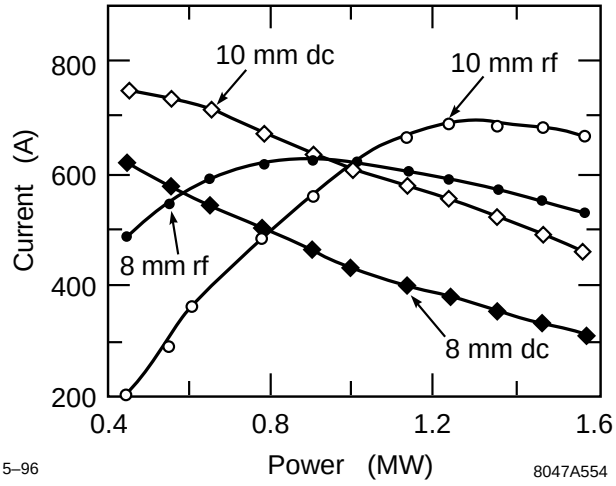


Figure A-22. *The DC and rf (11.4-GHz) current components as a function of drive power to the deflection cavity for two apertures.*

Designs of chopper systems have been extensively analyzed [Haimson 1965, Haimson 1970], and the original Choppertron was optimized for operation with the ETA II induction beam. The intent for the prototype experiment is to limit modifications of the modulator section of the Choppertron to adjustments in the drift length and the chopping slit aperture size. It is desirable for the radius of the beam to remain relatively constant in the chopper. For a given beam energy, current, and emittance, the radius is determined by the solenoidal field. Once the solenoidal field is determined for the desired beam radius, the drift section length is also fixed. The maximum deflection amplitude at the aperture is determined by the transverse momentum imparted to the beam by the deflection cavity. The deflection amplitude and chopping slit aperture determine the emittance growth and the modulated beam's characteristics.

A series of simulations were performed with the relativistic klystron design code RKS to determine the optimum deflection and aperture. Results are shown in Figures A-21 and A-22. The beam was assumed to be cylindrical with a maximum radius of 3.6 mm, waterbag distribution, and normalized, rms emittance (ε_x and ε_y) of 127 π-mm-mr at the chopper entrance. In Figure A-21, the emittance has been normalized with respect to the initial value. The emittance represents the area in trace space for all the particles within a 360° rf bunch length. The emittance increases with the chopping slit aperture as the simulations cover the range where a substantial current loss occurs on the aperture. Thus, the beam has been deflected to fill the acceptance of the aperture. The difference between the x and y emittances is due to the deflection in a single plane. In the simulations, x represents the horizontal plane.

Considering only the increase in emittance, it would appear that the smaller aperture is better. However, as shown in Figure A-22, the amount of current contained in a 240° bunch length is substantially reduced as the chopping slit aperture is reduced. We require at least 600 A of the initial DC current to remain in this bunch length for the adiabatic compressor section. To allow for current losses during the adiabatic compression, 660 A is a practical limit. Also, the beam must be sufficiently modulated for the idler cavities in the adiabatic compressor to function. The rf current component (modulation) initially increases with drive power, as shown in Figure A-22, with a decrease in DC current. The design goal of the chopper is to generate a 240° rf bucket containing 600 A that can be captured and bunched by the adiabatic compressor, with minimum emittance growth. Maximizing the rf current at the exit of the chopper will not optimize emittance growth with respect to DC current in the rf bucket. A 10-mm aperture at a drive power of about 0.8 MW produces a satisfactory modulation for growth in emittance, although further studies are required for an optimum design. Future simulations will use the results of EGUN simulations for the initial beam characteristics.

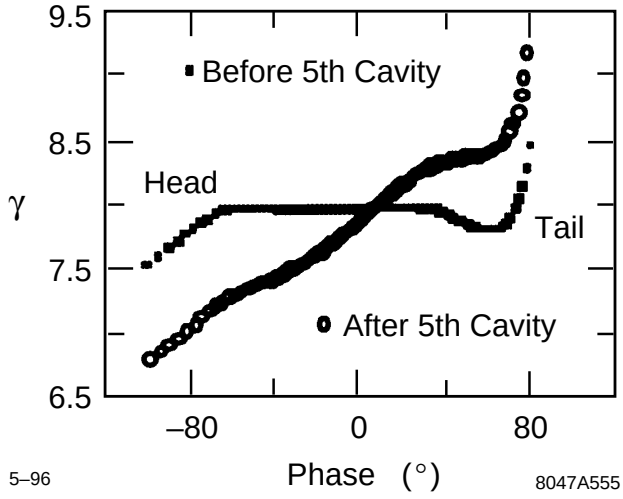


Figure A-23. *Simulations of longitudinal phase space before and after the 5th idler cavity.*

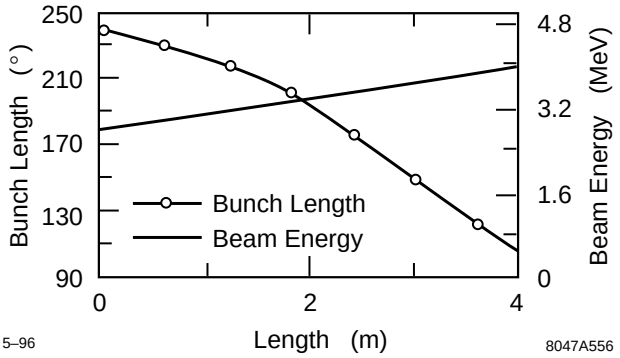


Figure A-24. *Simulations of adiabatic compressor effect on bunch compression and beam energy.*

A.5.4 Adiabatic Compressor

Adiabatic compression is used to achieve the desired beam characteristics at the entrance of the extraction section. In the adiabatic compressor, the bunch length is reduced from 240° to 110° with SW idler cavities while the beam is accelerated to an energy of 4 MeV. The accelerating gradient of the induction cells (300 kV/m) sets the minimum length of the adiabatic compressor to 4 m. Extensive 1-D numerical studies have been performed to determine the most efficient scheme for accomplishing this: The present design uses seven idler cavities appropriately spaced and detuned to progressively bunch the beam. The idler cavity has a resonant frequency higher than the drive frequency. The rf field in the cavity, excited by the beam, places itself at 90° with respect to the rf bucket. This leads to a loss of energy for electrons towards the front of the rf bucket, and a gain for those in the back.

The bunching effect of the idler cavity is illustrated in Figure A-23. As the beam drifts between the idler cavities, electrons in the tail of the rf bucket have a greater velocity leading to increased bunching. At the same time, space charge forces will increase the energy of electrons at the head of the bunch with respect to the tail slowing the bunching process. Appropriately designing and spacing the idler cavities can eventually bunch the bucket to the desired length. Table A-6 summarizes the important parameters of the adiabatic compressor, and Figure A-24 shows the overall effect on the beam.

A.5.5 RF Power Extraction

After leaving the adiabatic compressor, the beam enters the main RTA where power is extracted. Here the beam energy is periodically converted into rf energy (via output cavities) and then restored to its initial value (via induction modules). Stable propagation of the rf bucket traveling through many resonant cavities and achievable power output extraction have been studied numerically. Space-charge effects and energy spreads due to rf fields in the output structures tend to debunch the beam as it traverses the main RTA. To counteract this effect, inductively detuned output structures are used.

Cavity	Position (m)	Bunch length (°)	Beam E_k (MeV)	Resonant freq. (gHz)
1	0.0	240	2.8	11.82
2	0.6	229.7	2.98	11.82
3	1.2	217.9	3.17	11.82
4	1.8	201.7	3.36	11.74
5	2.4	177.6	3.53	11.74
6	3.0	148.6	3.69	11.74
7	3.6	122.0	3.85	11.74
exit	4.0	105.5	4.00	

Table A-6. *Parameters of the adiabatic compressor.*

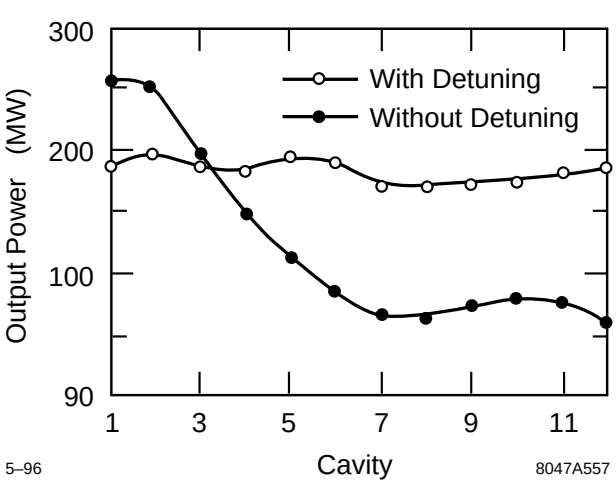


Figure A-25. *Simulations of the rf power generated per output, both with and without detuning of the output cavities.*

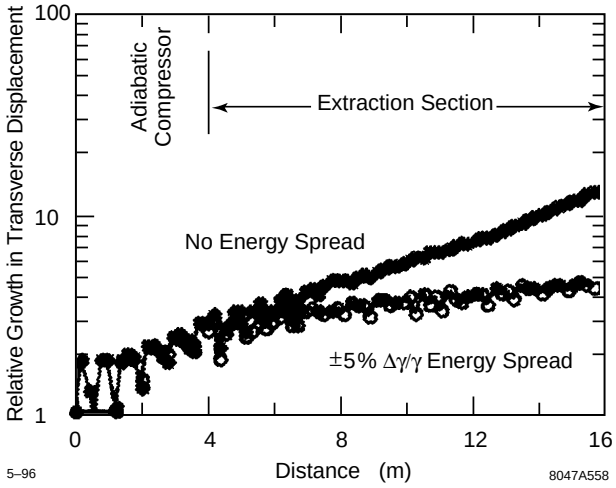


Figure A-26. *Simulations of growth in the transverse displacement of the beam centroid from deflection by HOMs in the induction cell.*

Both traveling wave (TW) and standing wave (SW) structures are being considered for the output structure design of the RTA. The TBNLC RK design used TW structures to reduce the surface fields associated with generating 360 MW per output. RTA is designed to generate 180 MW per output. Thus, inductively detuned SW cavities are a practical alternative. Furthermore, the RKS code is believed to incorrectly describe the radial beam dynamics for inductively detuned TW structures. [Giordano 1995] We have used SW cavities in our modeling to perform complete 2-D simulations and to validate the inductive detuning concept. The detuning concept for the SW output cavities is similar to that used in the adiabatic compressor cavities. The required detuning is also affected by the finite external Q of the cavities. Different cavities can have different Q values to optimize the output power and the bunching. The detuning mechanism is required for stable rf power production through the main RTA. See the simulation results in Figure A-25 Resonant frequency of the cavities was 11.566 GHz for this simulation.

A.5.6 Beam Dynamics Issues

Transverse instability of the beam due to the excitation of higher order modes (HOM) in the rf couplers and induction cell gaps is a serious issue for a long relativistic klystron. We do not anticipate beam loss due to the HOM excitation, but the effect should be measurable. The HOMs associated with three different components will predominate in the prototype. Fortunately, the frequency ranges do not overlap so that the effects can be studied separately.

The first is the approximately 300-MHz transverse mode in the accelerator induction cells. ETA II has operated with 3 kA of current through 60 induction cells of the same geometry, but with ferrite cores. The 20 accelerator cells in the RTA should not pose an instability problem with 1.2 kA of current, but it should be possible to detect weak beam oscillations at 300 MHz by using rf probes.

Of greater concern are the 60–80 prototype induction cells of the adiabatic compressor and extraction section. The transverse impedance of a cell is roughly proportional to the cell gap divided by the square of the beam-pipe radius. The transverse impedance of the smaller prototype cells is about four times that of the larger accelerator cells. Simulations of the growth in the transverse instability were performed using the OMICE code. Results are shown in Figure A-26. For the purpose of the simulation, the prototype cells were assumed to have the same characteristics as the TBNLC induction cells ($Z_{\perp 1} = 5,400\,\Omega/m, Q = 2$, resonant frequency = 3 GHz), the current was increased from 0 to 600 A in 50 ns, total pulse duration was 200 ns, and a step function offset was used as the excitation seed. One curve represents growth for a monoenergetic beam while the second has a $\pm 5\%$ energy spread over the rf bucket, imposed to illustrate the effect of Landau damping. The expected exponential growth is evident in Figure A-26, but well within the design goal of less than a factor of 100.

The third and most critical instability affecting transverse motion is caused by the rf output structures in the extraction section. Despite heavy damping of the structures and the effect of Landau damping, it was necessary to use the Betatron Node Scheme to suppress the transverse instability to a tolerable level for the TBNLC. Simulation results of the growth in the transverse instability through the extraction section are shown in Figure A-27. The rf characteristics of the three-cell traveling-wave output structures described in the TBNLC design study were used for the simulation ($Z_\perp/Q = 3\Omega$ per cell, dipole resonant frequency of 14 GHz, $Q = 10$ for the first and third cells and 3,500 for the second, and dipole phase advance of about $2\pi/3$). A monoenergetic beam with parameters the same as for the induction cell simulation was used for this simulation. A relative small change in the beam energy from that required for the Betatron Node Scheme can lead to substantial increase in the growth of the instability as indicated in Figure A-27. While the growth remains tolerable for both cases shown, the idler cavities in the adiabatic compressor require different transverse rf characteristics than the main RTA output cavities to avoid beam loss and/or adverse emittance growth. The difference in HOM power generated in the cavities is three orders of magnitude greater after 12 cavities when not operating at the energy for the Betatron Node Scheme. A measurement of the HOM component in the output power is expected to be sensitive to the effectiveness of the scheme.

Landau damping and the Betatron Node Scheme were both required for transverse stability in the TBNLC design. Neither are required for the operation of the RTA, but the effectiveness of both for a variety of operating parameters can be measured. The effect of the Betatron Node Scheme can be readily ascertained by measuring the HOM power generated in the output structure while rf loops distributed along the beam line can determine the effectiveness of Landau damping.

Longitudinal stability issues include both the rf bucket and the phase relationship between rf buckets. The rf bucket must remain appropriately bunched for stable rf current and power extraction as described in Section A.3.1. Simulation results of the rf power generated and bunching in the extraction section when the resonant frequency of all the detuned

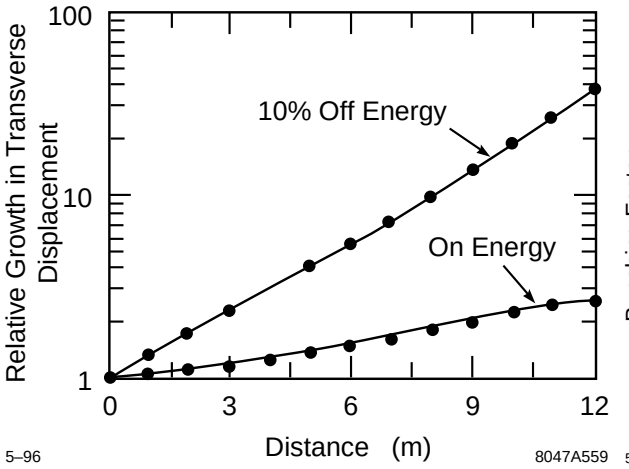


Figure A-27. *Simulations showing growth in the transverse displacement of the beam centroid from deflections by HOMs in the rf structures.*

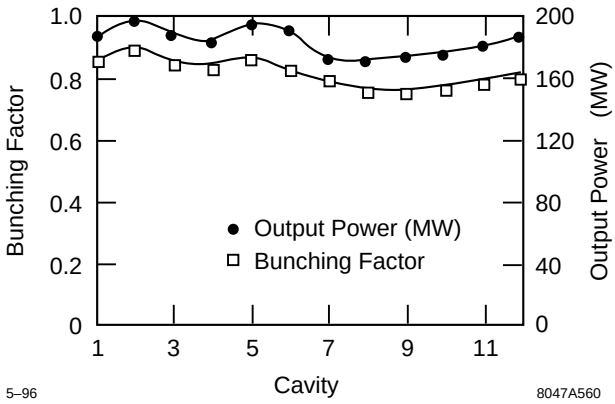


Figure A-28. *Evolution of bunching factor and power output in the main RTA with detuned rf output structures.*

output structures has been set to 11.566 GHz is shown in Figure A-28. The bunching factor is defined as:

$$b = \left| Re \left\{ \frac{1}{N} \sum_{n=1}^{N} e_n^{j\psi} \right\} \right| \quad , \tag{A.11}$$

where the sum is over all the electrons and ψ is the phase of the electron with respect to the center of the rf bucket. The relatively constant value of the bunching factor is a good indication of long term stability of the rf bucket.

The length of RTA is adequate for a meaningful measurement of the longitudinal beam dynamics involved in the detuning of the output structures to maintain the rf bucket. The synchrotron wavelength can be expressed as:

$$\lambda_s = 2\pi \left(\frac{\omega}{c\gamma^3} \frac{d\gamma}{dz} \right)^{-\frac{1}{2}} \quad . \tag{A.12}$$

At 4 MeV, accelerating gradient of 300 kV/ m, and 11.4 GHz, λ_s is ≈14 m. Numerical sensitivity studies indicate that rf output power is insensitive to energy variation and shows small variations ($< 4\%$) for current variations of $\pm 1\%$. The variations were imposed for the flattop portion of the beam pulse. Phase stability is not appreciably effected by current variations of less than $\pm 1\%$. However, phase sensitivity leads to a severe requirement on the average (head-to-tail) energy variation for the flat-top portion of the beam pulse. Phase variations are well modeled by the following first-order formula:

$$\frac{\Delta\psi}{\Delta z} = \frac{k}{\overline{\gamma}^3} \Delta\gamma \quad , \tag{A.13}$$

where $\Delta\psi$ is phase variation, $\Delta\gamma$ is head-to-tail energy variation over the pulse length, k is the free space wave number, and $\overline{\gamma}$ is the average beam energy. With the RTA parameters, assuming that field phase variation should not exceed $\pm 5^\circ$, the required pulse energy flatness (flattop) is estimated to be $\pm 0.3\%$ for an 8-m extraction section.

Beam emittance is an important parameter for the RTA. After the chopper, the focusing system is comprised of permanent quadrupole magnets. The ppm quadrupole focusing is important in the TBNLC design for cost and efficiency reasons. For the RTA, mechanical design constraints and the experimental goal of studying the Betatron Node Scheme require a pole field at a radius of 2 cm for the quadrupoles of about 870 gauss, a half-lattice period of

10 cm, occupancy factor of 0.5, and a phase advance of 72°. The normalized edge emittance must be no larger than 800 π-mm-mr to meet the design goal of an average beam radius (edge) in the extraction section of 4 mm.

Our goal is to limit the emittance growth in the injector from beam optics to a factor of three times the thermal source emittance of about 80 π-mm-mr (0.1 eV) for the cathode. The chopper is expected to increase the emittance by a factor of about 1.7. Thus it is very important to minimize sources of emittance growth such as nonzero magnetic flux at the cathode, magnet misalignments, solenoid to quadrupole matching, and higher order multipoles in the quadrupole magnets. The strong focusing used in the extraction section and the large beam energy spread will convert any transverse motion of the beam, *e.g.,* corkscrew motion or transverse instabilities, into an increase in effective emittance.

A.6 Conclusions

A preliminary design of an RK-TBA based rf power source for the NLC has been presented. The TBNLC system study focused on three major areas: (1) RK-TBA physics, (2) RK-TBA engineering and (3) RK-TBA costing. Here we have described the more important results related to physics and engineering from that study. In addition, we have described an experimental program to construct and test a prototype rf power source based on this design. The description of this program emphasized a number of the RK-TBA engineering issues studied in the TBNLC system study.

The issue of an economical and efficient rf power source is central to the development of future linear colliders. While the induction linac based TBA is known to have the potential for very high efficiency, research support received to date has been somewhat limited. General concerns about the TBA concept as a realizable power source include:

- The induction linac technology is not as mature as existing rf-klystron technology.
- The TBA beam dynamics, particularly BBU, is difficult.
- Induction linacs are perceived to be very expensive.

The objectives of the RK-TBA research program at LBNL and LNLL were motivated by these concerns.

In regards to the cost issue, past cost estimates were based on induction linacs built for other applications (short pulse, high-repetition rate, high current). Our belief has been that the induction linac design ought to be optimized for the specific mission. Hence, the TBNLC system study was specifically directed towards an rf power source suitable for the NLC. Likewise, the experimental program is developing a prototype rf power source that is suitable for powering the NLCTA.

Since cost is, and will be, a key issue for future colliders, we feel that it is essential to have a believable cost estimate. The only way to have a believable cost is to carry the engineering study to sufficient details to make a bottom-up estimate possible. We have tried to estimate cost for fabrication and assembly on the basis of engineering drawings and proposed mass-production procedures. For the induction modules, a major component of the machine, we have sent out for external bids for various parts, to cross-check our own estimates. Our estimate, for approximately ten different items, agreed with external quotes to a few percent on the total. All components are based on known technology, with relatively well-known costs. Hence for the engineering design proposed, we believe that our cost estimate is realistic.

Similarly, the efficiency estimate for our pulse power design is based on well-known numbers with today's technology. We have not made any extrapolations for possible future innovations. We might note here that as we have done more testing in the experimental program, the 40% number quoted in the TBNLC system study for the efficiency of the

pulse power system has been increased to 55% leading to a wall-plug-to-rf efficiency of 50%. As we learn more about the technology, we may introduce new design changes which could alter the cost and efficiency estimates. Yet, on the basis of the work that has been performed thus far, we can safely argue that the cost of upgrading the initial NLC configuration to one or more-TeV-c.m. energy is an attractive option.

In regard to the beam dynamics issues, we have offered conceptual solutions for longitudinal beam stability as well as BBU control for both the low-frequency component, associated with induction gaps, and the high-frequency component, associated with rf extraction cavities. These concepts are supported by detailed simulations. We hope to have demonstrated by these studies that the solutions offered have a reasonable chance of success.

We would like to stress that the reported work is an ongoing effort. Further theoretical and computation studies are being conducted. Specifically, cavity design work continues, in particular in relation to the issue of how to minimize the transverse defocusing on the beam envelope, and the induction cavity design is being optimized to achieve the needed longitudinal as well as transverse impedances. Feedback systems are being studied to relax the constraints on energy flatness and quadrupole field errors. Our experimental program allows for hardware development, verification of theory/simulations, and improvement of our cost estimates. At the scheduled completion of our experiments in 2002, the TBNLC could be seriously considered for a 1.5-TeV NLC upgrade.

References

[Allen 1989] M.A. Allen *et al.,* "High Gradient Electron Accelerator Powered by a Relativistic Klystron", *Phys. Rev. Lett.* **63**, 2472 (1989).

[Allen 1991] S.L. Allen *et al.,* "Measurements of Reduced Corkscrew Motion on the ETA-II Linear Induction Accelerator", *Proc. of the 1991 Part. Acc. Conf.*, 3094–3096 (1991).

[Barnard 1991] J. Barnard *et al.,* "Study of Recirculating Induction Accelerators as Drivers for Heavy Ion Fusion", UCRL–LR–108095 (1991).

[Brigg 1985] R.J. Brigg *et al.,* "Theoretical and Experimental Investigation of the Interaction Impedances and Q Values of the Accelerating Cells in the Advanced Test Accelerator", *Particle Acclerators* **18**, 41 (1985).

[Calame 1991] J.P. Calame and W.G.Lawson, "A Modified Method for Producing Carbon-Loaded Vacuum-Compatible Microwave Absorbers from a Porous Ceramic", *IEEE Trans. Electron. Devices* **38**–6, 1538–1543 (June 1991).

[Chen 1990] Y.-J. Chen, "Corkscrew Modes in Linear Accelerators", *Nucl. Instr. and Methods* **A292**, 455–464 (1990).

[Chen 1992] Y.-J. Chen, "Beam Control in the ETA-II Linear Induction Accelerator", *Proc. of the 1992 LINAC Conf.*, 540–544 (1992).

[Clark 1988] J. Clark *et al.*, "Design and Initial Operation of the ETA-II Induction Accelerator", *Proc. 14th Int'l Linac Conf.*, Williamsburg, VA, 19–23 (1988).

[Chattopadhyay 1990] S. Chattopadhyay, ed. "Impedance Beyond Cutoff", Special edition of *Particle Acclerators* **25**, Nos. 2–4 (1990).

[Deford 1989] J.F. Deford, G.D. Craig, and R.R. McLead, "The AMOS (Azimuthal Mode Simulator) Code", *Proc. of the 1989 Part. Acc. Conf.*, Chicago, IL (1989).

[Deford 1990a] J.F. Deford and G. Kamin, "Application of Linear Magnetic Loss of Ferrite to Induction Cavity Simulation", *Proc. 1990 Linear Accel. Conf.*, Albuquerque, NM, 384–386 (1990).

[Deford 1990b] J.F. Deford *et al.*, "The AMOS Wakefield Code", *Proc. Conf. on Computer Codes and the Linear Acc. Community*, Los Alamos, NM, 265–289 (1990).

[Fessenden 1992] T. Fessenden and C. Fong eds. "Induction Linac Systems Experiments Conceptual Design Report", LBL–PUB–5324 (1992).

[Giordano 1995] G. Giordano *et al.,* "Beam Dynamic Issues in an Extended Relativistic Klystron", *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Griffith 1990] L.V. Griffith and F.J. Deadrick, "Progress in ETA-II Magnetic Field Alignment Using Stretched Wire and Low Energy Electron Beam Techniques", *Proc. of the 1990 Linac Conf.*, 423–425 (1990).

[Haimson 1965] J. Haimson, "Injector and Waveguide Design Parameters for a High Energy Electron-Positron Linear Accelerator", *IEEE Trans. Nucl. Sci.* **NS-12**–3, 499–507 (June 1965).

[Haimson 1970] J. Haimson, "High Duty Factor Electron LINACS", in *Linear Accelerators,* P.M. Lapostolle and A.L. Septier eds., 462–466. (North-Holland, Amsterdam, 1970).

[Haimson 1989] J. Haimson, and B. Mecklenburg, "Design and Construction of a Chopper Driven 11.4-GHz Traveling Wave RF Generator", *Proc. of the 1989 Part. Accel. Conf.,* 243–245 (1989).

[Haimson 1992] J. Haimson, and B. Mecklenburg, "Suppression of beam induced pulse shortening modes in high power RF generator and TW output structures", *Proc. SPIE Symposium on Intense Microwave and Particle Beams III*, Vol. **1629–71**, 209 (1992).

[Halbach priv] K. Halbach, private communication.

[Halbach 1980] K. Halbach, *Nucl. Instr. and Methods* **169**, 1–10 (1980).

[Hogan 1992] W.J. Hogan, R. Bangerter, and G.L. Kulcinski, "Energy from Inertial Fusion", *Physics Today* **42**–9, 42–50 (1992).

[Houck 1992a] T.L. Houck, and G.A. Westenskow, "Status of the Choppertron Experiments", *Proc. 16th Int'l Linac Conf.*, Ottawa, Ontario, Canada, 498–450 (1992).

[Houck 1992b] T.L. Houck, G.A. Westenskow, and S.S. Yu, "BBU Code Development for High-Power Microwave Generators", *Proc. 16th Int'l Linac Conf.*, Ottawa, Ontario, Canada, 495 (1992).

[Houck 1993] T.L. Houck, "Design Study of a Microwave Driver for a Relativistic Klystron Two-Beam Accelerator", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC, 2590 (1993).

[Kapetanakos 1985] C.A. Kapetanakos, and P. Sprangle, "Ultra-high-current electron induction accelerators", *Physics Today* **38**, 58 (1985).

[Kroll 1990] N.M. Kroll, and D.U.L. Yu, "Computer determination of the external Q and resonant frequency of waveguide loaded structures", *Particle Acclerators* **34**, 231–250 (1990).

[Lee 1978] E.P. Lee, "Resistive Hose Instability of a Beam with the Bennett Profile", *Physics of Fluids* **21**, 1327 (1978).

[Li 1994] H. Li, T.L. Houck, S.S. Yu, and N. Goffeney, "Design Consideration of Relativistic Klystron Two-Beam Accelerator for Suppression of Beam-Breakup", *Proc. SPIE Symposium on Intense Microwave Pulses II*, **2154**–10, 91–98 (1994).

[Orzechowski 1986] T.J. Orzechowski *et al.*, "High-Efficiency Extraction of Microwave Radiation from a Tapered-Wiggler Free-Electron Laser", *Phys. Rev. Lett.* **57**, 2172 (1986).

[Panofsky 1968] W.K.H. Panofsky and M. Bander, "Asymptotic Theory of Beam Break-up in Linear Accelerators", *Rev. Sci. Instr.* **39**, 206 (1968).

[Ryne 1990] R.D. Ryne, and S.S. Yu, "Relativistic Klystron simulations using RKTW2D", *Proc. of the 1990 Linear Acc. Conf.*, Albuquerque, New Mexico, 177–179 (1990).

[Sessler 1982] A.M. Sessler, "The free-electron-laser as a power source for a high gradient accelerating structure", Workshop on Laser Acceleration of Particles, NY, *AIP Conference Proceedings* **91**, 154 (1982).

[Sessler 1987] A.M. Sessler and S.S. Yu, "Relativistic-Klystron Two-Beam-Accelerator", *Phys. Rev. Lett.* **58**–23, 2439–2442 (1987).

[Siemann 1993] R.H. Siemann, "Overview of Linear Collider Designs", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC, 532–536 (1993).

[Thompson 1993] K.A. Thompson *et al.*, "Design and simulation of accelerating structures for future linear colliders", SLAC–PUB–6032 (Nov. 1993).

[Westenskow 1994] G.A. Westenskow and T.L. Houck, "Results of the Reacceleration Experiment: Experimental Study of the Relativistic-Klystron Two-Beam-Accelerator Concept", *Proc. of the 10th Int'l Conf. on High Power Particle Beams*, San Diego, CA (1994).

Contributors

- Fred Deadrick
- Guido Giordano
- Enrique Henestroza
- Tim Houck
- Hai Li
- Steve Lidia
- Louis Reginato
- David Vanecek
- Glen Westenskow
- Simon Yu

B

A Second Interaction Region For Gamma-Gamma, Gamma-Electron and Electron-Electron Collisions

Contents

B.1 Introduction

To maximize the accessible high-energy physics, the NLC will have two interaction regions (IRs): one will study e^+e^- collisions and the other may study $\gamma\gamma$, γe^-, and e^-e^- collisions. In this appendix, we describe the final focus and interaction region required for $\gamma\gamma$ and γe^- collisions, henceforth referred to as IR2.

For both $\gamma\gamma$ and γe^- collisions, the required high-energy photons (γ beams) are most effectively produced via Compton backscattering of focused laser beams by the high-energy electron beams. The high-energy photon beams are then brought into collision with opposing electron and photon beams for γe^- collisions and $\gamma\gamma$ collisions, respectively. This region is distinctly different from the e^+e^- final focus and interaction region in that the final focus is optimized to produce rounder beams and that the IR must contain one IP for the luminosity collision and one or two conversion points where the photon beams are generated. With suitable laser and electron beam parameters, a luminosity of γe^- or $\gamma\gamma$ collisions comparable to that of the e^+e^- collisions can be achieved. The polarization of the high-energy photons can be controlled by the polarizations of the laser and the electron beams. With high luminosity and variable polarization, the $\gamma\gamma$ and γe^- collisions at TeV energies will significantly enhance the discovery potential and analytic power of a TeV linear collider complex. A conceptual layout of the NLC including the second IR for $\gamma\gamma$ collisions is shown in Figure B-1.

A review of $\gamma\gamma$ and γe^- colliders can be found in the proceedings of a workshop at Berkeley [Berkeley 1995]. The idea of incorporating γe^- or $\gamma\gamma$ collisions in a future linear collider via Compton backscattering of a laser beam [Arutyunian 1963] has been studied conceptually, especially by scientists from Novosibirsk [Ginzburg 1981, Ginzburg 1983, Ginzburg 1984, Telnov 1990, Telnov 1991, Telnov 1995]. The nonlinear QED experiment E-144 [Heinrich 1991] on the FFTB line at SLAC may be regarded as an essential "proof-of-principle" for future $\gamma\gamma$ or γe^- colliders.

The physics opportunities for γe^- and $\gamma\gamma$ collisions at the NLC are described in Section B.2. Some examples are [Brodksy 1994, Chanowitz 1994, Ginzburg 1994]:

- A $\gamma\gamma$ collider offers a unique opportunity for measuring the two-photon decay width of the Higgs boson, providing a glimpse of the mass scale beyond the TeV range.
- A $\gamma\gamma$ collider is well suited for searching for new charged particles, such as SUSY particles, leptoquarks, excited state of electrons, etc. because photons generally couple more effectively to these particles than do electrons or positrons.
- A $\gamma\gamma$ or γe^- collider serving as a W-factory, producing 10^6–10^7 Ws/year, allowing for a precision study of gauge boson interactions and a search for their possible anomalies.
- A γe^- collider is uniquely suited to studying the photon structure functions, etc..
- e^-e^- collisions (without conversion to γ rays) are interesting by themselves.

In the following, we describe a preliminary design of the IR2 for the NLC, with the goal of obtaining $L_{\gamma\gamma}$ about $10^{33}\,\mathrm{cm}^{-2}s^{-1}$ within a 10% bandwidth or several times $10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for a broad spectrum. In this design, we chose to employ the electron beam parameters for the e^+e^- collision before the final focus system (FFS). However, the FFS for the $\gamma\gamma$ collision is modified so that $\beta_x^* = \beta_y^* < 1\,\mathrm{mm}$. An FFS satisfying the luminosity requirements is worked out with a tolerance requirement similar to that of the e^+e^- FFS. An elaborate optical mirror system in the very constrained region around the vertex detector and quadrupoles brings the laser beam into a tight focus at the conversion point (CP) located 5 mm upstream of the interaction point (IP). The laser required for the Compton conversion must have a TW of peak power and tens of kW of average power. Such a laser can be built by either

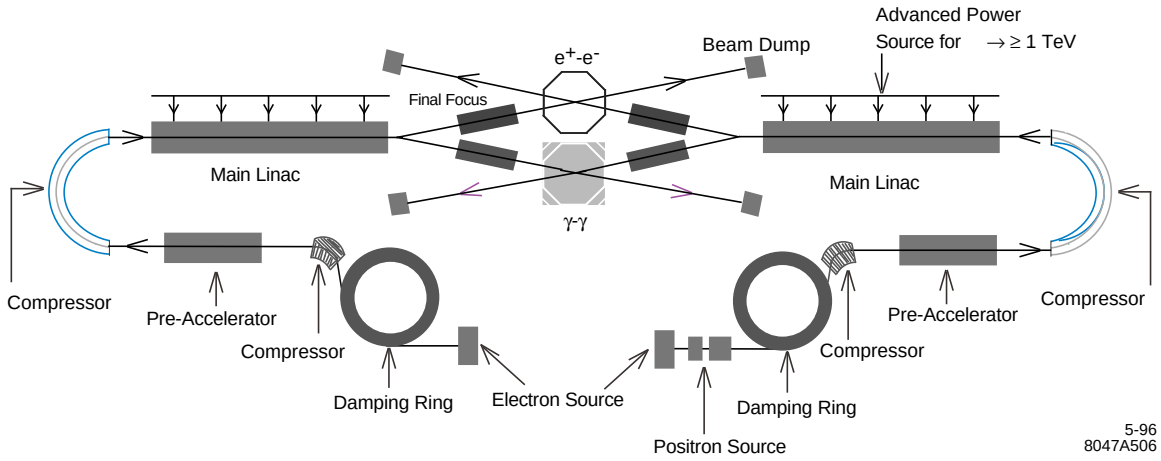


Figure B-1. *Next Linear Collider layout with a second IR for $\gamma\gamma$ collisions.*

combining diode pumping and chirped pulse amplification in solid-state lasers or by a free-electron laser driven by an induction linac and using chirped pulse amplification.

The phenomenon limiting the performance of the $\gamma\gamma$ collisions are different from those in the case of the e^+e^- collisions; the beamstrahlung is absent in the $\gamma\gamma$ collisions, while the e^+e^- pair creation is still important. Therefore the optimization of e-beam parameters for IR2 would be quite different from the case of IR1. Therefore it is worthwhile to revisit the damping rings and linac designs towards obtaining a smaller emittance and/or larger number of particles per bunch at a reduced pulse repetition rate if necessary. These more challenging topics are not pursued in this report.

B.2 Physics Opportunities at $\gamma\gamma$ Collider: The Higgs Sector and Other New Physics

Several review articles have been written on research that could be performed at the IR2 [Brodksy 1994, Chanowitz 1994, Ginzburg 1994]. In this section we focus on the new physics studies, especially concerning the properties of Higgs bosons. Other topics such as the study of the $t\bar{t}$ threshold region, and of the photon structure functions in the $e\gamma$ mode are omitted here.

B.2.1 The Higgs $\gamma\gamma$ Partial Width

One of the most interesting physics programs at a $\gamma\gamma$ collider is the measurement of the Higgs boson partial width into $\gamma\gamma$. This partial width is sensitive to physics beyond the Standard Model because heavy particles whose masses originate in the Higgs mechanism do not decouple in the one-loop diagram [Gunion 1993]. For instance, the fourth generation contributes to the partial width even in the limit where its mass becomes infinity. Therefore, a study of the partial width may indicate new physics.

A plot of $\Gamma(H \to \gamma\gamma)$ with new physics contributions is shown in Figure 2.1 of Ref. [Gunion 1993].

Here we discuss measurements of the partial width $\Gamma(H \to \gamma\gamma)$ for two cases separately, if H decays predominantly into $b\bar{b}$ ($m_H \leq 150\,\text{GeV}$ in the Standard Model) or WW and ZZ (for heavier Higgs in the Standard Model).

$\gamma\gamma \to H \to b\bar{b}$

For this measurement, it is preferred to use the electron helicity to make the photon energy spectrum peaked at its high end, and fix the ee center of mass energy to put the Higgs mass at the peak of the $s_{\gamma\gamma}$ spectrum. The most dominant background is the continuum production $\gamma\gamma \to b\bar{b}$ and $c\bar{c}$. It was pointed out [Barklow 1990] that a $J = 0$ combination of the photon helicities $(++)$ or $(--)$ significantly suppresses the background as m_b^2/s (m_c^2/s) because of the fermion chirality conservation. Furthermore, the continuum production prefers small angles while the signal is isotropic. A cut on the angle, *e.g.*, $|\cos\theta_b| < 0.7$, significantly suppresses the background. For a Standard Model Higgs boson in the intermediate mass range, $80 \leq m_H \leq 150\,\text{GeV}$, the signal cross section is 300–1000 fb level and well above the background after the cuts. Statistical accuracy of $\Gamma(H \to \gamma\gamma)$ is ~5% with an integrated luminosity of 20 fb^{-1}, assuming a b-tagging efficiency of 50% and a $c\bar{c}$-to-$b\bar{b}$ acceptance ratio of 5% [Borden 1993a].

The above analysis has two limitations. One is that the resolved photon contribution to the continuum $b\bar{b}$ production may be important [Eboli 1993]. It can, however, be suppressed by choosing the center-of-mass energy such that m_H lies at the maximum $\gamma\gamma$ energy. The $b\bar{b}$ events produced by resolved photon contribution has typically much lower energy and hence can be suppressed using a visible energy cut. The other limitation is $b\bar{b}g$ or $c\bar{c}g$ final states [Borden 1994, Jikia 1994a]. They appear at higher orders in α_s, but do not have m_f^2/s suppression even in a $J = 0$ helicity combination. Still, suitable kinematical cuts eliminate most of the backgrounds from bbg and ccg [Borden 1994], requiring at least five tracks with a large impact parameter >4 sigma (with $\sigma \simeq 30\,\mu$m) to reject cs. It was also pointed out that the previous studies did not optimize the center-of-mass energy to reduce the backgrounds. By putting m_H on the top of the $s_{\gamma\gamma}$ spectrum, a measurement of $\Gamma(H \to \gamma\gamma)$ is possible with 6% accuracy 20 fb^{-1} [Watanbe 1995].

$\gamma\gamma \to H \to ZZ$

If the Higgs boson is heavier and decays predominantly into vector bosons WW or ZZ, the $b\bar{b}$ mode discussed above is not useful. One cannot use the WW mode either because of its huge tree-level production cross section of ~100 pb. Even the ZZ mode suffers from one-loop production via the W-loop [Jikia 1993], but it is manageable for $m_H \leq 350\,\text{GeV}$. One can measure the partial width $\Gamma(H \to \gamma\gamma)$ at 10% level for small m_H, but the signal is lost for $m_H < 350\,\text{GeV}$ [Borden 1993b].

B.2.2 Higgs CP Eigenvalue

A measurement of the Higgs boson property special at a $\gamma\gamma$ collider is to decide definitively whether a particular Higgs boson is CP even or odd [Grzadkowski 1992, Kramer 1994]. The basic idea is that a CP-even Higgs boson, H^0, couples to the photon with $\mathcal{L} \sim H^0(\vec{E}\cdot\vec{E} - \vec{B}\cdot\vec{B})$, while a CP-odd one, A^0, couples with $\mathcal{L} \sim A^0\vec{E}\cdot\vec{B}$, where $\vec{E}$ is the electric and $\vec{B}$ the magnetic field strength of photon. If the two colliding photon beams are linearly polarized, their polarizations have to be parallel to produce a CP-even state H^0 while they ought to be perpendicular to produce a CP-odd state A^0. Therefore, the asymmetry

$$A \equiv \frac{\sigma(\text{parallel}) - \sigma(\text{antiparallel})}{\sigma(\text{parallel}) + \sigma(\text{antiparallel})} \tag{B.1}$$

is $+1$ for H^0 and -1 for A^0. The studies in Refs. [Grzadkowski 1992, Kramer 1994] showed that an integrated luminosity of 100 fb^{-1} is enough to determine the CP eigenvalue if they decay dominantly into $b\bar{b}$.

B.2.3 Higgs Boson Search

An advantage of a $\gamma\gamma$ collider is that one can use full center-of-mass energy to produce Higgs bosons in s-channel, while one may need to produce them in pairs at the e^+e^- mode. For instance, the heavy CP-even Higgs H^0 and CP-odd Higgs A^0 in the minimal supersymmetric standard model (MSSM) are produced in an association $e^+e^- \rightarrow H^0A^0$, while their production with Z^0 (Z^0H^0 or Z^0A^0 final states) are suppressed if $m_{A^0} \gtrsim 300\,\mathrm{GeV}$. On the other hand, a $\gamma\gamma$ collider can produce A^0 and H^0 states with its full center-of-mass energy, and can be used as a discovery machine even if their threshold lies beyond the e^+e^- center-of-mass energy. Using basically the same strategy in looking for the Standard Model Higgs decaying to $b\bar{b}$, one can cover a substantial region of the parameter space. The final states $t\bar{t}$ or $H^0 \rightarrow h^0h^0$ can be used as well. The $t\bar{t}$ final state suffers from continuum background, and the $t\bar{t}g$ final state has been calculated [Kamal 1995]. Even though more studies are necessary, the detection seems to be feasible.

There is a potential problem with their supersymmetric decay modes. For instance, $A^0 \rightarrow \tilde{\chi}^0_1\tilde{\chi}^0_1$ may be open and dominate the decay branching ratio, which does not leave any visible signature [Gunion 1995].

For a light Standard Model Higgs boson decaying into $b\bar{b}$, the high-energy part of the broad-band photon energy spectrum from the γ-conversion does an excellent job for the discovery. With $\sqrt{s_{ee}} = 500\,\mathrm{GeV}$, 10 fb^{-1} and broad-band spectrum, one can observe Higgs bosons for $m_H = 110$–$140\,\mathrm{GeV}$ [Baillargeon 1995]. Of course, with lower center-of-mass energy (*e.g.*, 350 GeV), discovery reach extends to lower mass (90 GeV). This capability is desired especially when the $\gamma\gamma$ collision operates at the second collision point of an e^+e^- collider and one cannot vary the center-of-mass energy freely.

For heavy Standard Model Higgs bosons decaying predominantly into WW and ZZ, one needs to go to WWH final state, and it requires a large luminosity. For instance, with $\sqrt{s_{ee}} = 1.5\,\mathrm{TeV}$ and 200 fb^{-1}, one can observe up to 700-GeV Higgs bosons [Jikia 1994b, Cheung 1994].

B.2.4 Strongly Interacting Electroweak Sector

The study of the electroweak symmetry breaking sector is difficult if it is strongly interacting as it is for any other colliders, *e.g.*, pp or e^+e^-. The main reasons for the difficulty are that there is no light degrees of freedom in the sector and the only signature is the tail of strong interaction among longitudinal W-boson or top quark above the TeV scale. Even though a $\gamma\gamma$ collider has a huge cross section to produce a W-pairs, they are predominantly transversely polarized and are not sensitive to the strong interactions.

There are discussions to study $WWWW$ or $WWZZ$ final states at $\gamma\gamma$ colliders [Jikia 1994b, Cheung 1994]. However, their study typically requires center-of-mass energy *higher* than their e^+e^- cousin, and probably not of a main target of the first stage $\gamma\gamma$ collider. For instance, it was discussed that at $\sqrt{s_{ee}} = 2\,\mathrm{TeV}$, one needs a luminosity more than 200 fb^{-1} to observe strong interaction among the W_Ls [Jikia 1994b].

Another possible signature of the strongly-interacting electroweak sector is the energy dependence of the $t\bar{t}$ production cross section. Suppose the top quark mass is generated by an effective four-fermion interaction, $\mathcal{L} \sim \frac{1}{\Lambda^2}t\bar{t}Q\bar{Q}$, where Λ is the scale of extended technicolor or its analog, and Q is a techniquark which condenses to break electroweak symmetry. Due to a loop diagram of techniquarks Q, the $t\bar{t}$ production cross section can be significantly reduced

[Asaka 1995]. A possible techni-eta meson may be observed at the $\gamma\gamma$ mode as well [Tandean 1995]. Clearly, more discussions and studies are necessary for the case of the strongly interacting electroweak sector.

B.2.5 Supersymmetry

If supersymmetry exists, charged superparticles can be produced at a $\gamma\gamma$ collider with reasonable cross sections. For many of them, W-pair is the main background. For instance, a pair of sleptons $\tilde{l}^+\tilde{l}^-$ can be produced which decays into $l^+\tilde{\chi}_1^0 l^-\tilde{\chi}_1^0$. While W-pairs can lead to the same signature, one can obtain a relatively clean sample of signals after suitable cuts [Kon 1993]. Mass measurement of sleptons and neutralinos can be done at a 5% level with 20 fb^{-1} [Murayama 1994]. Charginos suffer more from the W-pair background, and more studies are necessary.

The backscattered laser beam allows us to use the $e\gamma$ mode to extend the discovery reach of selectron $\tilde{e}$, and in the process $e\gamma \to \tilde{e}\tilde{\chi}_1^0$. A selectron can be produced even if the e^+e^- center-of-mass energy is below the threshold of its pair production, up to $m_{\tilde{e}} < \sqrt{s_{e\gamma}} - m_{\tilde{\chi}_1^0}$ [Kon 1992a].

B.2.6 Compositeness

If some of the particles in the Standard Model are a composite of more fundamental objects, they exhibit either (1) excited states decaying into the ground state by γ, Z, or g radiation, or (2) anomalous interactions at the low-energy limit of their form factors.

If the electron is a composite, one can look for its excited state e^* in the process $e\gamma \to e^* \to e\gamma$ [Kon 1992b].

If a W-boson is a composite, it may have an anomalous magnetic moment or electric quadrupole moment (assuming CP invariance). The huge W-pair production cross section from $\gamma\gamma$ allows us a precise measurement of such anomalous moments [Yehudai 1991, Gounaris 1995]. Another process $e^-\gamma \to \nu_e W^-$ can be also used [Yehudai 1990, Raidal 1995]. One can obtain constraints complementary to that from an e^+e^- mode.

B.3 Major Parameters

B.3.1 Basic Scheme

The basic scheme of the IR2 for $\gamma\gamma$ collisions is shown in Figure B-2. Two electron beams from their respective final-focus system (FFS) are heading toward the IP. At a location a short distance (5 mm for this design) upstream from the IP, referred to hereafter as the conversion point (CP), a laser beam is focused and Compton-backscattered by the electrons, resulting in a high-energy beam of photons. The photon beam follows the original electron motion (with a small angular spread of order $1/\gamma$) and arrives at the IP in a tight focus. It collides at the IP with an opposing high energy photon beam similarly produced by a second electron beam.

The spent electron beam, following its interaction at the CP, together with the photon beam will cause background γe^- and e^-e^- events as well as producing copious beamstrahlung photons, which will further increase the backgrounds, in interaction with the other electron beam. Extensive detector simulation is required to determine whether these background poses significant problem for a given experiment. One way to reduce the background is to place a bending

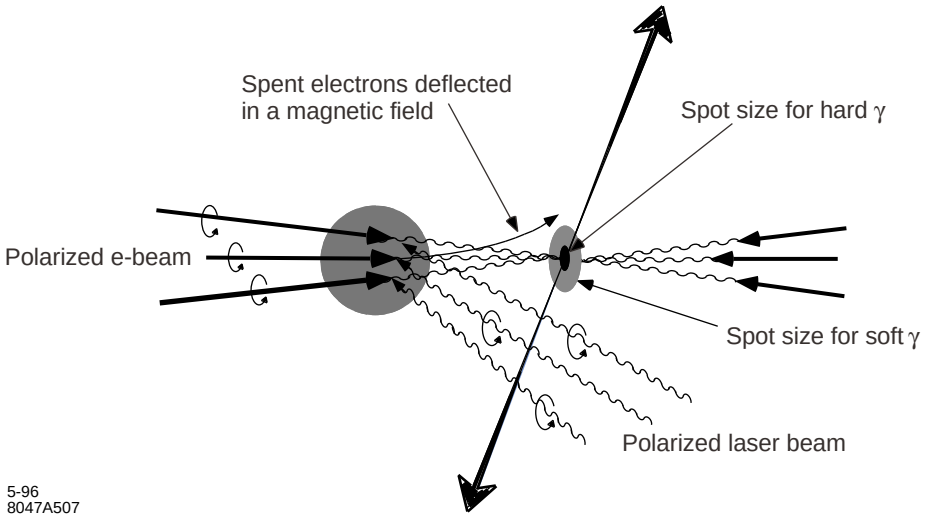


Figure B-2. *General scheme of $\gamma\gamma$ collision.*

magnet [Ginzburg 1983] between the CP and the IP, so that the spent beams miss each other. Implementing the sweeping magnet in the tight space in the IR2 is a challenging problem.

B.3.2 Laser Parameters

The laser beam must be chosen to optimize the generation of the gamma-rays via Compton scattering at the CP. The relevant laser parameters at the CP are summarized in Table B-1. The wavelength and the peak intensity of the required laser are similar to that available in the E-144 experiment at SLAC [Heinrich 1991]. However, the average power is two orders of magnitude larger than what is currently available. In computing the laser spotsize, it is useful to remember the following correspondence between the electron beam and the light beam:

$$\epsilon_x\,,\,\epsilon_y \;\leftrightarrow\; \frac{\lambda}{4\pi} \tag{B.2}$$

$$\beta_x^*\,,\,\beta_y^* \;\leftrightarrow\; \mathrm{Z}_R \tag{B.3}$$

The quantity Z_R is known as the Rayleigh length in optics literature. Thus, the rms spotsize $\sigma_{Lx} = \sigma_{Ly}$ and the cross section Σ at the focus are respectively given by

$$\sigma_{Lx} = \sqrt{\frac{\lambda}{4\pi}\mathrm{Z}_R}\,, \tag{B.4}$$

$$\Sigma \equiv 2\pi\sigma_{Lx}^2 = \frac{1}{2}\lambda\mathrm{Z}_R \tag{B.5}$$

The considerations leading to the parameters in Table B-1 are given in B.4.

B.3.3 Electron Beam Parameters

The electron beam parameters for the reference design at 500-GeV CM energy are summarized in Table B-2.

Wavelength	$\lambda = 1.053\ \mu$m
Micropulse energy	$A = 1$ J
Repetition rate	The same as the electron beam pulse rate (90 micropulses separated by 1.4 ns repeating at 180 Hz)
Rayleigh length	$Z_R = 0.1$ mm
Rms spotsize at waist	$\sigma_{Lx} = \sigma_{Ly} = 2.90\ \mu$m
Rms angular divergence	$\sigma'_{Lx} = \sigma'_{Ly} = 28.9$ mr
Rms micropulse length	$\sigma_{Lz} = 0.23$ mm
Peak intensity	$\approx 1 \times 10^{18}$ W/ cm^2
Peak power	0.5 TW
Average power	16.2 kW
Transverse coherence	Near diffraction limited
Polarization	Fully polarized with helicity switching capability

Table B-1. *Laser parameters for a reference design at $E_{CM} = 500$ GeV.*

The $\gamma\gamma$ or γe^- luminosity is approximately proportional to the e^-e^- geometric luminosity. Although we can in principle reexamine the design of the damping ring–linac complex from the point of maximizing the geometric e^-e^- luminosity, we have chosen for this initial design of the IR2 to use the same electron beam parameters before the IR1 for the e^+e^- collision. However, the beta-functions at the IP are chosen differently from the ones for the e^+e^- collision—with a relaxed β_y^* and a tighter β_x^*. This is due to the fact that the CP is separated from the IP (by 5 mm) to suppress the low-energy part of the $\gamma\gamma$ luminosity spectrum. The separation will introduce an increase in the spotsize of the gamma-ray photons at the IP due to their angular spread $\approx 1/\gamma$ relative to the electron beam. The vertical β_y^* could be larger than that in the case of the e^+e^- collision. It is necessary to reduce β_x^* to compensate the reduction in the luminosity when β_y^* is increased. With $\beta_x^* = \beta_y^* = 1$ mm, the geometric luminosity would be the same as in the case of the e^+e^- collision. The design goal for the FFS for the $\gamma\gamma$ collision is $\beta_x^* = \beta_y^* = 0.5$ mm. Note that, contrary to the e^+e^- case, there are no constraints on the beam profile at the IP arising from beamstrahlung effects.

The large crossing angle, 30 mr, is necessary in the collision scheme without a sweeping magnet, due to the large disruption of the low-energy electrons (coming from the high-order multiple scattering in the CP). The disruption is smaller when a sweeping magnet is employed.

A more detailed considerations leading to the parameters in Table B-2 are given in Section B.5.

B.4 CP Issues

B.4.1 Optimization of the Laser Parameters

Compton scattering of laser beam by relativistic electron beams is an efficient way to generate gamma-ray photons [Arutyunian 1963]. A review of the relevant kinematics can be found in [Telnov 1990].

The energy of the Compton-scattered photon is maximum when the scattered photon is in the direction of the incoming electron, *i.e.,* in the backscattering direction. The maximum energy is given by

$$\omega_{max} = \frac{x}{x+1} E_0, \tag{B.6}$$

Luminosity goal	$\sim 10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for 10 % BW
	$\sim 5\times 10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for broad band
Beam parameters before FFS	The same as e^+e^- design:
Electron energy	250 GeV
Rep. rate	90 bunches separated by 1.4 ns, 180 Hz
Particles per bunch	$N_e = 0.65\times 10^{10}$
Normalized rms emittance	$\gamma\epsilon_x = 5\times 10^{-6}$ mr, $\gamma\epsilon_y = 8\times 10^{-8}$ mr
Beta function at the IP	$\beta_x^* = \beta_y^* = 0.5$ mm
Rms spotsize at the IP	σ_x^*/σ_y^* =71.5/9.04 nm
Rms spotsize at the CP	σ_x^c/σ_y^c =718./90.9 nm
Rms angular divergence	$\sigma_x'/\sigma_y' = 143./18.1\mu$rad
Rms bunch length	$\sigma_z = 0.1$ mm
Polarization	Fully polarized with helicity switching capability
Collision scheme	Vertical offset or sweeping magnet
CP–IP distance	b=5 mm
Crossing angle	$\phi_c \leq 30$ mr

Table B-2. *Electron beam parameters for a reference design at $E_{CM} = 500$ GeV*

where

$$x = \frac{4E_0\hbar\omega_0}{m^2c^4} \simeq 15.3\left[\frac{E_0}{TeV}\right]\left[\frac{\hbar\omega_0}{eV}\right] \quad . \tag{B.7}$$

Here ω_0 is the laser frequency and E_0 is the initial energy of electrons. Hence, the energy of the backscattered photon increases with increasing value of the parameter x, but if x is larger than 4.8, high-energy photons can be lost due to e^+e^- pair creation in collision with unscattered laser photons (Breit-Wheeler process). Thus, the optimum value is $x = 4.8$, corresponding to the maximum photon energy $\omega_{max} = 0.81E_0$. For $E_0 = 0.25$ TeV, this leads to a laser wavelength of about 1 μm. Thus it is convenient to choose the wavelength of Nd:Glass laser, 1.05 μm.

Neglecting multiple scattering, and assuming that the laser profile seen by each electron is the same, the "conversion" probability of generating high-energy gamma photons per individual electron can be written as

$$n_\gamma = 1 - \exp(-q) \quad . \tag{B.8}$$

If the laser intensity along the axis is uniform

$$q = \sigma_c\frac{N_L}{\Sigma} = \frac{\sigma_c A}{\hbar\omega_o\Sigma} = \frac{\sigma_c I\tau_L}{\hbar\omega_0} \quad . \tag{B.9}$$

Here σ_c is the Compton cross section, which for $x = 4.8$ is $1.75\times 10^{-25}/cm^2$, N_L is the number of the laser photons, Σ is the transverse area of the laser spot, A is the the laser pulse energy, I is the laser intensity (power per unit area), and τ_L is the laser pulse length. The pulse energy corresponding to $q = 1$, *i.e.*, a conversion probability of 65%, is given by $A_0 = \hbar\omega_0\Sigma/\sigma_c \approx \hbar\pi c\mathrm{Z}_R/\sigma_c$. The increase in the conversion probability with a laser pulse energy larger than A_0 is relatively small, and, furthermore, is expensive due to higher laser power requirements. Therefore we should choose $A \sim A_0$, *i.e.*, $q \sim 1$. Thus the pulse energy is minimized when the laser spot is focused tightly to match the electron pulse shape. However, the focusing may not be made arbitrarily strong: the laser intensity I could become so large that nonlinear QED effects may spoil the conversion process.

The nonlinear effect is characterized by the quantity

$$\eta = \frac{eE}{\omega_0 mc},\ \eta^2 = 0.4\left[\frac{I}{10^{18}\mathrm{W/cm}^2}\right]\left[\frac{\lambda}{1.054\mu\mathrm{m}}\right]^2 \tag{B.10}$$

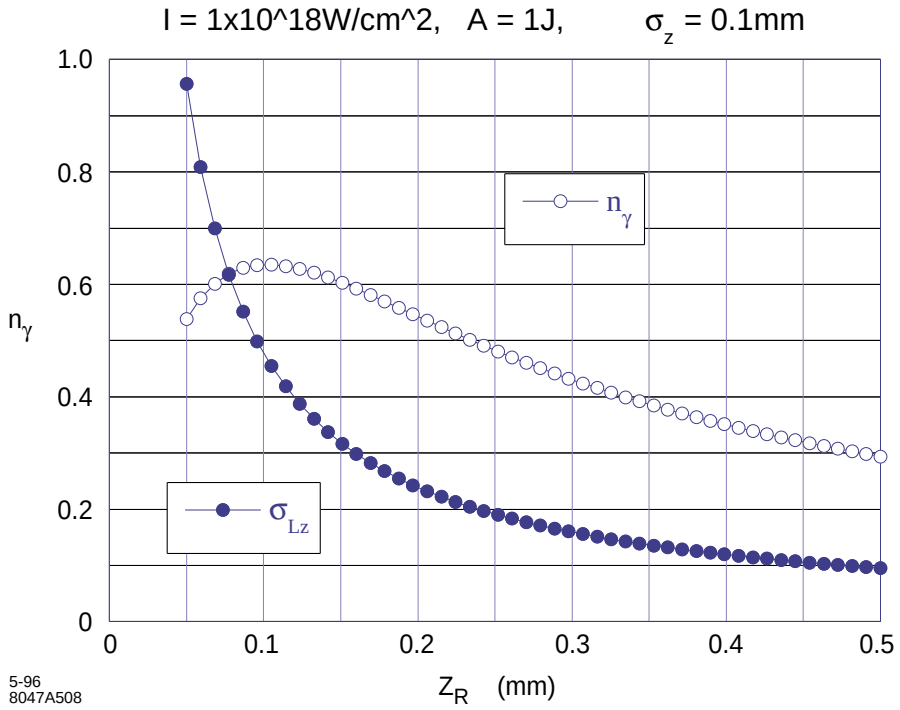


Figure B-3. *Compton Conversion efficiency as a function of z_r at fixed A and I.*

When $\eta^2 \gtrsim 1$, two or more laser photons can be scattered at the same time. The the maximum energy of the backscattered high energy photon in the non-linear Compton scattering involving n laser photons is given by

$$\omega_{max}^{n} = \frac{nx}{1+\eta^2+nx} E_0 \quad . \tag{B.11}$$

For $n = 1$, which corresponds to the single photon process, the maximum photon energy is smaller than that given by the linear approximation, Eq. B.6. This is not desirable since the $n = 1$ photons are usually the most useful ones. Another effect, which is essentially quantum mechanical, is that the pair production can now proceed via multiphoton scattering with a gamma-ray photon, leading to a depletion of the gamma-ray flux. In this design the laser intensity I is kept below 1×10^{18} W/cm^2.

A formula for the conversion efficiency n_γ, neglecting the nonlinear effect and multiple scattering, but taking into account the fact that different electrons see different laser profiles during the interaction, is derived in Section B.4.3. Given the laser pulse energy A and the intensity I, Eq. B.21 can be used to find the optimum value of Z_R and σ_{Lz} corresponding to the maximum conversion efficiency. The case for $A = 1$ J, $I = 1 \times 10^{18}$ W/cm^2, and the rms electron pulse length $\sigma_z = 0.1$ mm is shown in Figure B-3. It is seen that a maximum conversion efficiency $n_\gamma = 0.68$ can be achieved with $Z_R = 0.1$ mm. The corresponding laser pulse length (rms) is $\sigma_{Lz} = 0.23$ mm(which is larger than σ_z for the electrons).

The peak power corresponding to $A = 1$ J and pulse length $\tau \sim 2\sigma_{Lz}/c \simeq 1.8$ ps is about 0.5 TW. With the NLC pulse format of 90 micropulses repeated at 180 Hz, the required average laser power is 16.2 kW, which is rather high. The power can be reduced significantly if laser pulses can be reused by multipass optics or by storing the laser energy in an optical cavity.

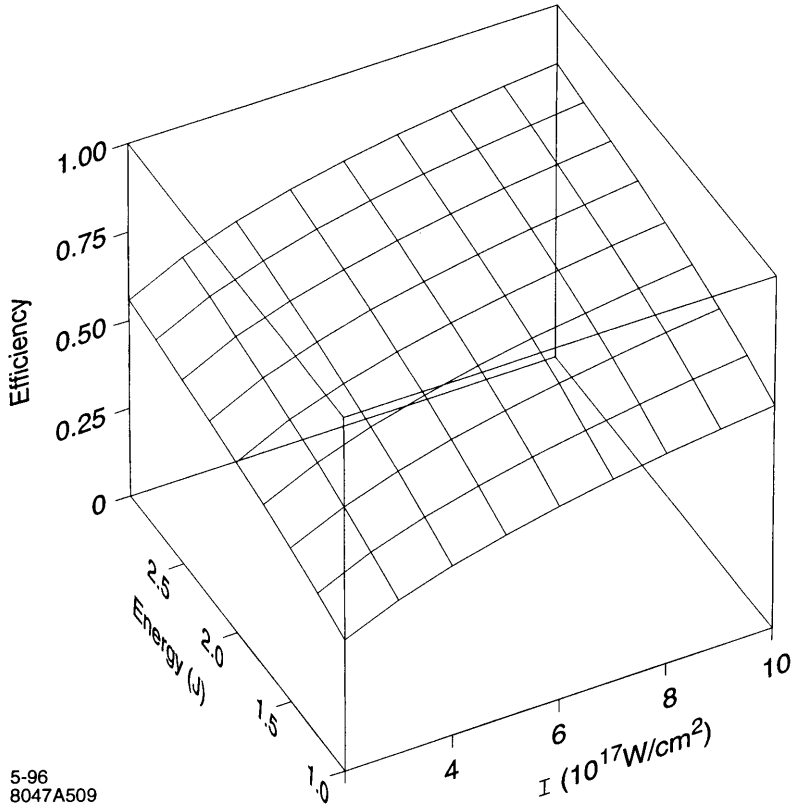


Figure B-4. *Optimized conversion efficiency as a function of pulse energy A and pulse intensity I.*

With the laser intensity $I = 1 \times 10^{18} W/\,\mathrm{cm}^2$, the nonlinear effect is not negligible as we will see in Section B.6. To study the dependence of the conversion efficiency on A and I, we show in Figure B-4 the maximum conversion efficiency as a function of the laser pulse energy and the laser intensity. Figure B-5 and B-6 give respectively the corresponding Rayleigh length and the rms laser pulse length. Figure B-7 gives the contour line corresponding to the conversion efficiency $\eta_\gamma = 0.65$, showing that as the intensity is reduced the pulse energy must be increased to maintain the same conversion efficiency. As an example, with $A = 2$J and $I = 5 \times 10^{17} W/\,\mathrm{cm}^2$, one can obtain the maximum conversion efficiency $n_\gamma = 0.648$ with $Z_R = 0.201\,\mathrm{mm}$ and $\sigma_{LZ} = 0.46\,\mathrm{mm}$. Another example is $A = 3J$ and $I = 3.3 \times 10^{17} W/\,\mathrm{cm}^2$ for which $n_\gamma = 0.65$, $Z_R = 0.347\,\mathrm{mm}$ and $\sigma_{LZ} = 0.76\,\mathrm{mm}$. The non-linear effects in this case are smaller than the $A = 1J$, $I = 1 \times 10^{18} W/\,\mathrm{cm}^2$ example, but the pulse energy is higher. Since the laser power is expensive, we will adopt in this report $A = 1$J, $I = 1 \times^{18} W/\,\mathrm{cm}^2$ as the reference case.

Transverse coherence of the laser beam is important in obtaining a diffraction-limited focal spot. Versatile polarization control is also important; the helicity of the laser light should be opposite to that of the electron beam to obtain a higher conversion rate and the γ photon spectrum peaked around ω_m. Controlling the γ photon polarization by controlling the laser photon polarization is an important technique for many $\gamma\gamma$ or γe^- experiments [Barklow 1990]. Switching of helicity is proposed to characterize all helicity components of the luminosities [Telnov 1995].

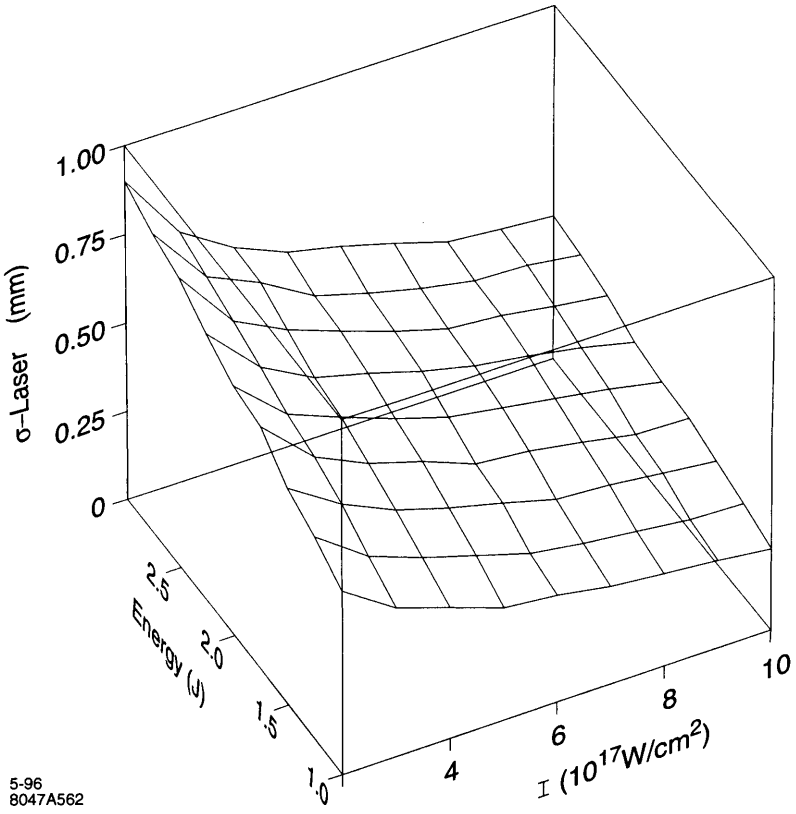


Figure B-5. *Optical pulse length corresponding to the optimized conversion efficiency.*

B.4.2 Low-Energy Electrons Due to High-Order Multiple Scattering

With the intense laser pulse required for an efficient conversion, the probability of multiple scattering is high, giving rise to soft electrons and photons. The multiple scattering process is roughly described by the Poisson distribution. A k-fold scattering has the probability

$$P_k \sim \frac{q^k e^{-q}}{k!} \quad , \tag{B.12}$$

and gives rise to a minimum electron energy E^k where [Telnov 1990]

$$E^k \sim \frac{E_0(1+\eta^2)}{1+\eta^2+kx} \quad . \tag{B.13}$$

For ten-fold multiple scattering, $k = 10$, the electron energy is about 2% of the incoming energy. There are about 1000 such particles, which could cause significant background signals if they are allowed to hit the quadrupole faces.

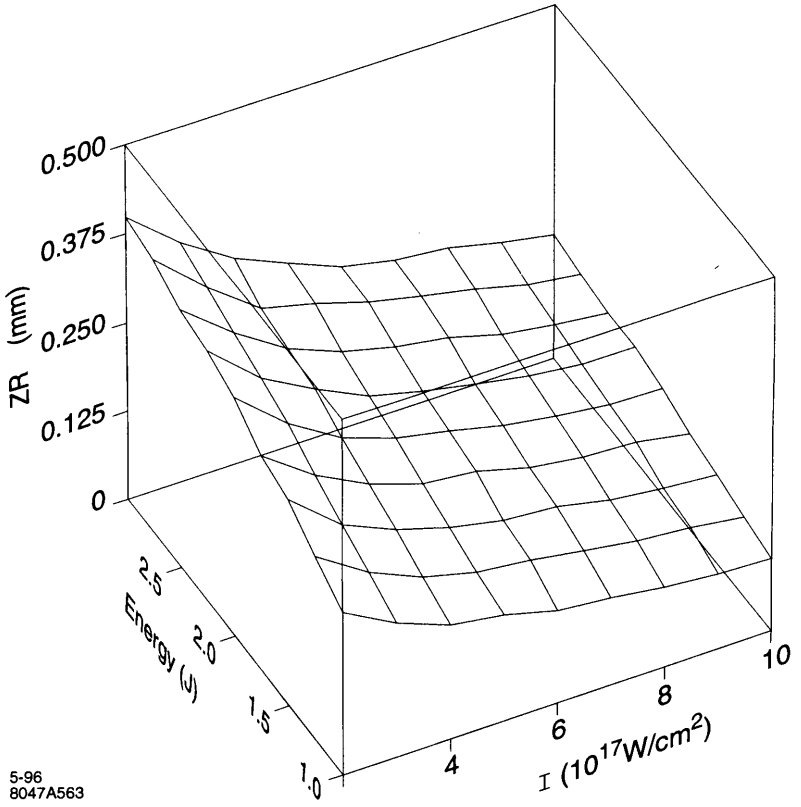


Figure B-6. *Rayleigh length corresponding to the optimized conversion efficiency.*

B.4.3 Compton Conversion Efficiency

Let an electron beam collide with a laser beam of density $n_L(\mathbf{x},t)$. The density of the electron beam will be written as $n_e(\mathbf{x}-\mathbf{v}t,t)$ to indicate that the beam moves with velocity $\mathbf{v}$. The number of scatterings in the space-time element $d\mathbf{x}dt = d\mathbf{x}'dt$, where $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ is given by [Landau 1987]

$$d\nu = \sigma_c v_{rel} n_e(\mathbf{x}',t) n_L(\mathbf{x},t) d\mathbf{x}'dt \quad . \tag{B.14}$$

Here σ_c is the Compton cross section, $v_{rel} = \sqrt{(\mathbf{v}-\mathbf{v}_L)^2 - (\mathbf{v}\times\mathbf{v}_L)^2}$ is the relative velocity, $\mathbf{v}_L = c\mathbf{n}$, and $\mathbf{n}$ is the direction of the laser propagation. Since we are interested in the case where the loss of the laser photons can be neglected, we may assume that the scattering does not change $n_L(\mathbf{x},t)$. On the other hand, the probability of Compton scattering per electron is large, and the electron after producing a γ photon after scattering may be regarded as lost. The rate of the loss is given by Eq. B.14 and can be written as follows:

$$-dn_e(\mathbf{x}',t)d\mathbf{x}' = v_{rel}\sigma_c n_e(\mathbf{x}',t) n_L(\mathbf{x}'+\mathbf{v}t,t) d\mathbf{x}'dt \quad . \tag{B.15}$$

From this, we derive

$$\frac{\partial}{\partial t} n_e(\mathbf{x}',t) = v_{rel}\sigma_c n_e(\mathbf{x}',t) n_L(\mathbf{x}'+\mathbf{v}t,t) \quad . \tag{B.16}$$

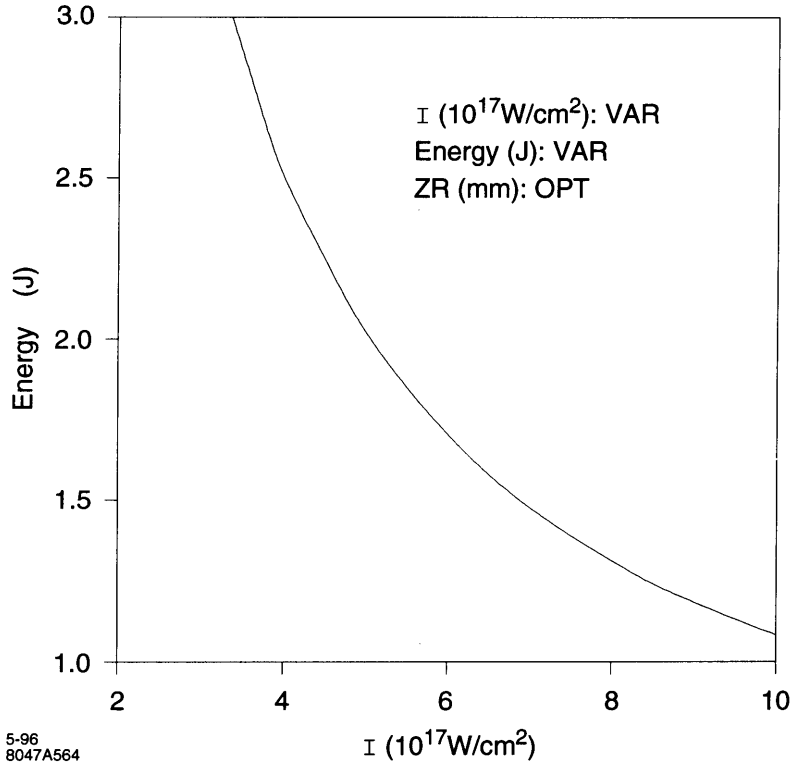


Figure B-7. *Optical pulse energy versus intensity at conversion efficiency = 0.65.*

The solution of Eq. B.16 is

$$n_e(\mathbf{x}', t) = n_e^i(\mathbf{x}') \exp\left[-\int_{-\infty}^{t} v_{rel}\sigma_c n_L(\mathbf{x}' + \mathbf{v}t', t')dt'\right] \quad . \tag{B.17}$$

Here $n_e^i(\mathbf{x}')$ is the initial electron distribution. The total number of the γ photons generated is obtained by integrating the above equation:

$$N_\gamma = \int d\mathbf{x}' n_e^i(\mathbf{x}') \left(1 - \exp\left[-\int_{-\infty}^{\infty} v_{rel}\sigma_c n_L(\mathbf{x}' + \mathbf{v}t', t')dt'\right]\right) \quad . \tag{B.18}$$

The transverse dimension of the electron beam is usually much smaller than that of the laser beam. For an electron beam traveling at an angle θ with respect to the z-axis, we have $x' = x + ct\sin\theta$, $y' = y$, and $z' = z - ct\cos\theta$. The initial electron density can be written as

$$n_e^i(\mathbf{x}') = \frac{N_e}{\sqrt{2\pi}\sigma_z}\delta(x')\delta(y')\exp(-z'^2/2\sigma_z^2) \quad . \tag{B.19}$$

Here N_e is the total number of electrons in the bunch, and σ_z is the rms bunch length. The laser pulse propagating along the negative z-direction can be written as

$$n_L(x, y, z, t) = \frac{N_L}{(2\pi)^{3/2}\sigma_{Lz}} \frac{\exp\left(-\frac{1}{2}\left[\frac{x^2+y^2}{\sigma_{Lx}(z)^2} + \frac{(z+ct)^2}{\sigma_{Lz}^2}\right]\right)}{\sigma_{Lx}(z)^2} \quad . \tag{B.20}$$

Here N_L is the number of the laser photons in the pulse, σ_{Lz} is the rms length, $\sigma_{Lx}(z)^2 = \sigma_{Lx}^{*2}(1+(z/Z_R)^2)$, σ_{Lx}^* is the rms spotsize, and Z_R is the Rayleigh length. The last two quantities are related by $\sigma_{Lx}^{*2} = Z_R\lambda/4\pi$, where λ is the laser wavelength.

In the special case of head-on collision, $\theta = 0$, and one derives from the above formulas the result

$$n_\gamma \equiv \frac{N_\gamma}{N_e} = 1 - \frac{1}{\sqrt{2\pi}\sigma_z}\int dz \exp\left(-\frac{z^2}{2\sigma_z^2} - U(z)\right) \tag{B.21}$$

where

$$U(z) = \frac{4\sigma_c N_L}{\sqrt{2\pi}\lambda Z_R \sigma_{Lz}}\int ds \frac{\exp\left[-\frac{2(s-z/2)^2}{\sigma_{Lz}^2}\right]}{1+s^2/Z_R^2} \quad . \tag{B.22}$$

The spectrum of the photons is given by

$$\frac{dn_\gamma}{dy} = n_\gamma \frac{1}{\sigma_c}\frac{d\sigma_c}{dy} \tag{B.23}$$

where $y = \hbar\omega/E_o$, and $d\sigma_c/dy$ is the differential scattering cross section of the Compton process, given, for example, by Eq. 3 of Ref. [Telnov 1995].

Although we have neglected the multiple scatterings in the derivation here, Eq. B.23 is expected to be valid near $y \sim y_{max} = x/(1+x)$ even for a "thick" target. On the other hand, the effect of multiple scattering will be large in the soft photon region, $y \sim 0$.

B.5 IP Issues

B.5.1 Optimization of Spectral Luminosity

The total $\gamma\gamma$ luminosity is approximately given by $n_\gamma^2 \simeq 0.4$ times the geometric e^-e^- luminosity (with zero offset). However, the spectral luminosity of the $\gamma\gamma$ collision depends strongly on the distance b between the CP and the IP. This is because the energy of the gamma photon depends sensitively on the scattering angle, being maximum in the original electron direction and decreasing rapidly away from this direction. Introducing the parameter $\rho = b/\gamma\sigma_y^*$, where σ_y^* is the vertical rms electron beam size (assumed to be smaller than the horizontal size) in the absence of the CP, the spectral luminosity is broadly distributed as a function of the c.m. energy of the two-photon system when $\rho \ll 1$. As ρ is increased, the low-energy part of the luminosity spectrum becomes suppressed due to the larger spotsize occupied by low-energy photons. Thus, the luminosity spectrum develops a well defined peak at the high-energy end with a bandwidth of about 20% when $\rho \geq 1$. This region is also characterized by a high degree of polarization. For most applications, one would choose $\rho \simeq 1$ to obtain a narrow spectrum without suffering a large luminosity reduction. In our case, this corresponds to $b \sim 5$ mm. The spectral peak at the high-energy end of the invariant mass distribution accounts for about 20% of the total $\gamma\gamma$ luminosity, or about 10% of the geometrical e^-e^- luminosity.

For e^+e^- collisions, the beam spot at the IP is normally designed to be flat to minimize the beamstrahlung effect. In $\gamma\gamma$ collisions, the vertical beam size, which is determined by the condition $\rho \simeq 1$, is larger than that in the e^+e^- collisions for a reasonable value of the CP-IP distance b. The horizontal spotsize should be reduced in proportion to achieve a comparable luminosity. Thus the FFS for $\gamma\gamma$ collision must provide a value of β_x^* which is smaller and β_y^* which is larger than the corresponding values for the e^+e^- design. With $\beta_x^* = \beta_y^* = 1$ mm, the geometric luminosity would be the same as in the case of the e^+e^- collision. We therefore aim for $\beta_x^* = \beta_y^* = 0.5$ mm for the $\gamma\gamma$ collision. In doing so, a proper account should be made of the Oide effect as well as the constraint that β_x and β_y be larger than

the bunch length. A design of the FFS similar to that of the e^+e^- case, but with final quadrupoles reversed in x and y is presented in Section B.9.

B.5.2 Polarization

By varying the polarization of the electron and the laser beams, the polarization of the high-energy photon beams can be tailored to fit the needs of the individual experiment. Controlling the polarization is also important in sharpening the spectral peak in the $\gamma\gamma$ luminosity. Due to the polarization dependence of the Compton scattering, the spectral peak present in the case of $\rho \geq 1$ is significantly enhanced by choosing the helicity of the laser photons to be of the opposite sign to the helicity of the electrons.

B.5.3 Collision of the Spent Electron Beam

The background due to γe^- and e^-e^- collisions as well as the collision of the beamstrahlung photons is large if the spent electron beams are allowed to collide at the IP. The collisions of these particles would also produce positrons and minijets. These unwanted collisions give rise to the background events. Whether these backgrounds pose a significant problem will depend on the nature of the particular experiment, and can only be evaluated after detailed detector simulation.

Among the background events, the γe^- collisions appear to be the most significant, with a luminosity roughly equal to the $\gamma\gamma$ luminosity. The ee luminosity due to the collision of the "spent" electrons (*i.e.*, after Compton conversion) is suppressed significantly (by a factor of 5) due to the fact that the spectrum of the spent electron is broad, leading to a large disruption at the IP.

A way to avoid the collision of the electron beams would be to sweep the spent electrons away from the IP by an external magnetic field. The magnetic field should extend longitudinally to about 1 cm with a strength of about T. Such a magnet could in principle be designed either with a superconducting [Telnov 1990] or with a pulsed conductor [Silvestrov un]. Installing the sweeping magnet to the tight space in the interaction region with a minimum obstruction to the detector is a major challenge.

A plasma lens to overfocus the spent electron beam has also been proposed [Rajagopalan 1994]. This scheme must inject gases to produce plasma and also remove them from the interaction region. Another proposal is to arrange the electron beams to repel and miss each other entirely in "a heads-up" collision [Balakin 1994]. For this scheme to work, the electron beam intensity needs to be much higher than that contemplated in most linear collider proposals. We have not studied these options in detail in this study.

B.5.4 Disruption of Low Energy Electrons

A characteristic angle for the full-energy primary, disrupted electrons is [Hollebeek 1981]:

$$\theta_d \equiv \frac{2Nr_e}{\gamma\sigma_x} \quad . \tag{B.24}$$

Thus the main fraction of the electrons after the IP will be deflected into an angular cone given by θ_d in which γ is replaced by an average value. However, Eq. B.24 is valid only when the deflection angle is smaller than $4\sigma_x/\sigma_z$

(we assume $\sigma_x \ll \sigma_z$). For very low energy electrons for which $\theta_d \gamma e^- \sigma_x/\sigma_z$, the deflection angle is given by [Telnov 1990]

$$\theta \sim \sqrt{\frac{4\pi N r_e}{\gamma \sigma_z}} \quad . \tag{B.25}$$

Low-energy electrons are generated by high-order multiple scattering at the CP, as discussed in Section B.4.2. For our parameters, a ten-fold or higher multiple scattering generates about one thousand electrons with energy as low as 2–3% of their initial energy . These particles will be deflected up to an angle of about 10 mr due to collision with the opposing electron beam. Since there are about 1000 such particles, which will contribute to the background signals if they are allowed to hit the quadrupole faces, the crossing angle should be larger than 10 mr plus an additional angle to clear the quadrupole faces closest to the IP. In the case of e^+e^- collisions, the quadrupole clearance is taken to be 20 mr. We therefore take the crossing angle for the $\gamma\gamma$ collision to be 30 mr. Crabbing the electron beam is essential for a $\gamma\gamma$ collider. The effect of the solenoidal field on beam collisions with a large crossing angle also needs to be studied.

B.6 Luminosity Calculations

B.6.1 Simulation Code Development

The physical processes occurring in the CP and the IP are complex and diverse, including linear and nonlinear, single and multiple Compton scattering at the CP, beamstrahlung, coherent and incoherent pair production, Bethe-Heitler and Landau Lifshitz processes at the IP. A reliable prediction of the $\gamma\gamma$ and γe^- luminosities and the backgrounds can only be done with a numerical code simulating the entire complex of CP and IP physics. It is desirable that several independent codes are available so that simulation results can be cross-checked. At the same time, simple analytical estimates for the relative importance of these processes are also highly desirable.

Ideally code for a full simulation of $\gamma\gamma$ or γe^- collisions must incorporate the following features [Chen 1995a]:

- The CP physics:
  - Linear and nonlinear Compton scattering: $e^\pm + laser \to e^\pm + \gamma$.
  - Linear and nonlinear Breit-Wheeler scattering: $\gamma + laser \to e^+e^-$.
- Beam propagation from the CP to IP including the effects of external magnetic fields (solenoidal field, sweeping magnet) and plasma lens (if any).
- The IP Physics:
  - Disruption effects; interaction of $e^\pm$ with the field of the opposing beam.
  - Beamstrahlung and coherent pair production via interaction of γ and $e^\pm$ with the collective field the opposing beam.
  - Incoherent processes (Bremsstrahlung: $ee \to ee\gamma$; Breit-Wheeler: $\gamma\gamma \to e^+e^-$; Bethe-Heitler: $e\gamma \to ee^+e^-$; Landau-Lifshitz: $ee \to eee^+e^-$).
- Beam propagation from the IP to the exit line.

The ABEL code (Analysis of Beam-beam Effects in Linear colliders) simulates the beam-beam interaction including disruption and beamstrahlung effects [Yokoya, 1986]. The code has been subsequently modified to include incoherent pair creation in the equivalent photon approximation (ABELMOD) [Tauchi 1993]. These codes were originally written for e^+e^- collisions but have been modified to simulate the e^-e^- collisions.

There are several Monte-Carlo codes for simulating the Compton conversion process: A code written by Horton-Smith [Horton-Smith phd] in connection with the E-144 experiment at SLAC, which takes into account the non-linear effect fully but is only applicable for unpolarized electrons; a code written by Ohgaki and Yokoya [Ohgaki 1995], which is based on the Compton scattering in linear approximation but is applicable for arbitrary electron and photon polarizations; a code by Telnov [Telnov 1995], which is similar to the previous one with the further approximation that all electrons see the same laser profile ("same-profile" approximation); a code recently written by Yokoya [Yokoya 1996], which takes into account the non-linearity of the Compton scattering, and can handle circularly polarized electrons.

The codes for the IP and the CP are being combined to a varying degree of sophistication to calculate the $\gamma\gamma$ and γe^- luminosities. In our preliminary calculation, we have used Telnov's code extensively, which includes the multiple Compton scattering effects in linear Compton approximation and the same-profile approximation at the CP, deflection by external magnetic field and synchrotron radiation in the region between the CP and the IP, the beamstrahlung and the coherent pair production at the IP. A similar code has been assembled by Takahashi [Takahashi 1996] based on Ohgaki's Compton conversion package and ABEL. A more refined code incorporating Yokoya's non-linear Compton conversion and the ABEL-MOD is being assembled as a collaborative effort between Hiroshima University, KEK, SLAC and LBL. This code is referred to as CAIN 1.1. Recently, Yokoya has written a new code, named CAIN 2.0, which does not share any subroutines with ABEL [Yokoya 1996].

A simpler code (named BERT) aimed at a careful study of the transport and disruption of the two opposing electron beams with arbitrary initial energy distributions is being developed by W. Fawley. Such a code will be useful, for example, in evaluating the heads-up collision scheme [Balakin 1994] to suppress e^-e^- collisions at the the IP. The results from these codes have been cross-checked where applicable, and are found to be in reasonable agreement with each other.

B.6.2 Simulation of the CP

Figure B-8 gives the γ-photon spectrum after the CP, using the electron beam parameters in Section B.3, and the laser parameters in Section B.4. They are obtained using the code written by Yokoya [Yokoya 1996]. The top, middle and the bottom graphs correspond respectively to the case $A = 3J$, $I = 3.3 \times 10^{17}\,\mathrm{W/cm^2}$, the case $A = 2J$, $I = 5 \times 10^{17}\,\mathrm{W/cm^2}$, and the case $A = 1J$, $I = 10^{18}\,\mathrm{W/cm^2}$, For the top graph, the non-linearity is small, and the spectrum near the maximum photon energy agrees well with the theoretical formula, Eq. B.23. The non-linearity is visible but not pronounced for the middle graph. For the bottom graph however, the smearing of the spectral peak at the high-energy end of the photon is clearly seen.

B.6.3 Telnov's Simulation Results for $\gamma\gamma$, γe^- and e^-e^- Luminosities

This section summarizes the results of simulation calculations using Telnov's code. As discussed in the above, the code is valid under two assumptions: First, the laser profiles seen by all electrons are the same, and second, the non-linear effect can be neglected. The error introduced by the first assumption appears to be not significant in our parameter regime. The validity of the second assumption depends on the laser intensity as discussed in Section B.6.2: The non-linearity is negligible for the case $A = 3\,\mathrm{J}$, $I = 3.3 \times 10^{17} W/cm^2$, and significant for the case $A = 1J, I =$

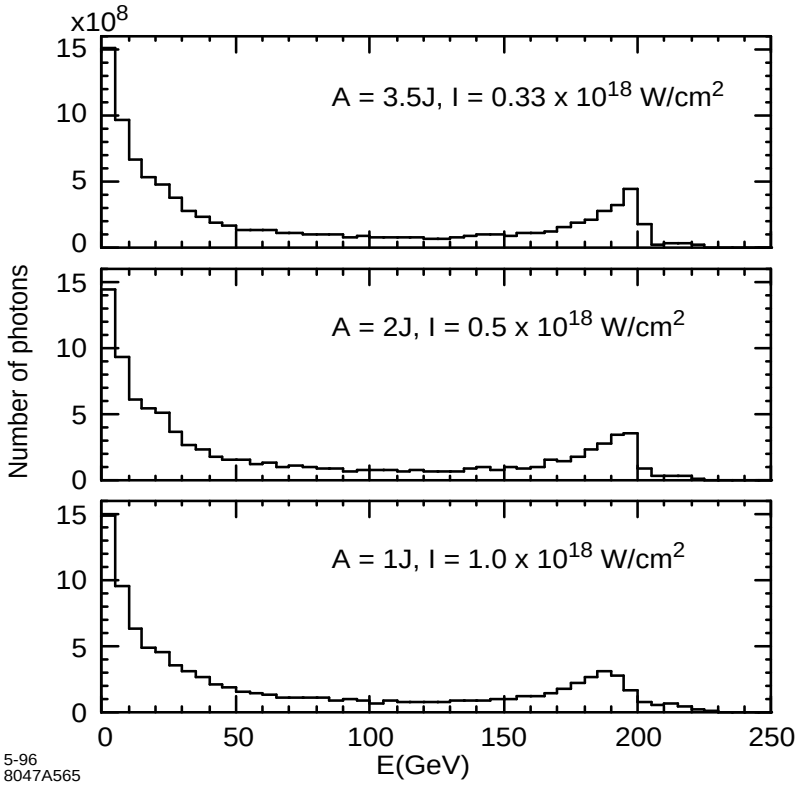


Figure B-8. *Spectrum of the γ-photons after the CP at different pulse energies and intensities.*

$1 \times 10^{18} W/$ cm^2. Although the second case is adopted as the baseline design in this report, we use Telnov's simulation code in this section because it is the code currently fully debugged. The result of Telnov's simulation appears to be in general agreement with that calculated by Takahashi using the linear version of CAIN-1.

Table B-3 summarizes the main results. Here $z =$ invariant mass of the colliding system$/2E_0$, where E_0 is the energy of the incoming electrons. The case (a) to (f) are for the collisions at various vertical offset Δy without the sweeping magnet. The case (d) is the same as (c) but suppressing the beamstrahlung at the IP. In the case (g), there is a 1-T sweeping magnet. The electron and the laser parameters are those given in Section B.3 and B.4, respectively. However, the distance between the CP and the IP is taken to be 7.8 mm for case (g), while it is 5 mm for all other cases.

Even without a sweeping magnet, the e^-e^- luminosity is significantly reduced, a factor of five already at a small offset, $\Delta y = 0.2\sigma_y$, and the reduction increases slowly as a function of the offset. The $\gamma\gamma$ luminosity and γe^- luminosity are roughly equal to the geometric luminosity up to $\Delta y \approx 1\sigma_y$. The $\gamma\gamma$ luminosity at high energy end, $z > 0.65$, is about 10% of the geometric luminosity. A significant fraction of the total $\gamma\gamma$ luminosity is therefore in the low-energy region, and arises from the collisions of the beamstrahlung photons generated at the IP by the interaction of the spent electron beams. The low-energy $\gamma\gamma$ luminosity, as well as the γe^- luminosity, would pose a significant background problem in the collision scheme without the sweeping magnet.

From the table, it is apparent that the luminosity distributions are not a very sensitive function of the offset Δ_y. The $\gamma\gamma$ luminosity at high energy end ($z \geq 0.65$) is practically constant. The ee luminosity becomes smaller by a factor of 2 from $\Delta y = 0.25\sigma_y$ to $\Delta y = 1\sigma_y$. However the ee luminosity is already smaller than the γe^- luminosity by about

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
$L_{ee}/L_{\rm geom}$	0.2	0.19	0.15	0.16	0.12	0.096	0
$L_{ee}(z>0.65)/L_{\rm geom}$	0.12	0.114	0.086	0.091	0.064	0.046	0
$L_{\gamma e\text{-}}/L_{\rm geom}$	1.12	1.04	0.93	0.52	0.79	0.706	0.1
$L_{\gamma e\text{-}}(z>0.65)/L_{\rm geom}$	0.26	0.24	0.2	0.22	0.18	0.143	0.017
$L_{\gamma\gamma}/L_{\rm geom}$	1.23	1.22	1.16	0.38	1.08	1.05	0.37
$L_{\gamma\gamma}(z>0.65)/L_{\rm geom}$	0.116	0.112	0.105	0.104	0.103	0.098	0.09
$L_{\gamma\gamma}(z>0.75)/L_{\rm geom}$	0.057	0.0545	0.0514	0.051	0.05	0.046	0.051
$\theta_{y,\max}$ (mr)	8	8	8	8	8	8	2.5
$E_{\min}$	3	3	3	3	3	3	3

[a] No magnet deflection, $\Delta_y = 0.25\sigma_y$
[b] " " " , $\Delta_y = 0.5\sigma_y$
[c] " " " , $\Delta_y = 0.75\sigma_y$
[d] " " " , $\Delta_y = 0.75\sigma_y$ without beamstrahlung
[e] " " " , $\Delta_y = 1\sigma_y$
[f] " " " , $\Delta_y = 1.25\sigma_y$
[g] With magnet deflection, $b = 0.78$ cm, $B = 10$ kGauss

Table B-3. *Luminosities in $\gamma\gamma$ collision (V. Telnov).*

a factor of five. Therefore the tolerance on Δy is rather relaxed; Δy up to about $1\sigma_y$ does not degrade the collision performance.

The column (g) in the table gives the result when a sweeping magnet is employed. Note that the high-energy $\gamma\gamma$ luminosity is about the same as before, but the background from the low-energy $\gamma\gamma$ or γe^- luminosities are significantly reduced.

The table also shows that the disruption angle for the low-energy particle ($E_{\min} = 3$ GeV) is ± 8 mr for collisions without a sweeping magnet. The low-energy particles are generated through multiple scattering at the CP. Analytical estimate shows that a ten-fold multiple scattering will generate of the order of one thousand particles with an energy of 2–3% of the initial electron energy (corresponding to about 5 GeV in the present case). These are deflected to an angular cone of about 10 mr. This implies that the crossing angle should be larger than 20 mr. The exit beam pipe must be designed to accept these particles to avoid the background events.

For the case (g) with a sweeping magnet, the disruption of the 3-GeV particle is much smaller, about 3 mr.

Figures B-9–B-11 give a more detailed picture for the collision with the vertical offset $\Delta y = 0.75\sigma_y$. Figure B-9 gives the luminosity distributions for $\gamma\gamma$, γe^-, and ee collisions. Figure B-10 gives the $\gamma\gamma$ luminosities, where the solid curve is the same as in Figure B-9 and the dashed curve is for the case where the beamstrahlung contribution is suppressed. Figure B-11 gives the energy distribution of the final electrons.

Figure B-12 shows that the $\gamma\gamma$ luminosities for different offsets between $0.25\sigma_y \le \Delta y \le 1.25\sigma_y$ are more or less the same with each other. Figure B-13 shows the distribution in the vertical angle of the disrupted electron beams at various separations. The shape of the angular disruption may be used to calibrate the offset distance experimentally.

Finally, Figure B-14 shows the luminosity distributions for the case where a sweeping magnet is used (the case (g) in Table B-3). The γe^- luminosity is significantly reduced, and the ee luminosity is entirely suppressed.

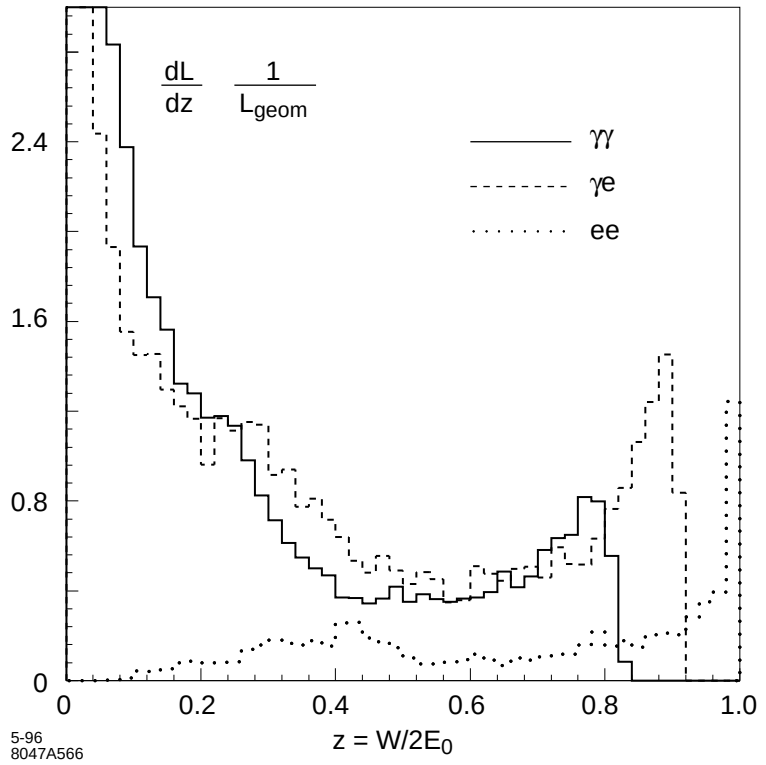


Figure B-9. *Luminosity distributions with $\Delta y = 0.75\sigma_y$ and no sweeping magnet.*

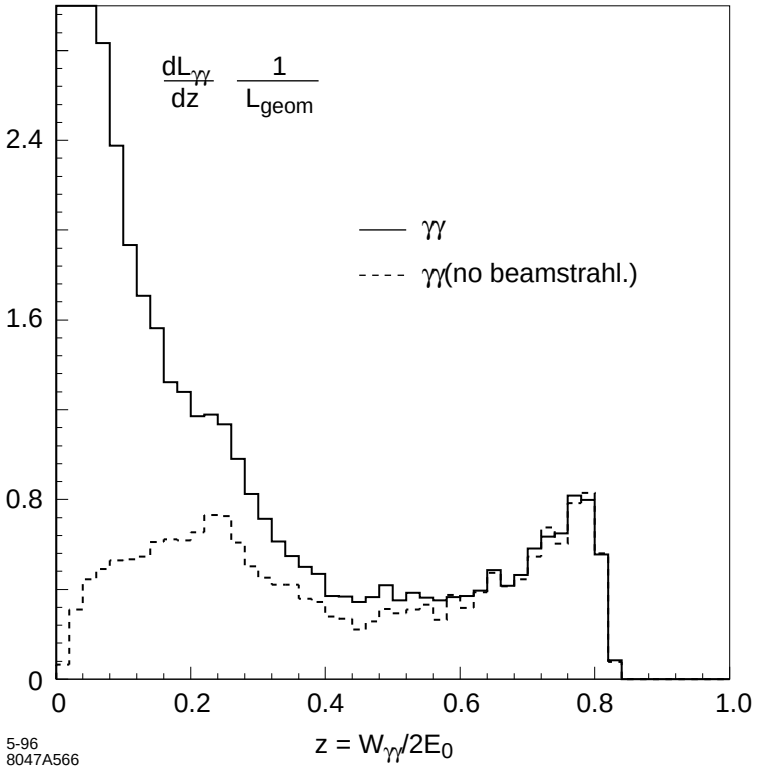


Figure B-10. *$\gamma\gamma$ luminosity for $\Delta y = 0.75\sigma_y$ and no sweeping magnet.*

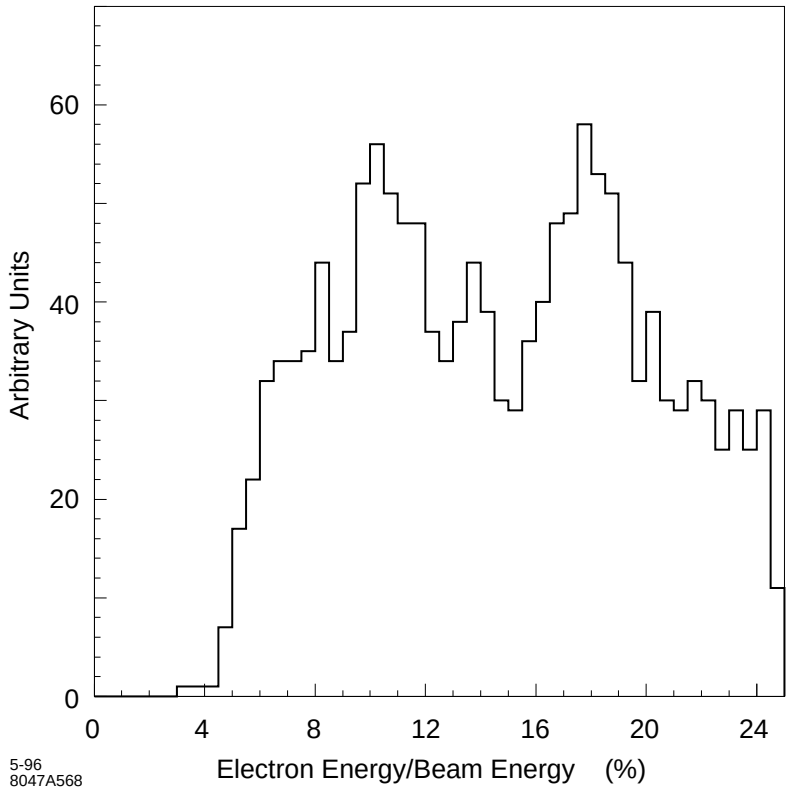


Figure B-11. *Energy distribution of final electrons for $\Delta y = 0.75\sigma_y$, no sweeping magnet.*

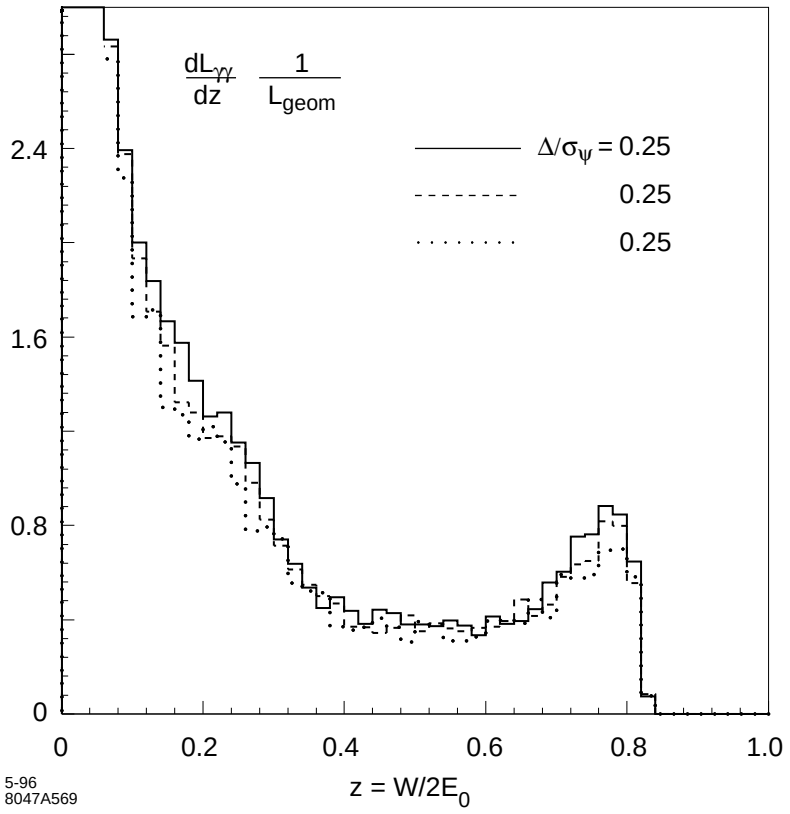


Figure B-12. *$\gamma\gamma$ luminosity distribution for various vertical separations, and no sweeping magnet.*

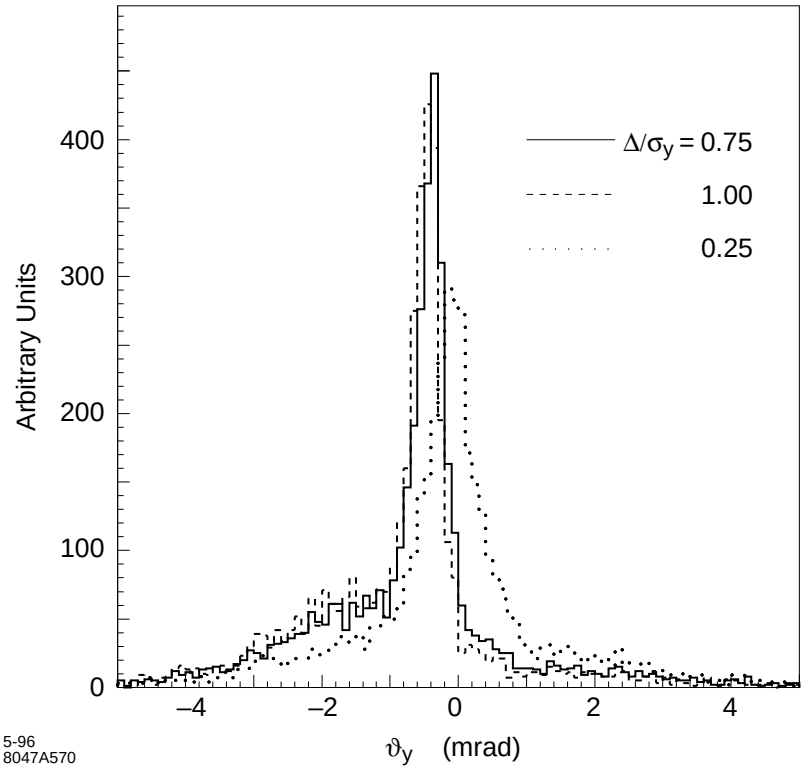


Figure B-13. *Distributions as a function of vertical angle for various vertical offset, and no sweeping magnet.*

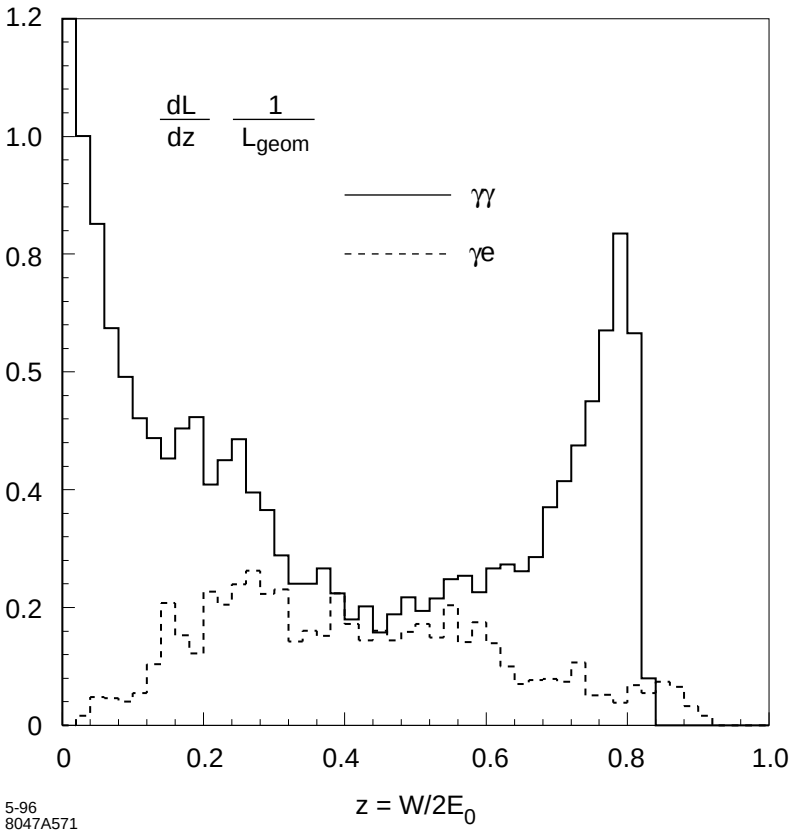


Figure B-14. *Luminosity distributions with magnetic deflection ($b = 0.78$ cm, $B = 1$ T).*

B.7 Backgrounds and Other Detector Considerations

B.7.1 Introduction

For the planned $\gamma\gamma$ physics program in IR2, the incoming polarized electron beams will be converted by an intense laser beam into hard photon beams. As a matter of routine operation in monitoring the incoming electron beam parameters, one or both of the beams will be left unconverted thus allowing high-luminosity studies of γe^- and e^-e^- collisions with the same detector. Given that the electron-photon CP will be located only a few millimeters from the IP, the resulting spent-electron beam background presents a significant challenge to the detector design for these interactions.

It is expected that the physics design criteria for the detector will be basically the same as those for the e^+e^- detector in IR1. Comprehensive, full solid-angle detectors are required to fully exploit the broad physics programs envisioned for both IRs, and the need for precision vertexing and excellent central tracking are identical for similar, specific physics reactions, such as Higgs production and detection in $\gamma\gamma \to H^0 \to b\overline{b}$ and $e^+e^- \to \nu\overline{\nu}H^0 \to \nu\overline{\nu}b\overline{b}$ through W fusion. The main differences expected in the IR2 design arise from the spent beam background and other detailed background differences, from the geometry of the laser optics and the luminosity monitoring systems, and from the unique physics opportunities offered by the different collision processes available.

In this section we detail the expected backgrounds for the $\gamma\gamma$ IR2 region and begin a discussion of detector design considerations. We provide a partial list of the unique physics requirements and a complete list of the machine-related backgrounds that are presently being considered. We review a set of detector geometry and performance parameters that are in use for the current level of simulation. Results of initial debugging runs of a detailed GEANT simulation and background analysis are also presented.

B.7.2 Physics Requirements

The Higgs $\gamma\gamma$ partial width measurement will be the central focus of the $\gamma\gamma$ physics program at a future linear collider. For a light mass Higgs, precision vertexing of b-quark jets will be essential in isolating a $H^0 \to b\overline{b}$ signal from charm and light quark backgrounds. For a larger mass Higgs, excellent electron and muon identification, and good jet energy measurements will be critical in identifying final state W bosons. In this case reasonable hermiticity would allow missing energy determination of neutrinos in the final state.

For detailed measurements of the photon structure function in γe^- collisions, good forward electron acceptance and energy measurements are required for accurate determination of q^2 and x. While in searches for singley produced supersymmetric (SUSY) particles, good electron and muon identification over the largest possible solid angle are important in testing various decay hypotheses.

In e^-e^- collisions, standard e^+e^- detector parameters are required in searches for new particles such as for a Z′. A comprehensive detector would be needed to explore fully any evidence for new physics, such as exotic doubley charged leptons or Higgses.

B.7.3 Backgrounds

Many of the background sources will be common to both the e^+e^- and $\gamma\gamma$ interaction regions. At this preliminary design stage, we rely heavily on the detailed background studies undertaken for the e^+e^- region as discussed in Chapter 12 of the main report.

The backgrounds from upstream sources such as the muon halo and quadrupole synchrotron radiation (QSR) backgrounds will be similar, but differing due to differences in beam parameters and collimation for the e^- beams needed for $\gamma\gamma$ collisions versus the very flat e^+e^- beams. Backgrounds from nearby sources, such as beam-gas scattering, should also be the same.

In the $\gamma\gamma$ interaction region, backgrounds due to the spent electron beams from the e^--γ conversion need particular attention. After the conversion, these charged beams are naturally deflected at the final focus by the long-range Coulomb interaction. This deflection will increase the angular divergence of the beams and also create beamstrahlung photons which will lead to additional backgrounds. As discussed in Section B.6, detailed simulations of the conversion process and interaction physics are being developed. A simple parameterization of this background is discussed below.

High-energy electron-positron pairs will be created due to nonlinear effects in the interaction of laser photons with high-energy photons at the conversion points. At the interaction point, beamstrahlung photons from the disrupted electron beams will interact with themselves, with the Weizsäcker-Williams virtual photons of the opposing e^- beam and with the oncoming high-energy photons to produce additional low-energy pairs. A display and simulation of the effect of these pairs are presented in a following section.

Physics backgrounds arise from the suppressed-luminosity collisions of the spent electron beams with each other and with high-energy photons, and from hadronic backgrounds from low-energy $\gamma\gamma$ interactions. Both backgrounds will add to the inherent backgrounds associated with any particular physics study. However, the backgrounds for $\gamma\gamma$ physics due to e^-e^- and γe^- collisions will provide in themselves additional physics opportunities, while the low-energy $\gamma\gamma$ interactions provide parasitic physics similar to what has been available at lower energy machines. More detailed simulation of the hadronic backgrounds from $\gamma\gamma$ interactions and studies of the background e^-e^- collisions are being planned.

B.7.4 Detector Considerations

For a comprehensive study of $\gamma\gamma$, γe^- and e^-e^- physics, the detector chosen for IR2 will be expected to provide precision vertexing for b-quark separation, and accurate momentum and energy measurements of electrons, muons, and jets up to the beam energy over its full solid-angle coverage. Particle identification of electrons and muons will be accomplished by the calorimeter and muon tracking systems, while a central tracker with dE/dx and other detection techniques would extend the electron-hadron separation over a larger solid angle. Excellent pattern recognition and fine segmentation is required to minimize the sensitivity to machine backgrounds.

Over a large rapidity (η) range, special consideration has to be given to monitoring both the total luminosity and the differential spin-dependent luminosity, $dL_J/d\sqrt{s}/d\eta$, for each of the processes, e^-e^- , γe^- , and $\gamma\gamma$. Detector issues arise from possible interference with the laser optics in the small angle region, and from the lower rates for particular interactions at larger angles which may allow only offline luminosity measurements. Small-angle Møller scattering would provide an excellent monitor of the colliding beams in the e^-e^- mode of operation. Placement of the luminosity detectors behind the laser optics should not limit online monitoring, or precision offline measurements at the few per cent level. The two-photon processes, $\gamma\gamma \to e^+e^-$ and $\gamma\gamma \to \mu^+\mu^-$, allow online monitoring at small angles and

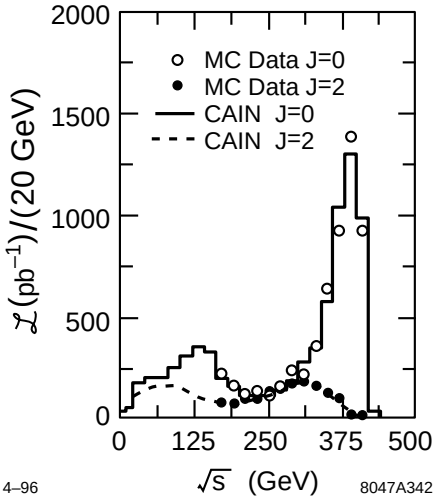


Figure B-15. *Comparison of measured and generated luminosity. The circles and asterisks are measured luminosity for $J = 0$ and $J = 2$ component while solid and dashed lines are generated luminosity for $J = 0$ and $J = 2$ respectively. The normalization corresponds to total integrated luminosity of $10 fb^{-1}$.*

precision measurements at intermediate angles for both the $\gamma\gamma$ and γe^- processes. The process of W pair production in two-photon reactions, $\gamma\gamma \to W^+W^-$, also provides an excellent measure of the $\gamma\gamma$ luminosity.

For a typical $\gamma\gamma$ collider luminosity distribution, the total weighted cross section for W pair production of about 50 pb and a selection efficiency of 15% for 4-jet reconstruction of the Ws yields 75K events per 10 fb^{-1} of integrated luminosity [Takahashi 1995]. For the spin-2 process $\gamma\gamma \to e^+e^-$ the corresponding total luminosity-weighted cross section with $\sqrt{s_{\gamma\gamma}} > 200\,\text{GeV}$ of 10 pb, and a detection efficiency of 70% yields 70 K events per 10 fb^{-1} [Takahashi 1995]. The spin-0 luminosity would be measured by flipping the electron and laser beam polarization simultaneously [Telnov 1993]. Figure B-15 compares the simulated $J = 0$ and $J = 2$ luminosity measurements with the generated luminosity spectra. The expected statistical errors in the W and lepton pair luminosity measurements are shown in Figure B-16.

In detailed GEANT Monte Carlo simulations of a generic detector, we have shared the basic geometry definitions for studies of both the e^+e^- and $\gamma\gamma$ interaction regions. The detector is taken to be 2 m in radius and ± 2.5 m along the beam with a 2-T solenoidal magnet field. At this stage, it consists of simple models of beam line elements and scoring planes only in the vertex and central tracking detector regions. The eventual calorimetry and muon tracking systems are not included, and the resulting detector self-shielding is not taken into account.

For the IR2 design studies, we increase the crossing angle from 20 to 30 mr as presently chosen for the laser optics design, and increase the acceptance of the outgoing quadrupoles from about $\sim$3 mr to $\sim$10 mr to transport the spent electron beam outside of the interaction region.

B.7.5 Initial Simulations

A GEANT drawing of the detector with an expanded vertical scale is shown in Figure B-17. At the center of the drawing are the scoring planes used in the vertex detector region from 2–10 cm in radius and ± 20 cm in length. On either side are the final-focus (FF) quadrupole magnets and beam line masking chosen to minimize backscattering into

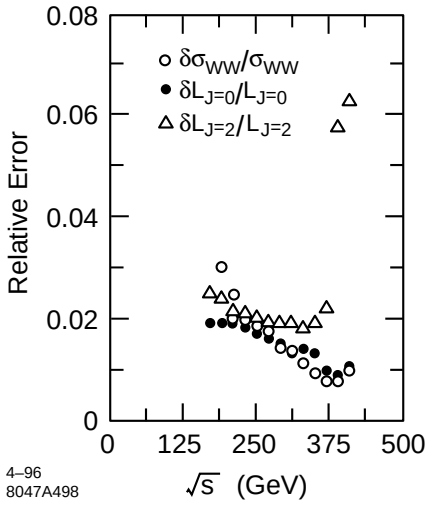


Figure B-16. *The statistical error in W pair production cross-section and luminosity measurement with integrated luminosity of $10fb^{-1}$. The circles are for W measurement while the asterisks and stars are for luminosity measurement for $J = 0$ and $J = 2$ component.*

the detector. For display purposes only, we plot 100 low-energy electron and positron tracks from a simulation of the beam-beam interaction in e^+e^- collisions. (We expect that the electron-positron pairs produced in the conversion and interaction points of the $\gamma\gamma$ region will be similar in energy and angular distributions.) The pairs were generated only in the forward direction of one of the beams. These tracks radiate photons as they spiral along the strong magnetic field lines to the face of the inboard FF quadrupole magnets. Secondary backgrounds are produced from electrons and photons that backscatter into the detector. Figure B-18 displays the hits in the central tracking chamber due to this background source for the $\simeq 10^4$ pair tracks expected to be produced at each bunch crossing. Earlier simulations of a conventional e^+e^- detector with an approximate crossing angle geometry indicated that this background should be tolerable [Ronan 1993].

To obtain an initial estimate of the spent electron beam backgrounds, we generate in our simulation electrons originating from the interaction point with a flat energy spectrum from 15 to 85% of the incoming beam energy. The generated angular spread of these electrons is parameterized by two Gaussian fits to independent detailed simulations of the conversion process. The fitted distributions in both transverse dimensions have a central component and a broader tail with angular spreads of $\sim$1 mr and $\sim$3–4 mr, respectively. On a highly expanded vertical scale, Figure B-19 displays 100 such spent electron tracks exiting through the downstream quadrupole magnets. More detailed simulations of the conversion process and tracking of the spent electrons through the interaction region are in progress. Also, realistic modeling of the fields within the quadrupole magnets are needed in the GEANT simulations to take account of the spent beam as it is transported out of the interaction region.

B.7.6 Conclusions

We are just beginning to understand the relevant design issues and parameters for a detector which could perform the physics anticipated at a future electron linear collider. We welcome volunteer help and encourage international collaboration in the development of simulation tools and eventually in the conceptual design of a detector for $\gamma\gamma$, γe^- and e^-e^- physics.

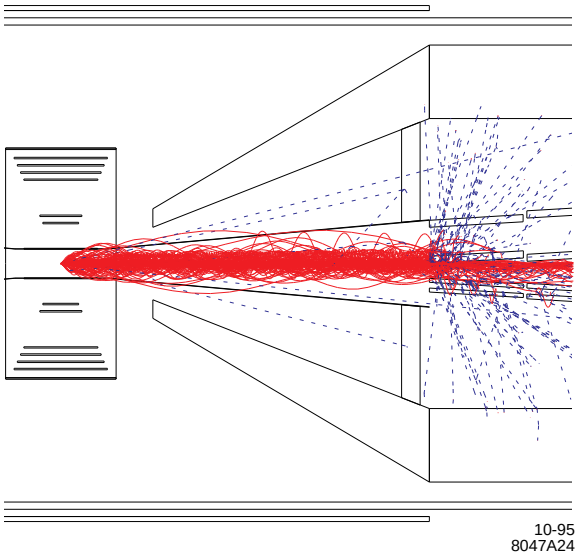


Figure B-17. *Low-energy electron-positron pair background tracks hitting the the face of downstream quadrupole magnets. Only 100 such tracks are shown.*

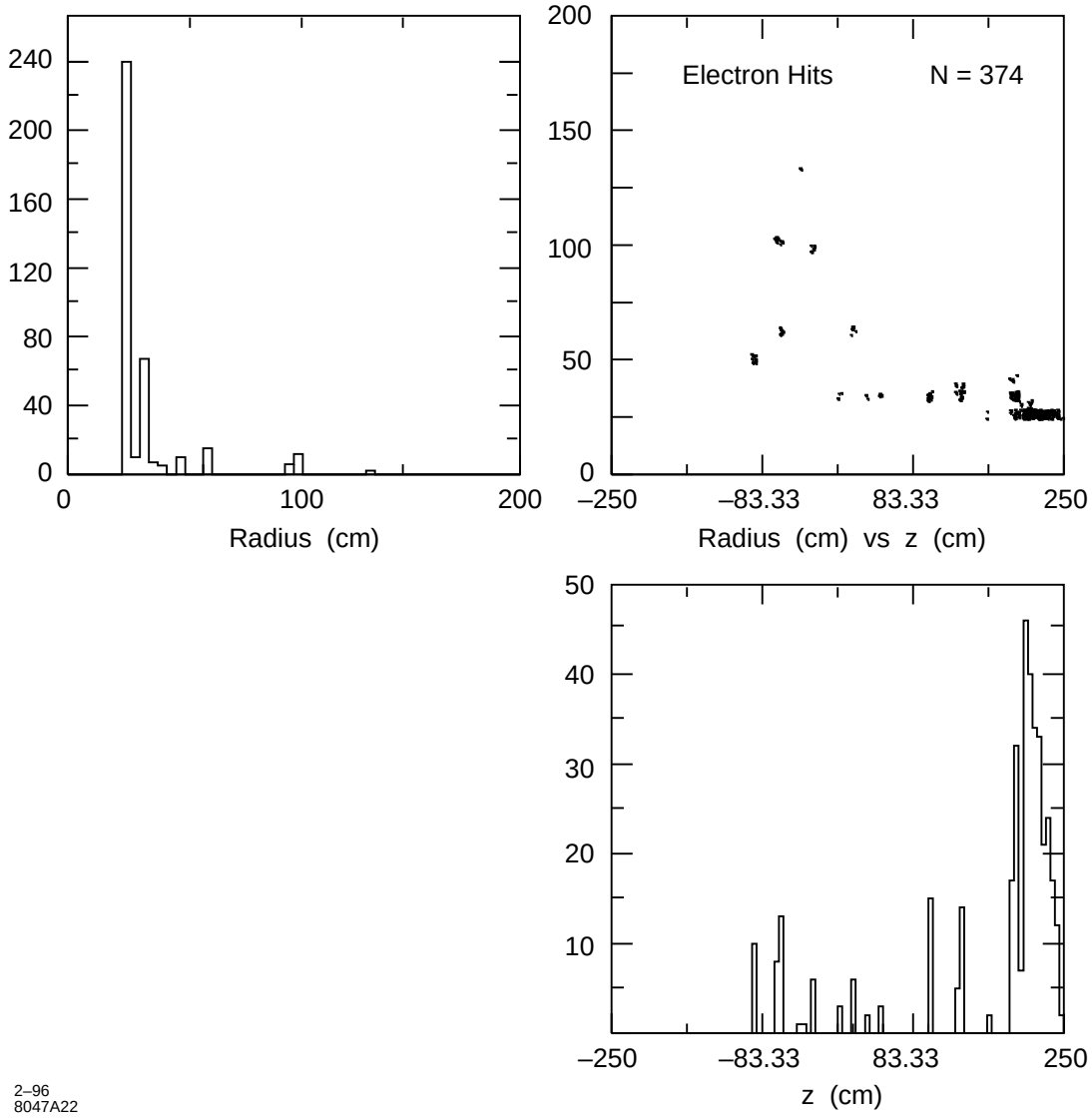


Figure B-18. *Scatter plot (2b), and radial and axial projections (2a,2c) of secondary background hits in central tracker from pair background.*

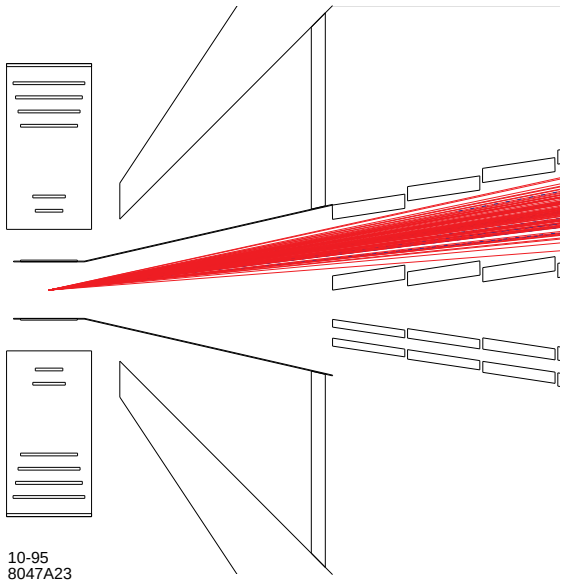


Figure B-19. *Spent electron beam exiting through downstream beam-line magnets. Only 100 rays are shown.*

B.8 Laser Optical Path in IR

B.8.1 Single-Pass Scheme

This section considers some of the practical issues in actually bringing the laser beams into (and out of) an interaction point centered in a HEP detector of the sort that is described in the previous section. A solution is shown in Figure B-20. The case considered here is the simplest one. Each electron beam has an opposing laser beam which is backscattered from it immediately prior to the interaction point. The laser beam, which is essentially unaltered in the process, is then disposed of. The beams are brought in from opposite sides and follow very similar (but not identical) paths along opposite directions. Before discussing the scheme in more detail, it is worthwhile to explain some of the constraints that are encountered in this problem.

- Because of the extremely high intensities involved, transmissive optics are, for the most part, not feasible. This occurs because in the final analysis, the index of refraction of all optical materials is a function of the intensity of the laser light in it. These nonlinearities give rise to a filamentation instability if the total length of material in the system is sufficiently large. The limitation from this so called "B-integral" problem [Siegman 1986] means that such devices as lenses, Pockels cells, and polarizers may not be used after the gratings in the compressor. In fact, the limitations on the design of the optical path produced by this problem are probably such as to just allow a pair of quartz windows for the beam to enter and leave the vacuum system of the accelerator. Beyond these two windows, all optical elements will presumably be reflective; *i.e.,* dielectric mirrors. This limitation also implies that the laser beam will be transported in vacuum.
- Focusing of the beam must be optimized to produce the desired peak intensity (limited by nonlinear effects in the backscattering process as discussed above) and a sufficiently long length of such intensity that most of the electron bunch is converted into photons. This optimization was discussed above. The practical consequence of this is that the "$\mathrm{f}_{\#}$" of the laser focusing will be fixed by this optimization. The $f_{\#}$ is defined roughly as the ratio of the focal length to the diameter of the focusing mirror, as illustrated in Figure B-21, assuming that the mirror is uniformly illuminated. For Gaussian beams, the diameter of the focusing mirror is somewhat arbitrary, but we will take it as the intensity $1/\mathrm{e}^2$ diameter. With this definition, one can show that $f_{\#} = 1/4\sigma_{x'}$, where $\sigma_{x'}$ is the rms angular divergence of the focussed laser beam. In the current design, we have $\sigma_{x'} \approx 27.9\,\mathrm{mr}$ and $f_{\#} = 8.65$.
- Effective use of the laser requires that the optical axis of the laser beam be parallel with the direction of the electron beam to within an angle small compared to the aspect ratio of the laser bunch. The degradation in the general case is given by [Xie 1995] (We do not consider the possibility of "crabbing" the laser beam here.)

$$\frac{L(\theta)}{L(0)} = \frac{1}{\sqrt{1+(\theta/\theta_0)^2}}, \quad \text{where} \quad \theta_0 = \sqrt{\frac{\sigma_{x1}^2+\sigma_{x2}^2}{\sigma_{z1}^2+\sigma_{z2}^2}} \quad , \tag{B.26}$$

  The σ_xs are the transverse sizes of the two crossing beams and the σ_zs are the longitudinal sizes. In our case, the expression for θ_0 is dominated by the size of the laser beam and reduces to $\theta_0 = \sigma_x/\sigma_z$. From the discussion above, one can infer that the aspect ratio of the laser is approximately $4\mathrm{f}_{\#}$. Since the angle between the optical axis and the electron beam is $1/2\mathrm{f}_{\#}$, the minimum degradation if the disrupted beam does not pass through the mirror is $1/\sqrt{5}$. Because of this degradation, a through-the-lens design has been implemented.
- Given both the high peak and average power involved, it is almost certainly necessary to transport the "spent" laser beam out of the IR to an external dump. This also allows external monitoring of the optical quality of the beam after it has gone through focus.

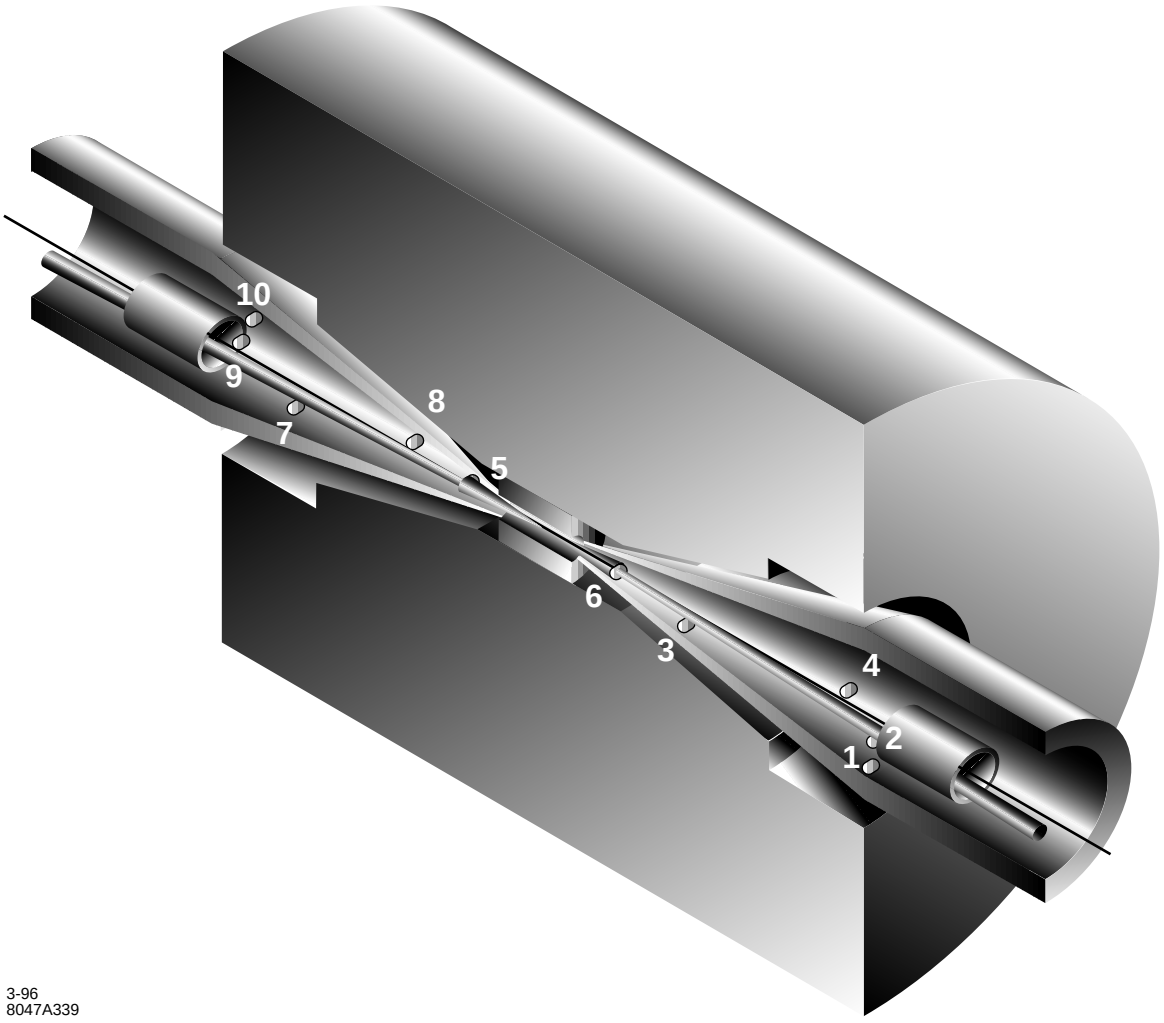


Figure B-20. *The layout of laser optical path in the IR2. The paths followed by the various beams are described in the text. The mirrors have an elliptical shape due to the overlap of the two circular beams. The shapes shown for the mirrors represent the outer edge of the flattop beams. While no provisions are shown for the edges of the mirrors or for mirror supports, space appears to be available.*

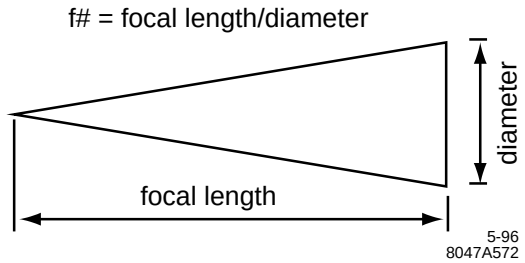


Figure B-21. *The definition of the $f_{\#}$ of an optical system.*

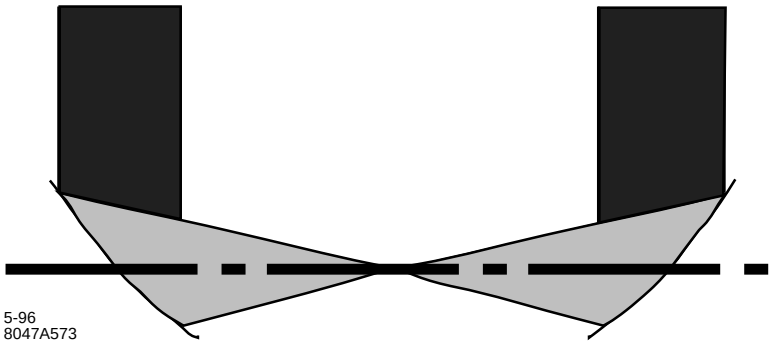


Figure B-22. *The off-axis paraboloid is both difficult to align and sends the beam path through space better occupied by the detector.*

- The two electron beams cross at only a small angle. This angle is limited by the creation of synchrotron radiation in the B-field of the detector. It must be greater than zero in order that the disrupted beam from the opposite side has a separate path by which to exit the detector. Since this angle, which in the current design is 30 mr, is small compared to the half angle associated with the focusing of the laser beam ($1/2f_{\#} \approx 57.8$ mr), it follows that the "used" beam from one side will land on the focusing optic for the other side.
- The two conversion points are separated (1 cm) in space so that softer photons (which are produced at larger angles) will diverge before the interaction point and not interact. This will introduce a "walk" into the paths of the two laser beams so that at a large distance from the IP the beams eventually separate.
- If the two beams share the same mirrors at some point (and this seems inevitable), then the mirrors probably should not be located at the points in space where the two pulses (one from each direction) overlap in time as this will significantly increase the peak intensity on the mirror.
- At first glance, there might seem to be advantages to bringing in the laser beam using an off-axis paraboloid as shown in Figure B-22. Such schemes suffer, however, both from difficulties in alignment and from a need to direct the beam through a region which is naturally occupied by the detector.
- Damage to the dielectric coatings on both the flat and curved mirrors used to transport the laser is an issue. Ultimately, this sets a limit on how close the closest optic can come to the interaction point.

The design for the laser-optical path is shown in Figure B-20. The central cylinder represents a vertex chamber. The more-or-less cylindrical object outside of that represents the rest of a generic cylindrical detector. A previous proposal [Miller 1995] in which the laser beam focuses twice has been implemented. This makes it possible to maintain near-normal incidence on all the optics in the system, and keep the laser beams in the vicinity of the conversion points inside a cylinder which is roughly defined by the outer radius of the final focusing optic used. The incident electron beams are shown as thin lines on the far side of the detector. The disrupted electron beams are shown as cones slightly offset from the incident beams. The axis of the detector is located between the incoming and outgoing electron beams.

As was mentioned above, the two laser beams enter from opposite sides of the detector and trace similar, but not identical, paths in opposite directions. These paths are distinct because the incoming electron beams make a small angle with each other, and because the conversion points are offset from each other by 1 cm.

The path followed by the laser beam incident from the right in Figure B-20 is as follows. A round and collimated (*i.e.*, parallel) beam incident upon the right (far) side of mirror 1 is reflected vertically onto mirror 2. It is, in turn, reflected

in a more-or-less horizontal direction onto the right side of mirror 3. This mirror reflects the beam onto the right side of mirror 4. At this point the beam is still parallel. Mirror 4 focuses the beam to a point over the center of mirror 6. This is done to minimize the clear region which must be provided as the beam is brought onto mirror 5 (the final focusing mirror for this side). After passing through this focus, the beam diverges and lands on the left side of mirror 5. Mirror 5 focuses the beam to a diffraction limited spot 0.5 cm past the interaction point where it backscatters from the electron beam which entered the detector from the same side.

At this point, the only task left is to extract the beam from the inside of the detector. Mirrors 6–10 are used to bring in the second laser beam to convert the second electron beam. Because of the very small angle between the two electron beams, and because the final focusing optic has its axis coincident with the electron beam, the "spent" beam will follow a path out of the IR which is nearly coincident with that of the other incoming beam. These paths are not exactly coincident because of the 20 mR angular offset of the two electron beams. This offset of laser beams is what produces the requirement for oval (as opposed to circular) mirrors.

After passing through the conversion point, the original laser beam diverges and strikes mirror 6. Mirror 6 will refocus the beam to a point just under mirror 5, from which it will diverge and land on mirror 7. Mirror 7 will recollimate the beam and send it onto mirror 8. After leaving mirror 7 the beam is once again parallel. It is then reflected from mirrors 9 and 10 in a manner similar to that by which is entered on the opposite side.

It must be noted that because mirrors 6–10 will be optimized to produce a diffraction-limited spot for the beam entering on that side, and because the beam exits at a slightly different angle, the optical path on exit will introduce aberrations into the beam. Whether or not these aberrations can be corrected subsequently so that the beam may be diagnosed after exiting the detector is an open question.

The mirrors in this scheme have been located where one finds the luminosity monitor in a conventional e^+e^- detector. Since there does not appear to be an analog of Bhabha scattering in $\gamma\gamma$ collisions, this is presumably not a major difficulty. If this region needs to be instrumented (to ensure, for instance, maximum hermiticity of the detector), then the material associated with the mirror is probably tolerable. A typical mirror will have a thickness of $\frac{1}{6}$ of its diameter. This will produce 10% of a radiation length of fused quartz over some fraction of the azimuth.

B.8.2 Optical Beam Focusing

Previous work [Meyerhofer 1995a] has sometimes assumed that Gaussian laser beams would be used at a $\gamma\gamma$ collider. This is not optimum for a couple of reasons. High-power lasers such as will be needed for this project are typically build with flattop beams to achieve the maximum fill factor. The idea is roughly the following. The cost of the laser depends on both the aperture and the peak intensity within that aperture. Having paid for both of these, the maximum energy is extracted by uniformly filling the aperture; *i.e.*, maximum fill factor. Since, as is explained below, it is probable that the disrupted electron beam will pass through the middle of the final laser focusing optic, a beam profile which peaks in the center is not optimum. Figure B-21 shows the definition of the $f_{\#}$ of a focused beam. The resulting properties of the beam at the best focus are given in Table B-4. It is seen that while the Gaussian profile provides simple analytical expressions near best focus, the flattop beam produces a much more complicated pattern as shown in Figure B-23.

Because the optimization has been done for Gaussian beams, a comparison has been made between the focal spots for the two cases. This is shown in Figure B-24. A flattop beam requires a slightly larger beam (smaller $f_{\#}$) to produce a similar spot. The exit hole required to let the disrupted beam pass through produces only a 2% energy loss in this case. Table B-5 gives a summary of the laser and electron beam parameters. Table B-6 is a summary of the parameters relevant to the laser optics design.

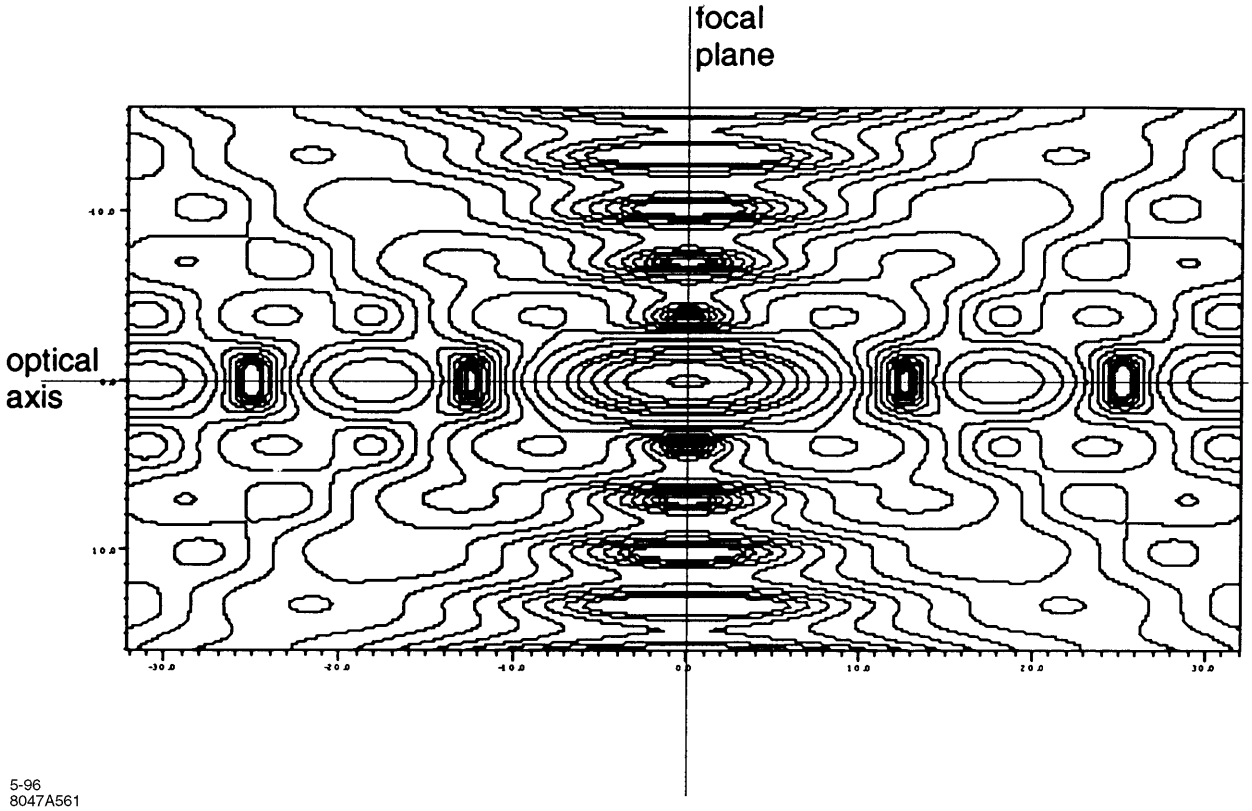


Figure B-23. *The behavior near best focus from a flattop beam.*

	Flattop	Gaussian
$I(r, z=0)$	$I_0\left[\frac{2J_1(r/s)}{r/s}\right]^2$	$I_0 \exp\left(-2\frac{r^2}{w^2(z)}\right)$
	$s = \frac{f_\# \lambda}{\pi}$	$w(z) = w_0\left[1+\left(\frac{z}{z_R}\right)^2\right]^{1/2}$
$I(r=0, z)$	$I_0\left[\frac{sin(z/z_0}{z/z_0}\right]^2$	$I_0\left[1+\left(\frac{z}{z_R}\right)^2\right]^{-1}$
	$z_0 = \frac{8}{\pi} f_\#^2 \lambda$	$z_R = \frac{4}{\pi} f_\#^2 \lambda$
Effective area	$\frac{f_\#^2 \lambda^2}{\pi}$	$\frac{\pi w_0^2}{2} = \frac{2f^2\lambda^2}{\pi}$

Table B-4. *The beam parameters at focus for Gaussian and flattop beams.*

Electron beams:	
μ-pulse / macro-pulse	90
repetition rate	180 Hz
angular offset	$\pm$15 mR
spent beam divergence	$\pm$10 mR
Laser beams:	
wavelength, λ	1.05 μm
beam profile	5-cm diam. flattop
energy	1 J / pulse
length	1.8 psec
μ -bunches	90 at 1.4 ns
repetition rate	180 Hz
power (per beam)	16 kW

Table B-5. *Beam parameters used in this study.*

Optics:	
$f_{\#}$ (flattop)	7 ($\pm$71 mr)
Distance to first mirror	35 cm
Area of first mirror	20 cm^2
Fractional area of first hole	2%
Fluence	0.05 J/ cm^2 /$pulse$
Vertex chamber:	
Inner radius/solid angle	4 cm/ 0.986
Outer radius/solid angle	12 cm/ 0.894
Total length	48 cm
Distance to first quad.:	200 cm
Masking:	100–150 mr

Table B-6. *Parameters of the design proposed in this study.*

B.8.3 Laser Damage of Optics

The particular optics used in this design are not expensive. They are neither extremely large, nor are the surfaces expected to be particularly complicated. On the other hand, their reliability will remain a critical issue if a $\gamma\gamma$ collider is to operate successfully. Definitive statements on this subject can not be made at this time because no data exists on the damage threshold for multilayer surfaces for the particular time structure of the laser pulse which would be required in this case. Data exists for the case of single picosecond scale pulses. Extensive data exists for much longer length pulses (nanosecond time scales), and for collections of nanosecond scale pulses which produce very high average powers. The three limits are considered below.

- On the time scale of a single pulse (1.8 ps), measurements have recently been made at LLNL on commercial multilayer surfaces [Stuart 1995a]. Damage thresholds in the range of 0.7 to 2 J/cm^2 have been observed. This is more than a factor of four above the anticipated fluence of 0.05 J/ cm^2.

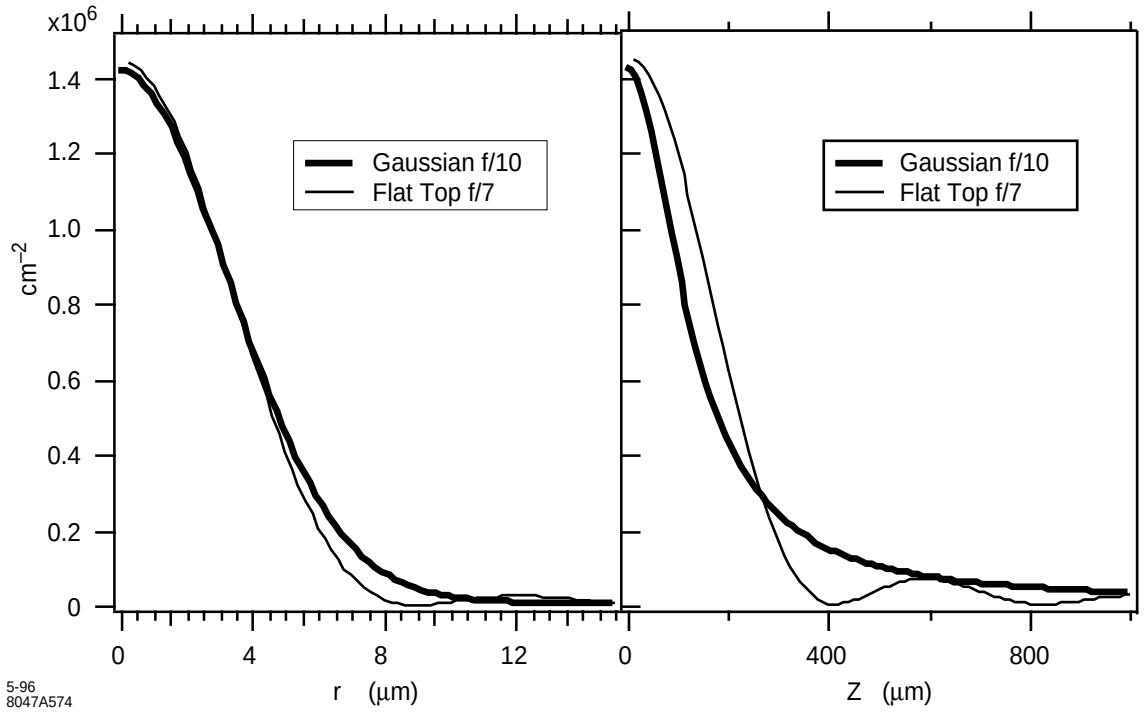


Figure B-24. *A comparison of the radial and axial dependence of the intensity for Gaussian and flattop beams. Similar sizes are obtained when the $f/\#$ of the flattop beam is about 0.7 times that of the Gaussian beam.*

- On nanosecond time scales, the situation is less clear. The anticipated fluence is 4.5 J/ cm^2 spread over 126 ns. The damage threshold for continuous pulses of this length is 100 to 200 J/cm^2. While this is well above the anticipated operating point, the validity of this averaging has not been demonstrated. No data exists for collections of picosecond pulses separated by a few nanoseconds and which accumulate to fluences of this magnitude.
- The average flux on the final optic is 0.83 kW/ cm^2. Since this is below the levels of 3 to 5 kW/cm^2 which are routinely used in the laser isotope separation program (AVLIS) at LLNL, it does not appear to be a problem.

It is worth noting that the overall scale of the optics, masking, and vertex chamber is set by the inner radius of the vertex chamber and by the damage threshold of the first optical element. If fluences higher than the design fluence are possible (and this certainly cannot be ruled out at this point), then all dimensions can be scaled down if it is desired to place the vertex chamber at a smaller radius.

At this point there is a clear need for a detailed optical design of the elements involved here. This would determine the particular surfaces required on the focusing elements to verify that a defraction-limited spot can be produced. It would also make it possible to determine the magnitude of the aberration introduced on the exit path and to determine if it is possible to correct these aberrations externally. Such corrections would be needed to use the transmitted beam to verify the quality of the focal spot.

There are also questions about the damage thresholds of the optics as well as questions about the optics resistance to damage by radiation.

The optical design here has been based on Table B-1, with A =1 J and $I = 1 \times 10^{18}$ W/cm^2. As discussed elsewhere, the non-linear effect in Compton conversion is not negligible. the non-linearity can be avoided by increasing A to 3 J and decreasing I by a factor of three. This will increase the requirement on the laser power, as well as power on the mirror by a factor of 3. The focusing optics also changes from f/7 to f/10, with the result that the mirrors are located a little further out from the IP.

B.8.4 Ideas on Relaxing Average Laser Power Requirement Via Multipass Optics

In a gamma-gamma collider, a high-energy electron bunch is converted into a burst of gamma rays by Compton scattering with an intense laser pulse. Assuming the laser source has the same pulse structure and repetition rate as that of the electron beam and assuming that each laser pulse is used only once, the required average laser power would be around 20 kW, three orders of magnitude higher than what has been achieved with any laser having TW peak power. However, the required average power can be significantly reduced if the laser pulses can be reused, because the laser pulse suffers little loss in energy after each scattering. In this section, we explore the possibilities of reusing the laser pulses with specially designed optics.

There are two approaches to this problem. The first one is a multipass approach in which a laser pulse is made to pass through the conversion point a finite number of times before being thrown away. In this way, the average power is reduced by reducing the number of pulses needed. The second approach may be called pulse stacking in which a train of weaker laser pulses are stacked up in an optical cavity to make a stronger pulse for intracavity conversion. In designing optics for both approaches, using transmissive optical elements should be avoided if possible to minimize the nonlinear transverse and longitudinal pulse distortion at high power.

To illustrate the idea, an example of multipass optics based on all reflective elements is schematically shown in Figure B-25, in which a laser pulse is made to pass through a CP eight times, four in each direction. This is done by using two mirror banks each with eight individual mirrors. Each individual mirror is placed in the numbered order along the optical pass according to the sequence the laser pulse is kicked. Such a kicking pattern guarantees the same pass length for each round trip the laser pulse makes passing through a CP. With some modification, the scheme in Figure B-25 can also be made to accommodate two conversion points separated by a few mm while requiring laser pulses to pass through the two conversion points in opposite direction.

The reduction factor in the required laser average power for the multipass approach is limited because the number of mirrors that can be utilized is limited due to the tight space near the interaction region, and due to the build-up of aberration. To make better use of the limited space, it is desirable to have laser pulses bounced back and forth along the same pass, thus forming a cavity mode. By stacking up weaker externally injected pulses inside a cavity, the factor of reduction in average power could be up to the cavity Q.

For gamma-gamma colliders the desirable cavity mode should have certain characteristics. The double confocal resonator seems to be an ideal choice for this purpose. A double confocal resonator is effectively made of two usual confocal resonators. The usual confocal resonator is formed by two curved mirrors separated by a distance equal to the radius of curvature of the mirrors. Putting two confocal resonators together, one gets a ring resonator with four identical mirrors separated by equal distance. Folding such a ring resonator with two flat mirrors one gets the double confocal resonator shown in Figure B-26.

The double confocal cavity has several advantages:

- A dominant cavity mode can be made to have a central peak at each focal point and an annular shape at each mirror location.

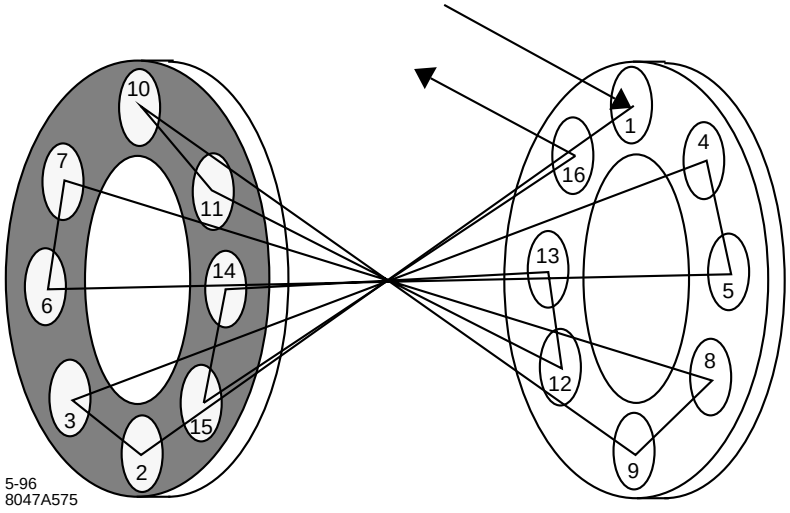


Figure B-25. *Multipass optics with reflecting mirrors.*

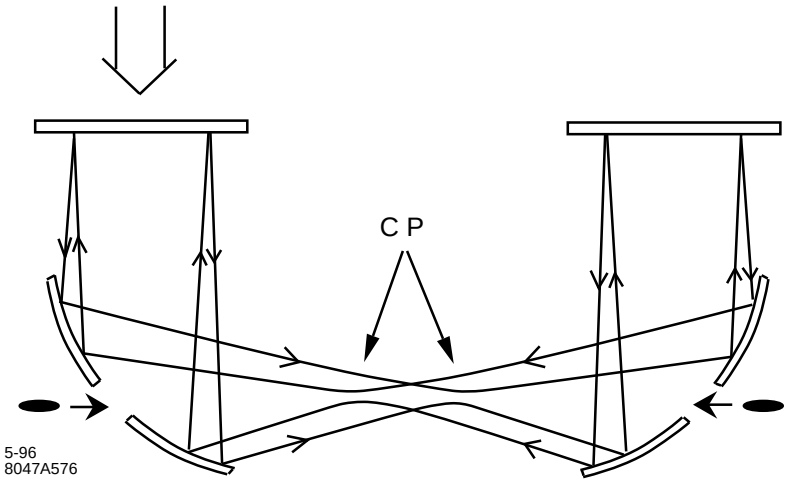


Figure B-26. *A double confocal resonator.*

- The focal spotsize or Rayleigh range is determined by the limiting intracavity apertures and are therefore easily adjustable.
- The annular mode distribution on the cavity mirrors allows large holes on the mirrors for electron beam and debris passage without sacrificing cavity Q.
- The annular shape on the mirrors also gives larger mode sizes thus reducing the power loading.
- With proper external mode for injection, two focal points can be formed each opposing to an incoming electron beam.
- In addition to these advantages, only one unavoidable transmissive window is used in Figure B-26 to couple laser pulse into the cavity in the vacuum system. Because a weaker pulse is needed for stacking, the power going through the window can be quite small compared to other approaches requiring full peak power injection.

The excitation of the desired mode in the cavity is largely dependent on the external mode preparation. Roughly speaking, the injection mode should have an annular amplitude distribution and a nearly flat phase front at the location of the transmissive window. If the external mode exactly matches the desired cavity mode, the cavity mode can be established right away, otherwise the resulting mode deterioration and slower intracavity power build-up could affect collider performance. There are ways to convert the usual Gaussian mode from a laser into an annular shape, for example by using an axicon or a profiled transmissive element. Fortunately, all these can be done outside the crowded interaction region and high-vacuum environment.

B.9 Gamma-Gamma Final Focus System

The goal of the FFS for the IR2 is to produce $\beta_x^* = \beta_y^* = 0.5\,\text{mm}$, as explained in Section B.3.3. However, the current version of the $\gamma\gamma$ final focus system has $\beta_x^* = 0.9\,\text{mm}$ and $\beta_y^* = 0.7\,\text{mm}$. This is not accidental because with these beta functions at the IP the chromaticity of the final focus doublet is the same as in the e^+e^- case. Efforts to find smaller β^* values resulted in higher x and y chromaticities, implying a greater sensitivity to the quadrupole placement tolerance and also a greater complexity and length of the FFS. The current solution of the FFS for the IR2 has a similar complexity as that for the IR1, being essentially an adaptation of the already existing solution for e^+e^- case to the new regime of the $\gamma\gamma$ collision.

B.9.1 Beam Parameters

Table B-7 lists the electron beam parameters necessary for a discussion of a final focus system of a $\gamma\gamma$ interaction region. Note that beta-functions at the IP are optimized differently from the ones for the e^+e^- collision—with a relaxed β_y^* and a tighter β_x^*. The following considerations are behind this choice: High-energy γ quanta appearing in a conversion of photons on electrons are emitted within the angular spread $1/\gamma$ towards a direction of the parent electron.

Thus, a spotsize of the high-energy γ beam at the IP reflects a spotsize of the electron beam plus additional contribution due to the angular spread of the γ quanta at the CP. This contribution, $\delta\sigma^*_{x,y}$, depends from a distance b from the CP to the IP

$$\delta\sigma^*_{x,y} = b/\gamma \quad . \tag{B.27}$$

E (gev)	250
ϵ_x (m×rad)	1×10^{-11}
ϵ_y (m×rad)	1×10^{-13}
β_x^* (mm)	0.9
β_y^* (mm)	0.7
Horizontal beam size at the IP (nm)	96
Vertical beam size at the IP (nm)	8.5

Table B-7. *Beam parameters*

We would like to have

$$\delta\sigma_{x,y}^* \lesssim \sigma_{x,y}^* \quad . \tag{B.28}$$

Therefore,

$$b \lesssim \sigma_{x,y}^* \gamma \quad . \tag{B.29}$$

Thus, it is more convenient to have larger β_y^* and, correspondingly, larger σ_y^* because it allows the CP to be placed further away from the IP. At the same time, it is necessary to reduce β_x^* to compensate the reduction in the luminosity when β_y^* is increased. Note that, contrary to the e^+e^- case, there are no constraints on the beam profile at the IP arising from the beamstrahlung effects.

B.9.2 Final Focus Doublet

For linear colliders, the chromaticity is defined as the change in the IP waist position when the particle has an energy unequal to the design energy. A change in the horizontal or vertical waist position is characterized by the presence of a term $x^{'2}$ or $y^{'2}$ in the beam line Hamiltonian, which is precisely the terms in the Hamiltonian of a drift. The horizontal or vertical chromaticity is thus characterized by the presence of terms $\delta x^{'2}$ or $\delta y^{'2}$, where δ is the fractional energy error. Since the change in the horizontal or vertical IP position due to the these terms is given by

$$\Delta x = \frac{\partial H}{\partial x'} \qquad \text{or} \qquad \Delta y = \frac{\partial H}{\partial y'} \tag{B.30}$$

and the derivative of a quadratic function introduces a factor of 2, it is usual to define a horizontal and vertical chromatic length L_x^c and L_y^c as the coefficient of $\delta x'^2$ or $\delta y'^2$ respectively.

The spread in spotsize from spread in incoming angle can be written

$$\frac{\Delta x}{\sigma_x^*} = \delta\xi_x \frac{x'}{\sigma_{x'}^*} \quad , \quad \text{where} \quad \xi_x \equiv \frac{L_x^c}{\beta_x^*} \quad . \tag{B.31}$$

The quantity ξ_x is called the chromaticity. It is a dimensionless number. A value of one would indicate that the chromatic aberration gives a contribution to the beam size equal to the linear term. the contribution to this term coming from the final doublet can be calculated by the formula

$$\int ds \; k_x(s)\beta_x(s) \quad . \tag{B.32}$$

since the phase advance from the IP to elements of the doublet is very close to $\pi/2$.

	Length, (m)	Gradient, (kg/mm)	Bore radius, (mm)	Beam-stay-clear
F-quad	1.43	4.50	3	$11\sigma_x, 50\sigma_y$
D-quad	1.16	-3.38	4	$19\sigma_x, 41\sigma_y$

Table B-8. *Doublet parameters*

It is well known that the chromaticity of a lattice with only quadrupoles will be non zero, and to compensate the chromaticity one must add sextupoles to the beam line. To minimize sextupole strength, one first seeks the quadrupole configuration that minimizes the chromaticity.

The minimum chromaticity from quadrupoles occurs by placing strong quadrupoles as close to the IP as possible, without interfering with the function of the detector. Doing this in both planes of course requires a doublet. The rays on the IP side of the doublet will be focused to the IP. The divergence of the rays on the upstream end of the doublet will have a divergence that is characterized by the length of typical beam line modules of phase advance $\pi/2$ or π. Thus the function of the doublet is to focus parallel rays to the IP.

Thus the doublet design can be separated from the beam line design. One first seeks the double giving the minimum chromaticity parameters, taking into account detector requirements, constraints of materials available for quadrupole fabrication, tolerances on quadrupole position and field strength, and synthrotron radiation within the doublet.

As a first attempt to design the final focus system for $\gamma\gamma$ collisions, we decided to keep chromaticity of the final focus doublet close to the chromaticity of the e^+e^- final focus. Thus, with $L^* = 2$ m, the minimum beam-stay-clear requirement of $10\sigma_{x,y}$, and the maximum pole-tip field in the permanent magnet quadrupoles of 1.35 T, we arrived at the doublet parameters described in the Table B-8. Note, that the quadrupole nearest to the IP is of the F-type and the length of the drift space between the F and D quadrupoles is 0.3 m. This doublet has an x-plane chromaticity $\xi_x = 3100$ and y-plane chromaticity $\xi_y = 24500$. These values are to be compared with $\xi_x = 1100$ and $\xi_y = 23000$ in the e^+e^- case.

The size of these chromaticities indicate the precision with which the chromatic correction must be made. This has, of course, direct implications for the system tolerances. It also places constraints on the synchrotron radiation allowed in the system, because the change of particle energy within the system implies a failure of the chromatic balance built into the optics. However, for the present design of the doublet, the increase of the beam spotsize at the IP due to the Oide effect is negligible.

B.9.3 Chromaticity Compensation

We follow a standard approach to the chromaticity compensation of the final focus doublet. Similar to the e^+e^- final focus system, the $\gamma\gamma$ final focus system consists of the five modules. These are, in order of their location beginning from the IP: final transformer (FT), vertical chromaticity correction section (CCY), beta-exchanger (BX), horizontal chromaticity correction section (CCX) and beta-matching section (BMS). All modules have exactly the same functions as in the e^+e^- case. Figure B-27 shows the beta and the dispersion functions along the final focus system from the entrance of the BMS to the IP. The total distance from the entrance of the BMS to the IP is about 1600 m.

Figure B-28 shows the vertical and horizontal beam size as a function of the momentum offset Δ of a mono-energetic beam. The spotsizes are given in units of the values for zero momentum offset. The momentum bandwidth for a 10% blow-up of either spotsize is $\pm 0.6\%$.

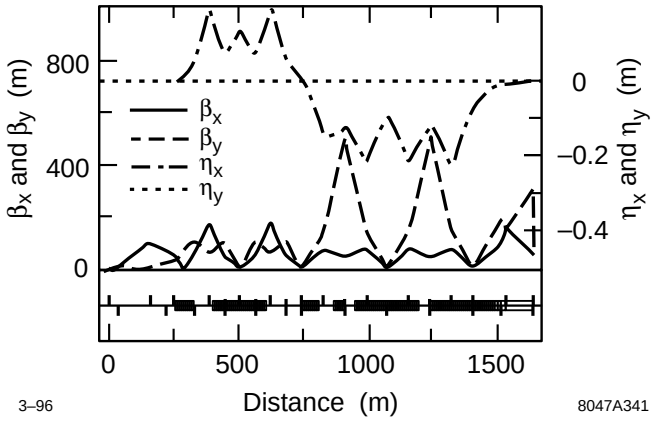


Figure B-27. *Horizontal and vertical beta functions from BMS to IP for 1600 m-long $\gamma\gamma$ final focus system at 500 GeV.*

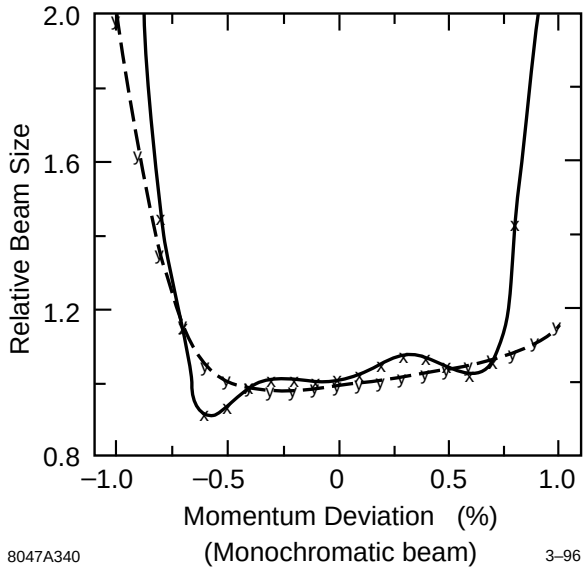


Figure B-28. *Relative increase of spotsizes as a function of momentum offset Δ of a monoenergetic beam.*

It was recently shown (Dick Helm) that by increasing the overall length of the Final Focus section to 1750 m one can have $\beta_x^* = \beta_y^* = 0.5$ mm with the energy bandwidth about $\pm 0.5\%$.

B.10 Extraction and Diagnostic Line

The extraction line has the important functions of optimizing luminosity, characterizing beam properties at the IP and transporting beams from the IP to a dump with minimal background in the particle detector from this or any secondary function along the way. Beam characterization includes measurements of current, position, profile, energy, polarization, and low-order correlations on a bunch-to-bunch basis for feedback and stabilization. A prototype optical

and diagnostic layout will be developed that provides such functions. In many respects it will be similar to the corresponding e^+e^- dump line that was presented in Section 11.8.

One presumed difference with that line is the difficulty of disposing of the high-power photon beam. We note that the outgoing beamstrahlung power approaches 10% of the incident, primary beam power for the e^+e^- IR. This was dealt with by implementing a common dump for both leptons and photons. The main question to be answered here is what distance is then available and whether it is reasonable from the standpoint of detector backgrounds and the various secondary functions that one would like to implement.

To accomplish all of the various tasks, we need to know the detailed composition and characteristics of the outgoing beam(s) under different, possible circumstances. Clearly, these characteristics depend on those assumed for the incoming beam. Based on some assumed set of incoming beam parameters, the transport optics, and the CP and IP conditions, we can then make predictions for the outgoing beams that are used to guide the design of the outgoing, beam line optics. Typically, the outgoing beams have a significant number of pairs as well as the primary, degraded electrons from the Compton conversion and more photons than leptons from multiple scattering within the strong laser conversion field.

Thus, the primary difference for this IP is the degraded electron beam that is highly disrupted but also necessarily includes a significant fraction of electrons with their full incident energy.

Clearly, the detector fields begin to have a serious influence on the incoming and outgoing beam characteristics with such large crossing angles. Thus, if background simulations for the detector imply an unacceptable situation, this procedure has to be iterated until a consistent solution is achieved. This is just now being done for the e^+e^- channel as described in Chapter 12.

While our overall goal is to optimize the luminosity while disposing the various beam components into their respective dumps, it is also important to provide any monitoring and feedback that can optimize the usable collision rate at the IP. Thus, beyond simply dumping the beam, there are other functions that run from the absolutely necessary to the desirable that will be considered as part of the dump line in roughly descending order of importance:

- Beam control and stabilization.
- Diagnostics and monitoring—including luminosity and polarization.
- e, μ, n, γ secondary beams and parasitic experiments.
- Polarized sources for γ, μ and e^+ beams.
- Energy recovery.

Separate discussions on these possibilities can be found in Section 11.8 of the ZDR for the e^+e^- channel. While the outgoing line is more difficult here, it can be accomplished in a similar way to that discussed in Section 11.8 because that design was made to be compatible with such a possibility there.

B.11 Laser Technology I: Solid State Lasers

As we have seen in Section B.3.2, the optical beam for $\gamma\gamma$ or γe^- colliders consists of a sequence of TW micropulses, each a few ps long, with an average power of tens of kW. The requirements are summarized in Table B-9. Unless multipass optics can be used in the interaction region, these are the requirements for the lasers. Solid state lasers

Energy per micropulse	1 J
Pulse duration	1.8 ps
Focusing f-number	7
Wavelength	0.8-1 μm
Pulse structure	90 micropulses(1.4-ns separation)
	180-Hz macropulse rate
Average power	16 kW

Table B-9. *Laser requirements for NLC γ-γ option*

meeting the requirements would probably be available for the colliders to 500×500 GeV, but FEL would be used for higher energy colliders.

The solid state lasers for the $\gamma\gamma$ colliders have been discussed in two recent articles [Meyerhofer 1995a, Clayton 1994, Clayton 1995]. There will be two of these laser systems, one each for each colliding electron beam. While the energy, pulse duration, and focusing can be met with currently operating lasers, [Perry 1994a], these lasers have not yet met the average power requirements. The average power of high peak power systems has, however, been increasing rapidly recently, driven by activities such as the Isotope Separation program at LLNL and facilitated by the development of high power laser diode pump sources. The system requirements could also be reduced by using a multipass conversion point. It is expected that the system requirements will be met with a series of 1-kW, diode-pumped, solid-state, chirped pulse amplification laser systems. These unit cells will be fed by a single, phase-locked oscillator to insure timing stability.

Many of the components of the required laser system can be achieved with technology which is currently being developed for applications other than the $\gamma\gamma$ collider. As a result, a single unit cell prototype laser module could be developed over the next few years.

B.11.1 Laser Materials

Although both dye and excimer laser systems can easily meet the short-pulse requirement of the NLC, achieving the energy and beam quality requirements with lasers based on these materials would be difficult and expensive. The difficulty is associated with the low saturation fluence (energy storage) of these materials. The saturation fluence and upper state lifetime limit the amount of energy which can be stored and extracted per unit area (volume) from a laser material. Pulse energies on the order of one J would require laser apertures of approximately 500 cm^2 for dye- and excimer-based systems. Solid-state lasers offer an increase in saturation fluence between two and three orders of magnitude greater than dye or excimer lasers. This makes possible the development of extremely compact, high-energy lasers based on solid-state lasing media. Unfortunately, the high saturation fluence of solid-state materials cannot be accessed directly with short-pulses due to limitations on peak power imposed by the nonlinear refractive index.

In principle, production of a 1-J, 1-ps laser pulse could be accomplished by producing a low-energy 1-ps, 1.053-μm laser pulse and directly amplifying it in a Nd:Glass laser chain. Because of the high saturation fluence, the final amplifier cross section could be less than 1 cm^2. The extracted power density would be in excess of 1 TW/ cm^2, close to the damage threshold of most materials [Stuart 1995a]. A second, more severe limitation on the amplification chain is provided by the nonlinear index of refraction in the material [Siegman 1986]. Self-focusing and filamentation of the

laser pulse can occur when the accumulated nonlinear phase (B-integral) exceeds π,

$$B = \frac{2\pi}{\lambda}\int_0^L n_2 I(t)dz = \frac{2\pi}{\lambda} n_2 I_{\rm peak} L \quad , \tag{B.33}$$

where n_2 is the nonlinear index of refraction and L is the length of the medium. The power density in a medium is thus limited by the B-integral. For a given length amplifier, the peak intensity must satisfy $I_{\rm peak} < \lambda/2n_2 L$. For a 10-cm-long Nd:Glass laser amplifier, the peak intensity must be less than 2 GW/ cm^2. For a pulse length of 1 ps and an energy of 1 J, the cross-sectional area must be greater than $A_L > E/I_{\rm peak}\tau = 500\,{\rm cm}^2$, which is to be compared with the $> 1\,{\rm cm}^2$ estimated from the saturation fluence. One solution to keeping the B-integral small is to use a larger diameter amplifier. Unfortunately, the use of a large diameter amplifier for short-pulse amplification eliminates the advantage of solid-state media.

B.11.2 Chirped-Pulse Amplification

This problem can be overcome by the use of Chirped-Pulse Amplification (CPA) [Strickland 1985]. This technique allows smaller amplifiers to be used, which reduces the cost and increases the repetition rate. Chirped-Pulse Amplification circumvents self-focusing by temporally stretching the pulse before amplification and recompressing it afterwards. This reduces the B-integral in the lasing medium by the compression ratio, the ratio of the stretched to compressed pulse durations. While optical fibers and prisms can be used to stretch or compress a pulse, the simplest technique involves a grating pair. A grating pair can be used to impart a positive [Martinez 1987] or negative [Treacy 1969] chirp to a short pulse. The sign of the chirp is defined by the time derivative of the frequency.

The CPA concept is shown in Figure B-29. A short, low-energy pulse is generated in an oscillator. The pulse is then stretched by a factor greater than 1000 in grating pair. To obtain a positive chirp a telescope is used between the grating pairs to invert the sign of the natural negative chirp associated with grating dispersion. To limit the size, complexity and cost of the telescope, a positive chirp is generated when the beam size and energy are low (before amplification). The pulse is now long enough for safe amplification to high energy. The pulse is recompressed in a second, grating pair (no telescope), generating a high-energy, ultrashort laser pulse.

The initial grating pair imparts a phase delay proportional to the frequency. This produces a pulse which has different frequencies spread out in time (chirped pulse). The second grating pair imparts a phase delay which is the inverse of the first grating pair, thereby removing the chirp and recovering the short pulse. Ideally, with the amplifiers placed between the two sets of gratings, the only change in the chirped pulse is its amplitude and the temporal characteristics of the pulse at the input and output are the same. Unfortunately the amplifier can modify both the amplitude [Perry 1990a, Chuang 1993] and phase [Chuang 1993] structure of the pulse. The amplitude changes include gain narrowing and pulling [Perry 1990a]. The primary phase change is self-phase modulation of the chirped pulse in the amplifier chain [Chuang 1993, Perry 1994b]. All of these effects must be minimized in order to obtain optimal pulses upon compression.

B.11.3 High-intensity, Short-pulse Laser Systems

There have been very dramatic advances in the technology and application of chirped-pulse amplification technology in the past five years [Perry 1994a]. Original CPA systems employed conventional mode-locked Nd:YAG or Nd:YLF oscillators producing transform-limited 50–100 ps pulses. The bandwidth necessary to achieve picosecond pulses was obtained by passing the 50–100-ps output of the oscillator through long (50–1000 m) lengths of single-mode fiber. Self-phase modulation in the fiber produced a chirped pulse with increased bandwidth. The pulse duration was

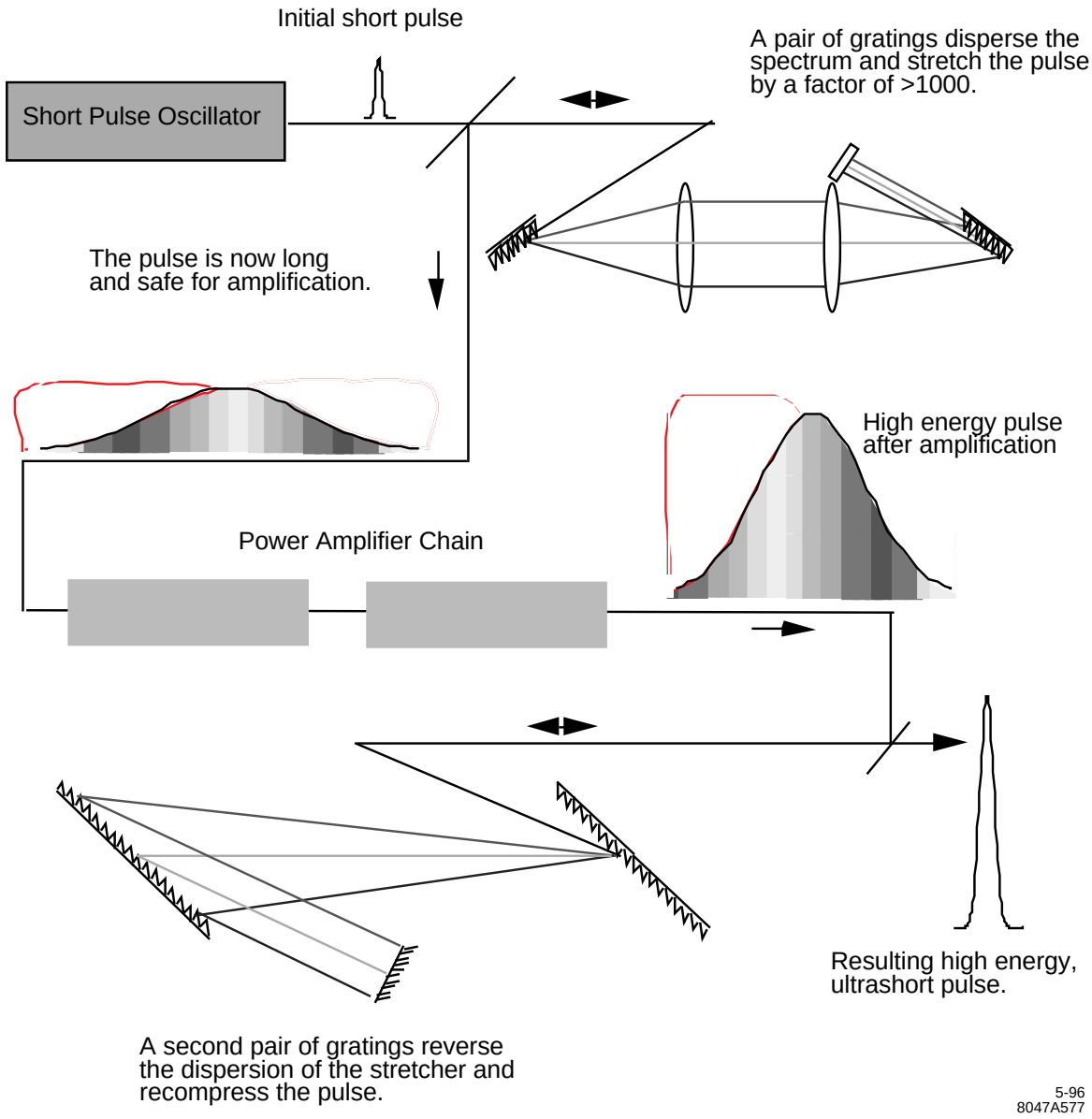


Figure B-29. *Chirped-pulse amplification concept.*

increased from the original 50–100 ps by group-velocity dispersion in the fiber or the addition of a grating-pair pulse stretcher as mentioned previously. This method for obtaining stretched, chirped pulses has been rendered obsolete by the development of Kerr lens mode-locked oscillators which directly produce transform-limited 10–1000 fsec pulses. These pulses are then directly stretched in time by factors on the order of 10^4 by the grating stretcher. Elimination of the complex nonlinear interaction in the fiber is one of the many important advances which have led to the rapid and widespread acceptance of chirped-pulse amplification lasers.

The pulse exiting the stretcher has an energy of 0.10-1 nJ and a duration of $\sim$1 ns. In the power amplifier section, the pulse must be amplified by a factor of 10^9 to the Joule level. This is typically achieved through the use of a multipass regenerative amplifier and a series of single or double pass amplifiers. In a regenerative amplifier, a pulse is switched into a cavity and makes multiple passes through an amplifier and is subsequently switched out. The regenerative amplifier brings the energy to the few milliJoule level (an increase of seven orders of magnitude) while the single and double pass amplifiers provide the remainder of the gain. Because of the losses inherent in the amplifiers, the total gain is of order 10^{10}–10^{11}.

After amplification, the pulse is compressed back to a short pulse by a pair of compression gratings. The damage threshold of these diffraction gratings is one of the most important limits on CPA lasers. The gratings must exhibit high damage threshold for both the long duration of the stretched pulse (ns) and the short duration of the compressed pulse (ps). Since the physical mechanism of optical damage changes from the nanosecond to picosecond regime [Stuart 1995a], producing high-damage threshold diffraction gratings is both a scientific and technical challenge. Commercially-available metallic gratings exhibit damage thresholds as high as 250 mJ/ cm^2 for ps pulses. Recent advances in grating technology [Boyd 1995] have increased the damage threshold of metallic gratings to over 400 mJ/ cm^2 for pulses in the range 0.1 to 50 ps. While these gratings do enable a factor of two increase in the peak power density achievable with most CPA systems, they will not be useful for the combined high average and high peak power requirement of the NLC. A new concept in diffraction gratings based on multilayer dielectric materials [Perry 1996] can, in principle, meet both the average power and peak power requirements of the NLC. These gratings have achieved a damage threshold of 600 mJ/ cm^2 for 100 fsec pulses [Perry 1995]. They should exhibit a significantly increased threshold for average power damage relative to thin-film metallic gratings, however this is yet to be demonstrated.

One must also be concerned about power density of the compressed pulse in transporting it to the interaction point. If the intensity exceeds 10^{12} W/ cm^2 it may cause plasma formation on solid surfaces [Stuart 1995a]. In addition, the nonlinear index of refraction of the pulse in air, in any windows, and in the focusing elements will destroy the beam quality of the pulse and make it impossible to focus. For example a 10^{12} W/ cm^2 pulse passing through a 1-cm-thick glass window ($n_2 = 3 \times 10 - 16$ cm^2/ W) the B-integral is 6π so that significant self-focusing would occur. The output diameter of the TW pulse must be large enough so that the cumulative B-integral in the transport line is limited to less than approximately π. This means that in practice, the intensity of the compressed pulse should be kept below 10^{11} W/ cm^2 before focusing, implying that the beam cross section must exceed 10 cm^2.

B.11.4 Synchronization and Repetition Rate

Two additional topics are important in the design of the laser system for a $\gamma\gamma$ collider. The laser beam must be synchronized to the electron beam to a fraction of the laser pulse duration and the repetition rate of the laser system must be matched to the repetition rate of the linac. The laser can by synchronized to the electron beam by driving the laser oscillator at a subharmonic of the linac rf. The phase relationship between the rf and the laser pulses can be maintained by electronic feedback [Rodwell 1986, Rodwell 1989]. Subpicosecond timing jitter has been demonstrated using these techniques [Rodwell 1986, Rodwell 1989]. A typical mode-locked oscillator operates with a driving frequency of 35–60 MHz, producing a 70–120-MHz pulse train. The length of the cavity is matched to driving

Central wavelength	1.053 μm
Gain bandwidth	$\sim$200 Å
Saturation fluence	5 J/ cm^2
Nonlinear index, n$_2$	6×10^{16} cm^2/ W

Table B-10. *Selected properties of Nd:phosphate glass lasing materials.*

frequency to produce the mode-locked, short duration, pulse train. The pulse train is detected with a fast photo-diode. The photo-diode signal is mixed with the input rf and error signal is used to phase shift the rf driving the mode-locker [Rodwell 1989]. It is important to note that to maintain the synchronization, the oscillator must be actively mode-locked. In addition, the path length of the laser after the oscillator, through the laser system and through the transport must be stable to less than the pulse duration, 0.3 mm for a 1-ps laser pulse.

B.11.5 1-ps, 1-J Laser System for Nonlinear QED Experiments

Recently a 0.5-Hz repetition rate, 1-μm, 1-ps, 1-J, chirped-pulse amplification (CPA) laser system has been demonstrated using a flashlamp-pumped, Nd:glass, zig-zag slab amplifier [Bamber 1995]. The system has been installed at the Stanford Linear Accelerator Center as part of the E-144 experiment to study nonlinear QED [Bula 1992] in collisions with 50-GeV electrons. With the exception of repetition rate and pulse format, many of the performance parameters required for $\gamma\gamma$ colliders are being examined with this system. In addition to demonstrating the laser energy, and pulse width, the jitter of the laser pulse with respect to the linac rf is currently less than 2 ps [Bamber 1995]. Compton-scattered gammas and recoil electrons due to multiphoton Compton scattering have been observed [Bula 1995]. Unfortunately, the average power of this laser system is four orders of magnitude less than required for the NLC $\gamma\gamma$ collider.

B.11.6 Average Power

As mentioned previously, the single pulse laser requirements for converting the NLC e^+e^- collider into a $\gamma\gamma$ or γe^- collider can be met by solid-state lasing materials. A number of different solid state materials are used for short-pulse, high-intensity, laser systems [Kmetec 1991, Perry 1991, Ditmire 1993, Beaudoin 1992, Salin 1991, White 1992] Both Ti:Sapphire [Salin 1991, Stuart 1995b] and Nd:Glass [Strickland 1985] have been used to generate high-intensity, ultrafast, laser pulses with wavelengths in excess of 1 μm. The advantage of using Ti:Sapphire is that its larger gain bandwidth allows shorter pulses to be generated and amplified, whereas a pure Nd:Glass system is limited to pulse durations of order 1 ps, which is sufficient for this application. The saturation fluence of Nd:Glass is approximately an order of magnitude higher than Ti:Sapphire, making Nd:Glass an attractive candidate for the NLC. Unfortunately, the thermal conductivity and thermal shock limit are low for glasses. As a result, although conventional Nd:glasses can meet the peak power requirements of the NLC, they cannot meet the average power requirements. The relevant properties of Nd:Glass are listed in Table B-10.

New glass hosts currently under development offer a nearly factor of two increase in the thermal shock limit. These glasses could, in principle, make possible a diode-pumped Nd:Glass based NLC laser. No laser has yet been constructed from these advanced glasses. However, two kilowatt-class (long pulse) lasers are under development at LLNL utilizing these new glasses. Performance data from these lasers will be invaluable in analyzing the suitability of glass-based systems for the NLC.

In addition to developments in laser glass, recent advances in crystal hosts are also encouraging for NLC options. New crystals which have been "engineered" for diode-pumping and high average power operation are now emerging. One attractive candidate is Yb:S-FAP. This is a fluroapatite crystal host for the Yb lasing ion. The material functions well as a laser near 1 μm and has sufficient bandwidth to support pulses of 2–5 ps in duration. It has thermo-mechanical properties which are substantially better than even the advanced glasses. Furthermore, it has a long upper-state lifetime ($\approx$1 ms) and an absorption band at 900 nm which make it nearly ideal for diode pumping with efficient AlGaAs diodes. The Yb:S-FAP crystal can not yet be grown in sufficient sizes to meet NLC requirements. However, a large amount of effort is currently being devoted to further developing S-FAP and related crystal hosts. It is reasonable to expect that large scale crystals would be available within one to two years.

In short, there are several options for meeting the average power requirements of the NLC. These options include: 1) direct, diode-pumped Nd:Glass based lasers incorporating advanced athermal glass, 2) direct, diode-pumped, broad-bandwidth crystals (*e.g.,* Yb:S-FAP or others) and, 3) two-stage laser-pumped lasers such as a long pulse ($\approx$10 ns) neodymium based laser pumping a short-pulse Ti:Sapphire laser. We have not yet performed an optimization study for the NLC laser which would compare the performance and cost of these various options. Such a study would be part of the conceptual design of the NLC $\gamma\gamma$ collider.

In addition to requiring advances in high average power laser materials, advances in diode laser technology are also required to meet the NLC specifications. However, there are major efforts on advancing diode laser technology already underway as part of both military- and civilian-led projects. High average power diode laser arrays which would meet the requirements of the NLC are already under development at LLNL and elsewhere. Current high peak power diode arrays have generated 1.45-kW average power [Beach 1994]. The continued development of diode laser technology and the associated thermo-mechanical systems will be only moderately influenced by the approval of the NLC. Instead, the NLC will reap the benefit of substantial development effort which is expected to produce diode packages which can meet the NLC requirements well in advance of the NLC construction schedule.

B.11.7 NLC Laser Concept

The proposed laser system for the NLC $\gamma\gamma$ option consists of two $\sim$16-kW laser systems built out of 1-kW unit cells. A schematic of the unit cell is shown in Figure B-30. All of the cells are fed by a single, phase-stabilized oscillator, ensuring synchronization of all of the laser pulses with the electron beam. Each of the unit cells consists of a series of diode-pumped, solid-state, laser amplifiers. The pulses are subsequently compressed in a grating pair and stacked into a single pulse train.

Pulse stacking from individual unit cells into a single pulse train occurs via polarization switching as shown in Figure B-31. The output of a single unit cell is a small pulse train which is s-polarized (linearly polarized out of the paper in the figure). These pulses are reflected along the primary axis by a thin film polarizer. These polarizers are designed to reflect s-polarized light with greater than 99% efficiency while simultaneously transmitting p-polarized light with similar efficiency. Light from the first unit cell passes through a Pockels cell which has an applied voltage sufficient to provide a half-wave retardation. This rotates the polarization of the pulse 90° from s-polarization to p-polarization (linearly polarized in the plane of the paper). This p-polarized light now passes through the next thin film polarizer. The s-polarized pulses from the next unit cell are reflected from the thin-film polarizer. Combined with the pulses from the first unit cell, we have two sets of orthogonally polarized pulses incident on the next Pockels cell. The Pockels cell is initially held at ground while the p-polarized pulses from the first unit cell pass. After these pulses pass, the Pockels cell voltage is switched to half-wave voltage. The s-polarized pulses from the second unit cell are rotated to p-polarization upon transmission through the Pockels cell. The result is now a combined train of p-polarized pulses along the same optical axis. The procedure is repeated for each subsequent cell. In this scheme, the pulse-compression

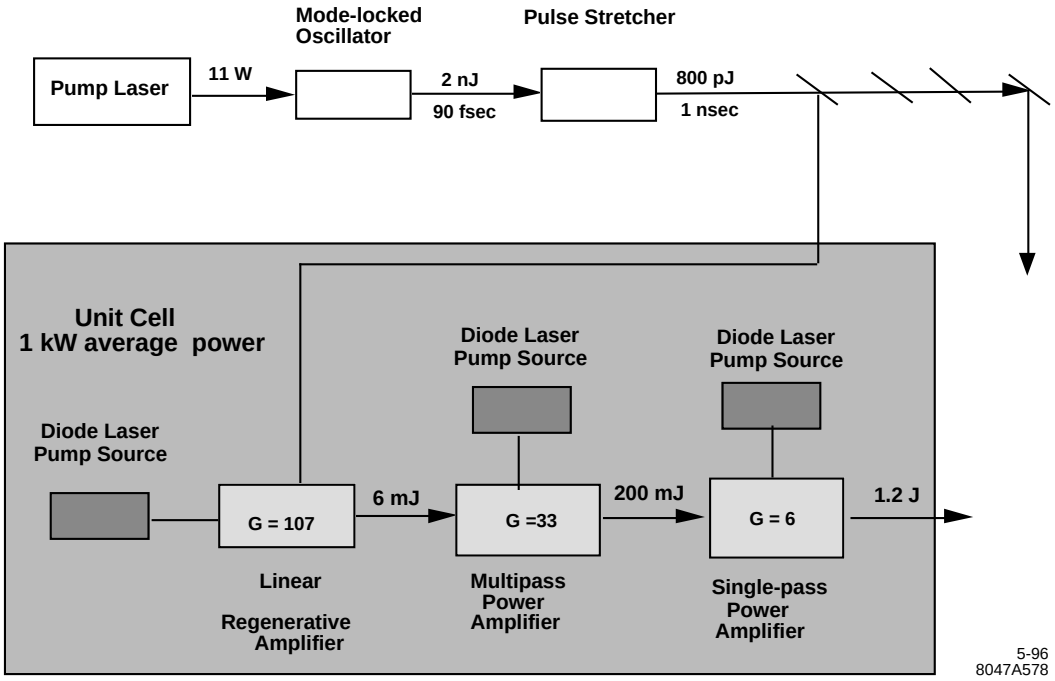


Figure B-30. *1-kW average power unit cell.*

gratings probably should be placed after the pulse stacking so that all of the pulses will be compressed in the same grating pair and to keep the B-integral down.

B.11.8 A Ring Configuration for Multiplexing and Polarization Control

The baseline design would be to use each pulse once, with a single pass through a Pockels cell for polarization control. However, we will explore during the design the possibilities of re-using the pulse, thus significantly reducing the average laser power requirements, and hence cost. This was discussed in Section B.8.4. Here we discuss in detail an example of such schemes based on a ring regenerative device first demonstrated at LLNL in 1990 [Perry 1990b]. The regenerative ring (Figure B-32) is a photon recirculator which is based on the original ring regenerative amplifier concept. The pulse is initially p-polarized (from the pulse stacker) and enters the cavity through a thin-film polarizer, TFP 1. The Pockels cell is initially at ground providing no phase retardation. The half-wave plate ($\lambda/2$) rotates the polarization 90° (now s-polarized, out of the plane of the paper). The s-polarized pulse reflects off of the high reflector and enters the vacuum chamber through a window. The laser beam is directed to the IR in the manner described in Section B.8.1 and directed out of the IR through a second window, before striking the original polarizer (TFP 1). The s-polarized pulse now reflects off of the polarizer and passes on its original path through the Pockels cell. However, the Pockels cell is now switched to half-wave voltage providing a 180° phase retardation which rotates the polarization back to the plane of the paper (p-polarization). The pulse is now trapped in the cavity and retraces its original path. Beam quality (focusabilty) is maintained by constructing the cavity either as a relay-imaged ring or as a TEM_{00} resonator.

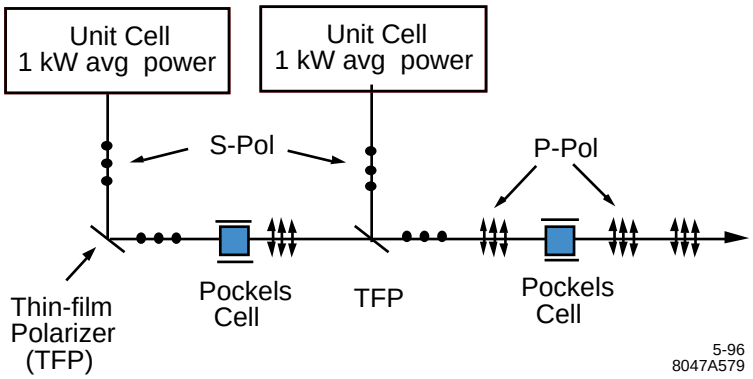


Figure B-31. *Pulse stacking/combination from individual unit cells onto a single optical axis is accomplished with electro-optic polarization switching.*

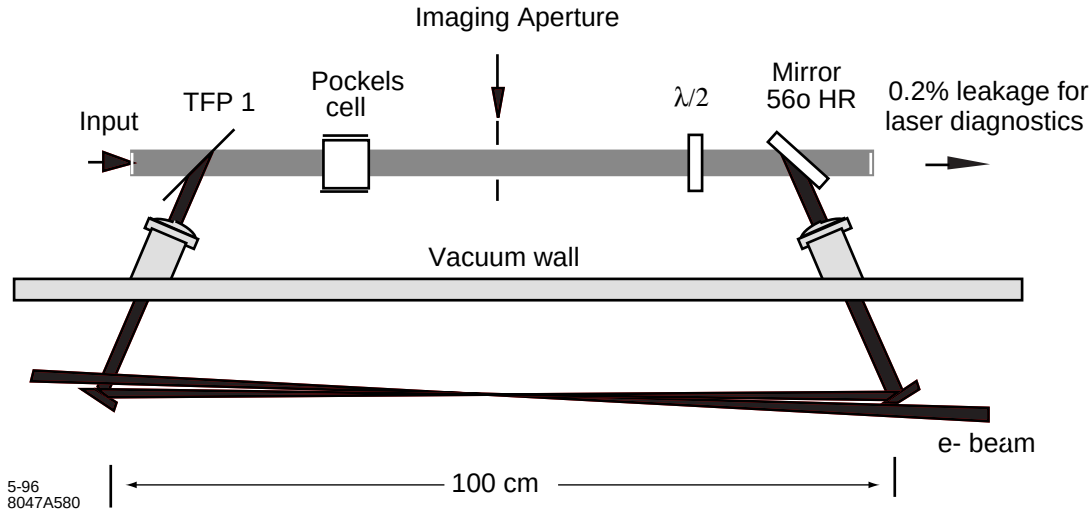


Figure B-32. *Relay imaged ring regenerative cavity.*

The ring is constructed such that its cavity length exactly matches a multiple of the separation between microbunches of the electron beam. In this case, each time the laser pulse traverses the ring, it strikes a new electron microbunch at the IP. This effectively multiplies the repetition rate of the laser by the number of passes around the ring. A typical cavity round-trip time is 10 ns corresponding to a 3-m cavity. However, this can easily be adjusted to match the optimum electron bunch spacing as dictated by the rf accelerator. We have constructed rings for a round-trip time as short as 3 ns to over 30 ns.

The pulse will slowly decay in energy with each pass of the ring (ring down). The rate of decay is determined by the optical quality and reflectivity of the cavity optics. We have achieved a net cavity loss as low as 4% in a ring cavity of a design similar to Figure B-32. With extremely high quality optics as are commonly used in the Atomic Vapor Laser Isotope Separation (AVLIS) program, we should be able to achieve a cavity loss lower than 2%. With a 2% loss, 80 round trips drops the pulse energy to 20% of the original input. Even with only moderate quality optics, we routinely achieve 70 round trips in existing systems.

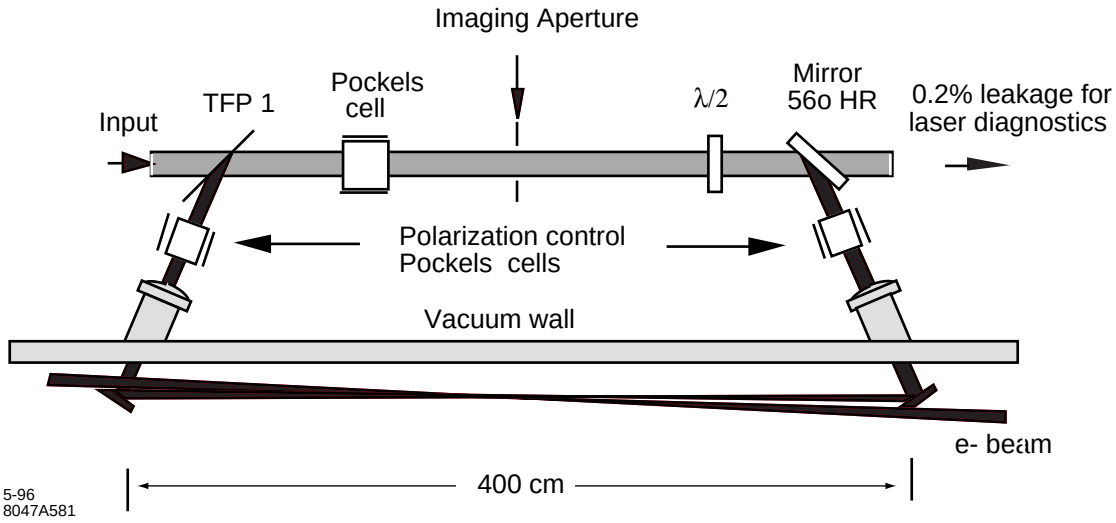


Figure B-33. *Laser photon recirculator and polarization control at the interaction point.*

For the picosecond pulses envisioned for the $\gamma\gamma$ collider, self-focusing and self-phase modulation will limit the number of round-trips achievable for the standard design of Figure B-32. These problems can be overcome by incorporating a passive pulse stretcher/compressor into the ring and/or using all reflective focusing of the beam to the interaction point. Optimization of these designs and the effect on system performance is an important development task.

B.11.9 Polarization Control at the Interaction Point

Helicity control is a unique requirement of the laser system for the $\gamma\gamma$ collider. The desire to change the polarization from linear to circular and back on either a macrobunch or several macrobunch time scale in order to investigate the helicity dependence of various $\gamma\gamma$ reactions is a complicating factor to the laser design. The strong polarization dependence of the laser amplifier section, pulse compressor and pulse stacker demand that all helicity manipulation be performed after the pulse has left these elements. The use of the photon recirculator of Figure B-32 is amenable to polarization control by addition of two additional Pockels cells as shown in Figure B-33. Before entering the vacuum chamber, the pulse passes through a Pockels cell which is either held at ground to produce no phase retardation and leaves the polarization unchanged (linearly polarized) or is switched to quarter-wave voltage to produce circularly polarized light. This Pockels cell can be easily switched at a 5-ms repetition rate (180 Hz) to enable switching between linear and circular polarization at the interaction region on alternating macropulses.

The ring scheme for reusing the laser beam and polarization control requires an extensive R&D. It should be demonstrated that there is the necessary space around the detector and that the B-integral in the Pockels cells, or the large cross-sectional areas of the cells, does not pose a problem.

These laser concepts should be considered preliminary at this time. Substantial conceptual design and optimization is yet to be done. A schematic of the conceptual design organization and issues to be addressed is shown in Figure B-34.

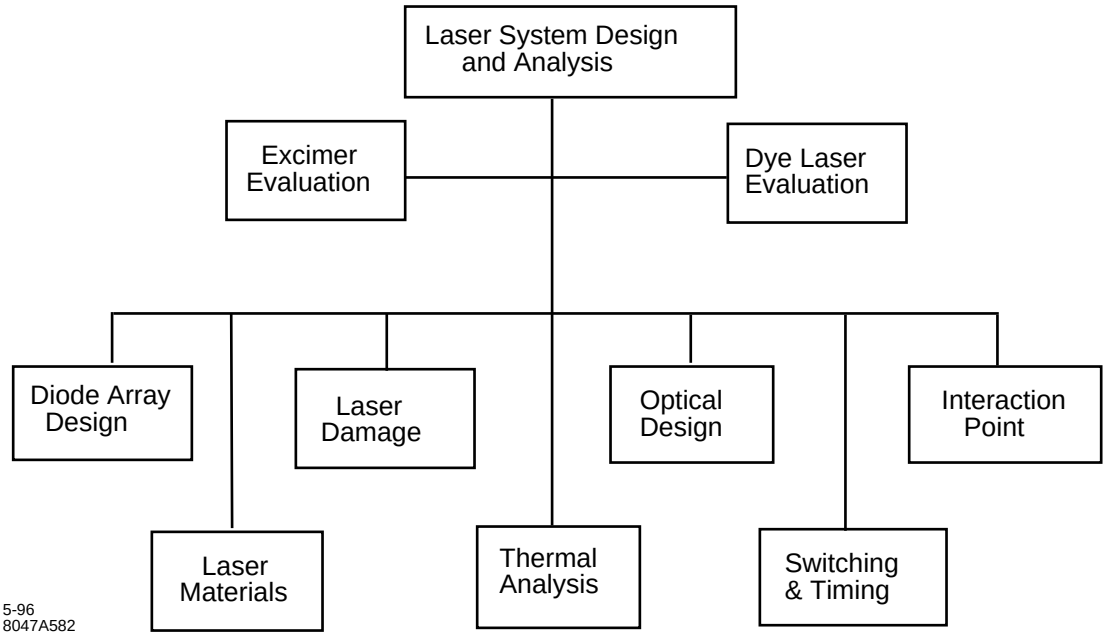


Figure B-34. *Conceptual design organization for NLC $\gamma\gamma$ collider laser system.*

B.11.10 Conclusion

In summary, many of the technological advances required for the NLC $\gamma\gamma$ option have recently been achieved. The $\gamma\gamma$ collider portion of the NLC benefits substantially from the large national efforts which are devoted to the development of high peak power and high average power laser systems. While substantial design, optimization and development still needs to be done, our preliminary study suggests that within the next few years, many of the required laser components will be demonstrated and a prototype NLC laser module could be developed.

B.12 Free-Electron Lasers

Free-electron lasers (FEL) are another option for photon colliders, and they are especially interesting for higher energy colliders, where the required wavelength of the laser is longer than $\sim 1\ \mu$m, for which solid-state lasers do not presently exist. Several schemes have been proposed based on different combinations of FEL oscillators, amplifiers, and optical switching techniques. A scheme based on the chirped pulse amplification and compression, similar to the technique used in solid state lasers but replacing the amplifier with an FEL driven by an induction linac, is another attractive option.

Electron Beam Parameters:	
I	1 kA
E	100 MeV
$\Delta E/E$	10^{-3}
Rms normalized emittance (ϵ_N)	50×10^{-6} m-rad
Betatron wavelength (λ_β)	11.3 m
Undulator Parameters:	
λ_W	4.0 cm
K	1.4
Length of uniform section	7.2 m
Length of tapered section	10 m
Micropulse FEL Power:	
Input power	100 kW
Power after uniform section	140 MW(power gain length = 1 m)
Power after tapered section	1.6 GW
Energy per micropulse	$1.6\,\text{GW} \times 1.4\,\text{ns} = 2.24\,\text{J}$

Table B-11. *Parameters for an FEL Pulse Compression Scheme.*

B.12.1 An FEL Scheme Using Induction Linac and Chirped Pulse Amplification Technique

With the usual high-gain FEL amplifiers, it is difficult to produce the laser pulses of the characteristics outlined in this section. This can be readily understood if we note that the saturation power in high-gain FELs is given approximately by $P_{\text{sat}} \sim \rho P_{\text{beam}}$ where ρ is the FEL scaling parameter [Bonifacio 1984] and P_{beam} is the electron beam power; $P_{\text{beam}} = EI$, where E and I are the beam energy and current respectively. Suppose we require the pulse energy $A = P_{\text{sat}}\tau$ to be about 1 J. Assuming $\rho \sim 10^{-2}$, $I \sim 1$ kA, and $\tau = 2$ ps, we find that E needs to be about 50 GeV. On the other hand, producing $\lambda = 1$ μmFEL with such a high-energy electron beam requires a strong and long (100-m) wiggler magnet.

The discussion above also points to the solution of the problem. Namely if the pulse length were much larger, about 1 ns, then the required electron beam energy becomes $E = 100$ MeV, which is quite reasonable for a 1-μm FEL. Thus the solution is to amplify the 1-ns pulse and later compress it to a few ps. A laser pulse can be compressed if it is chirped. Thus we are led to the idea of employing the chirped pulse amplification technique to FEL [Telnov 1991] extensively developed for solid-state lasers [Perry 1994a].

The scheme is schematically illustrated in Figure B-35. A solid-state laser produces a sequence of 1.8-ps, 0.14-mJ micropulses, with the same time structure as the collider beam, with an average power of 2.3 W. The micropulses are stretched (and hence chirped) to slightly less than 1.4 ns by means of a dispersive element schematically represented by a grating pair in the figure. The resulting optical beam which becomes amplified in a high-gain FEL driven by an induction linac producing $1.4 \times 90 = 130$-ns-long electron pulses at a 180-Hz repetition rate. The energy of each induction linac pulse may be modulated to match the chirped optical beam as indicated. Each amplified micropulse contains a few Joules of energy. The micropulses are compressed by another dispersive element to 1.8 ps. Of course, these optical components must designed to withstand high peak power (as is already true in SS laser compression systems) and high average power.

The electron beam and wiggler parameters required are shown in Table B-11.

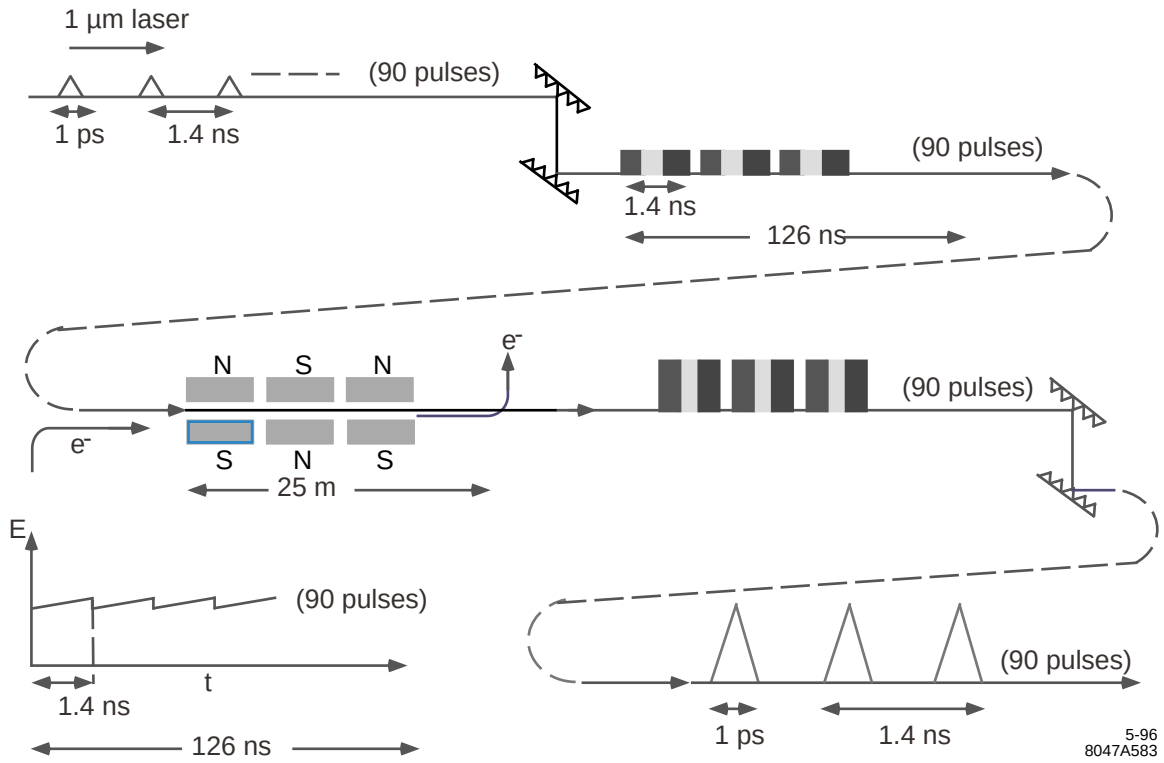


Figure B-35. *FEL pulse stretching/amplification/compressing scheme.*

A design of the induction linac producing the required electron beams is challenging but within the current state-of-the-art, and will be discussed in Section 9.2.3. The FEL consists of two sections, a 7.2-m-long uniform section in which the power gain length is about 1 m, and the input peak power of 100 kW is amplified to 140 MW. It is followed by a 7.2-m-long tapered section to further amplify the power to 1.6 GW. Thus, the energy contained in each micropulse is $1.6\,\mathrm{GW} \times 1.4\,\mathrm{ns} \simeq 2.24\,\mathrm{J}$, which is larger than 1 J required for a conversion efficiency of 68%.

A scheme to amplify chirped pulses, similar to the one discussed here, but based on a regenerative FEL amplifier driven by an rf linac, was independently proposed recently by an LANL group [Chan 1995]. The scheme requires the use of intracavity optical switches operating in a high-power environment and focusing mirrors operating at a peak power density of $1\,\mathrm{GW/cm^2}$.

B.12.2 Chirping Requirement and Tolerance

The compression of the input pulse with electric field amplitude $E_i(t)$ to the output amplitude $E_f(t)$ can be represented by the linear transformation

$$\begin{aligned} E_f(t) &= \frac{1}{\sqrt{2\pi}} \int G(t-t') E_i(t') dt', \\ G(t) &= \frac{1}{\sqrt{\tilde{\mu}}} \exp\left(\frac{i t^2 \tilde{\mu}}{2} - i\omega_0 t\right) \end{aligned} \tag{B.34}$$

where $1/\tilde{\mu}$ is the time delay per unit frequency interval introduced by the dispersive element. An input chirped pulse may be represented by

$$E_i(t) = \exp\left(-i\left[\omega_0 t + \frac{\mu}{2}t^2\right] - \frac{t^2}{4\sigma_{\tau i}^2}\right) \quad , \tag{B.35}$$

where $\sigma_{\tau i}$ is the rms pulse length, and μ is the frequency change per unit time interval (chirping rate). Inserting Eq. B.35 into Eq. B.34, we find that the output pulse is also Gaussian, with the rms pulse length

$$\sigma_{\tau f} = \sqrt{\frac{1}{4\sigma_{\tau i}^2\tilde{\mu}^2} + \sigma_{\tau i}^2\left(\frac{\mu}{\tilde{\mu}} - 1\right)^2}\,. \tag{B.36}$$

Thus, if the grating is designed so that $\tilde{\mu} = \mu$, then $\sigma_{\tau f} = 1/2\mu\sigma_{\tau i}$. If $\sigma_{\tau i} = 1.4\,$nsand $\sigma_{\tau f} = 1.8\,$ps, then we require $\tilde{\mu} = 2 \times 10^{20}/\mathrm{s}^2$. The total frequency chirping over the uncompressed pulse is $\Delta\omega \equiv= 2\mu\sigma_{\tau i} = 5.6 \times 10^{11}/$ or for $1\,\mu$m radiation $\Delta\omega/\omega = 2.8 \times 10^{-4}$. This is well within the gain bandwidth of the FEL. Therefore, it may not be necessary to modulate the electron beam energy.

There may be jitter introduced by the FEL amplification process. The jitter must be small so that $\sigma_{\tau i}^2(\mu/\tilde{\mu} - 1)^2 \ll 1/4\sigma_{\tau i}^2\tilde{\mu}^2$ or $\mu - \tilde{\mu} \ll 1/2\sigma_{\tau i}^2$. Thus the frequency error over the pulse must satisfy $\Delta\omega/\omega \ll 1/\omega\sigma_{\tau i} = 3.8 \times 10^{-7}$. To see the implication of the above requirement, we consider a high gain FEL in the exponential growth regime. The phase error due to a fluctuation $\Delta\rho$ in ρ is $\Delta\phi \sim 2\pi\Delta\rho L_W/\lambda_W$, where L_W is the length of the wiggler. Since ρ is proportional to $I^{1/3}/\gamma$, we have $\Delta\rho/\rho \sim (\Delta I/3I - \Delta\gamma/\gamma)$. Hence, $\Delta\omega/\omega = \Delta\phi/\omega\sigma_{\tau i} = (\lambda/c\sigma_{\tau i})(L_W/\lambda_W)\rho(\Delta I/3I - \Delta\gamma/\gamma)$. Taking $L_W = 20$m, $\lambda_W = 4\,$cm, $\lambda = 1\,\mu$m, $\rho = 2 \times 10^{-3}$, we obtain $\Delta\omega/\omega \sim 10^{-5}(\delta I/3I - \Delta\gamma/\gamma)$. Therefore, if the fluctuations in $\Delta I/I$ and $\Delta\gamma/\gamma$ are a few percent, then the compression requirement is satisfied. Such a tolerance should be achievable.

B.12.3 Induction Linac Driver

The $\gamma\gamma$ induction linac driver beam parameters required for driving the FEL are listed in Table B-12 with beam parameters of several other induction linacs. The ATA was used to drive the Paladin FEL [Orzechowski 1990], the ETA II was used to drive the Intense Microwave Prototype (IMP) FEL [Allen 1992], and the TBNLC is a proposed induction accelerator driven relativistic klystron for powering the NLC [Caporaso 1995]. The $\gamma\gamma$ Induction Linac Driver will require significant improvement in the areas of beam brightness (B_N) and energy flatness (ΔE/E) with respect to measured values for ATA or ETA II. Note that the electron beam of the last induction linac built at LLNL (ETA II) had a measured brightness 1/5 of that required. ETA II, however, had much better beam quality than the earlier ATA. Thus, although induction linacs built to date have not produced the beam quality required for the $\gamma\gamma$ induction linac driver, it is reasonable to expect that improved computer modeling of induction injectors will lead to designs with higher beam brightness meeting the FEL requirements for $\gamma\gamma$ colliders. The use of photocathodes in induction injectors could also be explored to achieve the required beam quality.

ATA and ETA II represent about 10-year-old induction accelerator technology. A program to construct a prototype induction linac for the TBNLC is in progress at LBNL. Simulations have shown that the induction linac injector can produce a B_N of about $1.5 \times 10^9\ A/m^2 - r^2$ for a 1.2-kA, 2.8-MeV beam. A demonstration of this brightness will be significant to the $\gamma\gamma$ induction linac driver as its injector represents the largest source of brightness degradation (emittance growth). Energy regulation to achieve a ΔE/E of $\pm$0.2% is planned for the prototype, and will determine the feasibility of a ΔE/E of $\pm$0.1% for the $\gamma\gamma$ driver.

Beam energy is primarily a cost issue. Induction modules can be added to provide the desired beam energy. However, as the length of the induction linac increases, transverse instability of the beam motion can become a difficulty. The

Parameter	$\gamma\gamma$ Induction Linac Driver	ATA (Paladin)	ETA II (IMP)	TBNLC Injector/main
E_b (MeV)	100	45	7	2.8/10[a]
I_b (kA)	1	0.8	2.4	1.2/0.6
ε_N(edge)	300	650(300 A)	900	400/800
(mm-mr)		1,116 (2.7 kA)		
$B_N(A/m^2 - r^2)$	2.2×10^9	1.4×10^8 (300 A) 4.6×10^6 (2.7 kA)	4×10^8 (1.5 kA) (2.5 MeV)	$1.5 \times 10^9/2 \times 10^8$
Flattop	150	>20	40	200
Pulse Length (ns)				
ΔE/E	$\pm$ 0.1%		$\pm$1%	$\pm$0.2%

[a] Final beam energy without rf power extraction would be 90 MeV.

Table B-12. *Comparison of the $\gamma\gamma$ Induction Linac Driver to other induction accelerators.*

beam tube radius of the induction cells can be increased to lower the growth in the transverse instability as described below. The larger inner radius will lead to a larger induction core volume increasing the core cost and reducing efficiency. The transverse instability mechanism is well understood for induction linacs [Capporaso]. Assuming constant acceleration ($\gamma = \gamma_0 + \lambda z$) and a solenoid focusing field proportional to the beam energy, a figure of merit for transverse instability is given by the product of the betatron phase advance (φ_β) times the number of e-folds (f_e) of gain:

$$\varphi_\beta f_e = \frac{2}{\lambda^2}\frac{(\omega_0 Z_\perp)}{L_g}\frac{I}{I_0}\left(\sqrt{\gamma_f} - \sqrt{\gamma_0}\right)^2 \quad , \tag{B.37}$$

where I is the beam current, I_0 = 17.03 kA, L_g is the separation between induction module gaps, $(\omega_0 Z_\perp)$ is the transverse impedance of the gaps, and λ is the gap energy increase ($\Delta\gamma$) divided by L_g. The transverse impedance scales approximately as:

$$(\omega_0 Z_\perp) = \eta\frac{4w}{b^2} \quad , \tag{B.38}$$

where w is the gap width, b is the inner radius, and η is a design factor of order unity ($\eta = 1.3$ for ATA).

Three e-folds of gain (f_e = 3) in an accelerator length of about 16 betatron wavelengths $(\varphi_\beta = 100)$ would be a reasonably conservative design for controlling transverse instability. Assuming a maximum field stress in the gap of 100 kV/cm ($\Delta\gamma$ = 0.196 and w = 1.0 cm), gap spacing of 30.0 cm, injector voltage of 1.5 MV, and η = 0.7 ("good" design), Eqs. B.37 and B.38 are used to find b $\geq$ 11.2 cm. The focusing field will increase from 183 G at 1.5 MeV to 1.3 kg at 100 MeV. The linac will have approximately 1,000 induction modules and be over 300 m in length.

Induction cells with 11-cm beam pipe radius will have about twice the core volume of the TBNLC cells (pipe radius = 5 cm). This volume increase will not only lead to added cost for core material, but also require an appreciable increase in the number of pulse power units than that of the TBNLC design. Although the design parameters have not been optimized, the $\gamma\gamma$ Induction Linac Driver is expected to cost significantly more per unit length than the TBNLC induction linac and have lower wall plug to beam power conversion efficiency.

References

[Allen 1992] S.L. Allen and E.T. Scharlemann, "The ETA II Linear Induction Accelerator and IM Wiggler: A High-Average-Power Millimeter-Wave Free-Electron-Laser for Plasma Heating", *Proc. 9th Intl. Conf. on High Power Part. Beams*, Washington, DC, (1992).

[Arutyunian 1963] F.R. Arutyunian and V.A. Tumanian, *Phys. Lett.* **4** 176 (1963); R. H. Milburn, *Phys. Rev. Lett.* **10** 75 (1963).

[Asaka 1995] T. Asaka, N. Maekawa, T. Moroi, Y. Shobuda, Y. Sumino, TU-483, May 1995, talk given at 5th Workshop on Japan Linear Collider (JLC), Tsukuba, Japan, 16-17 Feb 1995; e-Print Archive: hep-ph/9505371.

[Baillargeon 1995] M. Baillargeon, G. Belanger, and F.Boudjema, *Phys. Rev.* **D51**, 4712 (1995).

[Balakin 1994] V. Balakin, A. Sery, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355** 157 (1995).

[Bamber 1995] C. Bamber, T. Blalock, S. Boege, J. Kelly, T. Kotseroglou, A.C. Melissinos, D.D. Meyerhofer, W. Ragg, and M. Shoup III, *Opt. Lett.* submitted (1995).

[Barklow 1990] T. Barklow, SLAC–PUB–5364 (1990).

[Barletta 1993] W. Barletta *et al.*,, (Plasma Lens Collaboration), "Proposal for Plasma Lens Experiments at the Final Focus Test Beam", SLAC–Proposal–E–150 (April 1993).

[Beach 1994] R. Beach *et al.*, Conf. on Lasers and Electro-optics (CLEO), Anaheim, CA (1994).

[Beaudoin 1992] Y. Beaudoin, C.Y. Chien, J.S. Coe, J.L. Tapi, and G. Mourou, *Opt. Lett.* **17** 865–867 (1992).

[Ben-Menachem 1993] S. Ben-Menachem and P. Chen, SLAC–436, in *Proc. of the 5th Int. Workshop on Next-Generation Linear Colliders*, 426 (1993).

[Berestetskii 1982] V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, *Quantum Electrodynamics*, Section 101 (Pergamon Press 1982).

[Berkeley 1995] *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, 1994; *Nucl. Instr. Methods A* **355**, 1–194 (1995).

[Billhardt 1993] F. Billhardt, M. Kalashnikov, P.V. Nickles, and I. Will, *Opt. Commun.* **98**, 99–104 (1993).

[Bonifacio 1984] R. Bonifacio, C. Pellegrini and L.M. Narducci, *Opt. Commun.* **50**, 373 (1984).

[Borden 1993a] D.L. Borden, plenary talk presented at Workshop on Physics and Experiments with Linear e^+e^- Colliders, Waikoloa, Hawaii, 26–30 April 1993, ed. F.A. Harris, S.L. Olsen, S. Pakvasa, and X. Tata, (World Scientific, Singapore, 1993).

[Borden 1993b] D.L. Borden, D.A. Bauer, D.O. Caldwell, *Phys. Rev.* **D48**, 4018 (1993).

[Borden 1994] D.L. Borden, V.A. Khoze, W.J. Stirling, J. Ohnemus, *Phys. Rev.* **D50**, 4499 (1994).

[Boyd 1995] R.D. Boyd, J.A. Britten, D.E. Decker, B.W. Shore, B.C. Stuart, M.D. Perry, and L. Li, *Appl. Opt.* **34**, 1697–1706 (1995).

[Brodksy 1994] S. Brodsky, P. Zerwas, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 19 (1995).

[Bula 1992] C. Bula *et al.*, Study of QED at Critical Field Strength at SLAC, E-144 proposal (1992).

[Bula 1995] C. Bula *et al., Phys. Rev. Lett.* , submitted (1995).

[CAIN] CAIN (Conglomérat d' ABEL et d' Interactions Non-Linéaires) being developed by P. Chen, G. Horton-Smith, T. Ohgaki, A. Spitkovsky, A.W. Weidemann, and K. Yokoya.

[Caporaso 1995] "Relativistic-Klystron Two-Beam Accelerator Based Power Source for a 1-TeV Center-of-Mass Next Linear Collider: Preliminary Design Reports", LBID–2085 and UCRL–ID–119906, (Feb. 1995); G.J. Caporaso and A.G. Cole, "High Current Electron Transport", in *AIP Conference Proceedings* 249, The Physics of Particle Accelerators, eds. M. Month and M. Diens, AIP, NY, 1662–1672.

[Capporaso] G.J. Caporaso and A.G. Cole, "High Current Electron Transport", in AIP Conference Proceedings 249, The Physics of Particle Accelerators, eds. M. Month and M. Diens, 1662–1672, (AIP, NY).

[Chan 1995] K.C.D. Chan, J.C. Goldstein, D.C. Nguyen, and H. Tekeda, "A Chirped Pulse Regenerative - Amplifier FEL for the Gamma-Gamma Collider", LANL preprint, LA–UR–95–1502, presented at the 1995 Part. Acc. Conf., Dallas, TX (1995).

[Chanowitz 1994] M. Chanowitz, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355** 42 (1995).

[Chen 1989] P. Chen, V. Telnov, *Phys. Rev. Lett.* **63** 1796 (1989).

[Chen 1995a] P. Chen, G. Horton-Smith, T. Ohgaki, A.W. Weidemann, K. Yokoya, *Nucl. Instr. and Methods* **A35** 107 (1995).

[Chen 1995b] P. Chen, D. Bullock and D. Yu, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355** 130 (1995).

[Cheung 1994] Kingman Cheung, *Phys. Rev.* **D50**, 4290 (1994).

[Chien 1995] C.Y. Chien, G. Korn, J.S. Coe, J. Squier, G. Mourou, and R.S. Craxton, *Opt. Lett.* 20, 353-355 (1995).

[Chuang 1991] Y.-H. Chuang, D.D. Meyerhofer, S. Augst, H. Chen, J. Peatross, and S. Uchida, *J. Opt. Soc. Am.* B 8, 1226–1235 (1991).

[Chuang 1993] Y.-H. Chuang, L. Zheng, and D.D. Meyerhofer, *IEEE J. Quantum Electron.* **29** 270–280 (1993).

[Clayton 1994] C. Clayton, N. Kurnit, and D. Meyerhofer, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994).

[Clayton 1995] C.E. Clayton, N.A. Kurnit, and D.D. Meyerhofer, *Nucl. Instr. and Methods* **A355**, 121–129 (1995).

[Corkum 1988] P.B. Corkum, F. Brunel, N.K. Sherman, and T. Srinivasan-Rao, *Phys. Rev. Lett.* **61**, 2886 (1988).

[Ditmire 1993] T. Ditmire and M.D. Perry, *Opt. Lett.* **18** 426 (1993).

[Eboli 1993] O.J.P. Eboli, M.C. Gonzalez-Garcia, F. Halzen, D. Zeppenfeld, *Phys. Rev.* **D48**, 1430 (1993); O.J.P. Eboli, talk given at Workshop on Gamma-Gamma Colliders, Lawrence Berkeley Laboratory, Berkeley, CA, USA (March 28–31, 1994).

[Fawley 1980] W.M. Fawley and E.P. Lee, "Modelling of Beam Focusing and Kink Instability for Colliding Relativistic Electron and Positron Beams", LLNL report *UCID–18584*, (1980).

[Fawley 1987] W.M. Fawley and E.P. Lee, E.P. "Modelling of Beam Focusing and Kink Instability Particle in Cell Simulations of Disruption", in *Proc. New Developments in Par. Accel. Techniques*, ed. S. Turner, **CERN 870-11**, **ECFA 87/110**, 605–609 (1987).

[Ferray 1990] M. Ferray, L.A. Lompre, O. Gobert, A. L'Huillier, G. Mainfray, C. Manus, A. Sanchez, and A.S. Gomes, *Opt. Commun.* **75**, 278 (1990).

[Ginzburg 1981] I. Ginzburg, G. Kotkin, V. Serbo, V. Telnov, *Pizma ZhETF* **34** 514 (1981); *JETP Lett.* **34** 491 (1982); *Prep. INF* 81–50, Novosibirsk, in English (Feb. 1981).

[Ginzburg 1983] I. Ginzburg, G. Kotkin, V. Serbo, V. Telnov, *Nucl. Instr. and Methods* **205**, 47 (1983); *Prep. INP* 81–92, (Novosibirsk, Aug. 1981).

[Ginzburg 1984] I. Ginzburg, G. Kotkin, S. Panfil, V. Serbo, V. Telnov, *Nucl. Instr. and Methods* **219**, 5 (1984).

[Ginzburg 1994] I. Ginzburg, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 63 (1995).

[Gounaris 1995] G.J. Gounaris, J. Layssac, F.M. Renard, PM–95–11, May 1995. 23 pp.; e-Print Archive: hep-ph/9505430

[Grzadkowski 1992] B. Grzadkowski and J.F. Gunion, *Phys. Lett.* **B294**, 361 (1992).

[Gunion 1993] J.F. Gunion and H.E. Haber, *Phys. Rev.* **D48**, 5109 (1993).

[Gunion 1995] J.F. Gunion, J.G. Kelly, J. Ohnemus, *Phys. Rev.* **D51**, 2101 (1995).

[Heinrich 1991] J.G. Heinrich, C. Lu, K.T. McDonald, C. Bamber, A.C. Melissinos, D. Meyerhofer, Y. Semertzidis, Pisin Chen, J.E. Spencer, R.B. Palmer, "Proposal for a Study of QED at Critical Field Strengths in Intense Laser–High-Energy Electron Collisions", SLAC–Proposal–E–144 (October 1991).

[Heinz 1989] P. Heinz and A. Laubereau, *J. Opt. Soc. Am.* **B 6**, 1574 (1989).

[Hiramatsu 1995] S. Hiramatsu, S. Hashimoto and Y. Ishida, *Proc. of Workshop on Gamma-Gamma Colliders,* Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 133 (1995).

[Hollebeek 1981] R. Hollebeek, "Disruption Limits for Linear Colliders", *Nucl. Instr. and Methods* **184**, 333 (1981).

[Horton-Smith phd] G. Horton-Smith, Ph.D. Thesis, SLAC.

[Injean 1994] H. Injean *et al.*, "Diode array, kilowatt laser development", Conference on Lasers and Electro-Optics (CLEO), pp. CThC1 (Anaheim, CA, 1994).

[Jikia 1993] G.V. Jikia, *Phys. Lett.* **B298**, 224 (1993); *Nucl. Phys.* **B405**, 24 (1993).

[Jikia 1994a] G. Jikia and A. Tkabladze, invited talk at Workshop on Gamma-Gamma Colliders, Berkeley, CA, 28-31 (Mar 1994).

[Jikia 1994b] G. Jikia, IFVE-94-77, 34 pp. (Jul 1994); e-Print Archive: hep-ph/9407393.

[Kamal 1995] B. Kamal, Z. Merebashvili, A.P. Contogouris, *Phys. Rev.* **D51**, 4808 (1995).

[Kim un] K.-J. Kim, A. Sessler, M. Xie, report al LC95, to be published.

[Kim 1994] Prepared by K.-J. Kim, P. Pierini, A. Sessler, V. Telnov at LC95, KEK, Japan, (March 1994).

[Kmetec 1991] J.D. Kmetec, J.J. Macklin, and J.F. Young, *Opt. Lett.* **16**, 1001–1003 (1991).

[Kon 1992a] T. Kon and A. Goto, *Phys. Lett.* **B295**, 324 (1992); F. Cuypers, G.J. van Oldenborgh, R. Rückl, *Nucl. Phys.* **B383**, 45 (1992).

[Kon 1992b] T. Kon, I. Ito, and Y. Chikashige, *Phys. Lett.* **B287**, 277 (1992).

[Kon 1993] T. Kon, *Phys. Lett.* **B316**, 181 (1993); F. Cuypers, G.J. van Oldenborgh, R. Rückl, *Nucl. Phys.* **B409**, 144 (1993).

[Kondratenko 1983] A. Kondratenko, E. Pakhtusova, E. Saldin, *Dokl. Akad. Nauk* **264**, 849 (1982).

[Kramer 1994] M. Kramer, J. Kuhn, M.L. Stong, and P.M. Zerwas, *Z. Phys.* **C64**, 21 (1994).

[Landau 1987] L.D. Landau and E.M. Lifshitz, "The Classical Theory of Field", Course of Theoretical Physics 2, (Pergamon Press, 1987).

[Martinez 1987] O.E. Martinez, *IEEE J. Quantum Electron.* **QE-23**, 59–64 (1987).

[Meyerhofer 1995a] D. Meyerhofer, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 113–120 (1995).

[Meyerhofer 1995b] D.D. Meyerhofer, *Nucl. Instr. and Methods* **A355** 113-120 (1995).

[Miller 1995] D.J. Miller, *Nucl. Instr. and Methods* **A355**, 101 (1995).

[Morton 1995] P. Morton and S. Chattopadhyay, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 138 (1995).

[Murayama 1994] H. Murayama, talk at Workshop on Gamma-Gamma Colliders, Berkeley, CA (28-31 Mar 1994).

[Norem 1995] J. Norem, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 166 (1995).

[Ohgaki 1995] T. Ohgaki and X. Yokoya, to be published.

[Orzechowski 1990] Reference on Paladin–T.J. Orzechowski at 1990 USPA School.

[Perry 1990a] M. Perry and F. Patterson, *Opt. Lett.* **15**, 381 (1990).

[Perry 1990b] M.D. Perry and F.G. Peterson, LLNL Internal Documents (1990); M.D. Perry, T. Ditmire, and D. Strickland, *Opt. Lett.* **17**, 601 (1992).

[Perry 1991] M.D. Perry and R. Olson, *Laser Focus World* bf 27, 69-74 (1991).

[Perry 1994a] For a review, see M. Perry and G. Mourou, M.D. Perry and G. Mourou, *Science* **264**, 917-924 (1994).

[Perry 1994b] M.D. Perry, T. Ditmire, and B.C. Stuart, *Opt. Lett.* **19**, 2149 (1994).

[Perry 1995] M.D. Perry, R.D. Boyd, J.A. Britten, D. Decker, B.W. Shore, C. Shannon, and E. Shults, *Opt. Lett.* **20**, 940–942 (1995).

[Perry 1996] M.D. Perry, J.A. Britten, R.D. Boyd, H. Nguyen, B.W. Shore, U.S.Patent Pending.

[Raidal 1995] M. Raidal, *Nucl. Phys.* **B441**, 49 (1995).

[Rajagopalan 1994] S. Rajagopalan, D. Cline and P. Chen, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 169 (1995).

[Rajagopalan 1996] S. Rajagopalan, D.B. Cline, P. Chen, "Application of a Plasma Lens to $e^-\gamma$ Colliders", these Proceedings.

[Richard 1995] F. Richard, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 92 (1995).

[Rodwell 1986] M.J.W. Rodwell, K.J. Weingarten, D.M. Bloom, T. Baer, and B.H. Kolner, *Opt. Lett.* **11**, 638-640 (1986).

[Rodwell 1989] M.J.W. Rodwell, D.M. Bloom, and K.J. Weingarten, *IEEE J. Quantum Electron.* **25**, 817–827 (1989).

[Ronan 1993] M. Ronan, *Proc. of Workshop on Phys. and Exper. with Linear Colliders*, Waikoloa, Hawaii, US (1993).

[Salin 1991] F. Salin, C. Rouyer, J. Squier, S. Coe, and G. Mourou, *Opt. Commun.* **84**, 67 (1991).

[Saldin 1995] E. Saldin *et al.*, Preprint JINR E-9-94-74, Dubna,1994, Submitted to Particle and Accelerators; *Nucl. Instr. Methods A* **355**, 171; DESY 94-243,1994 (1995).

[Siegman 1986] A.E. Siegman, *Lasers* (University Science, Mill Valley, Ca, 1986).

[Silvestrov un] G. Silvestrov, V. Telnov (unpublished).

[Squier 1991] J.A. Squier, F. Salin, J.S. Coe, P. Bado, and G.A. Mourou, *Opt. Lett.* **15**, 85 (1991).

[Strickland 1985] D. Strickland and G. Mourou, *Opt. Commun.* **56**, 219 (1985); P. Maine, D. Strickland, P. Bado, M. Pessot, and G. Mourou, *IEEE J. Quantum Electron.* **QE-24**, 398 (1988).

[Stuart 1995a] B.C. Stuart *et al.*, submitted to JOSAB, UCRL–JC–120, 225.

[Stuart 1995b] B. Stuart, S. Herman, and M.D. Perry, *IEEE J. Quantum Electron.* **31**, 528 (1995).

[Stuart 1995c] B. Stuart, M.D. Feit, M.D. Perry, A.M. Rubenchick, and B.W. Shore, *Phys. Rev. Lett.* **74**, 2248 (1995).

[Takahashi 1994] T. Takahashi, "Physics of e^+e^-, γe^- and $\gamma\gamma$ collisions at Linear Accelerators", *Proc. INS Workshop*, Tokyo, Japan, 93 (1994).

[Takahashi 1995] T. Takahashi, Hiroshima Univ. Preprint, HUPD–9526 (1995).

[Takahashi 1996] T. Takahashi, in preparation.

[Tandean 1995] J. Tandean, *Phys. Rev.* **D52**, 1398 (1995).

[Tauchi 1993] T. Tauchi, K. Yokoya, P. Chen, *Particle Acclerators* **41**, 29 (1993).

[Telnov 1990] V. Telnov, *Nucl. Instr. Methods A* **294**, 72 (1990).

[Telnov 1991] V. Telnov, private communication and *Proc. of Workshop on Phys. and Exper. with Linear Colliders*, Lapland, Finland, 551 (1991).

[Telnov 1993] V. Telnov, *Proc. of Workshop on Physics and Experiments with Linear Colliders*, Wokoloa, Hawaii (World Scientific, 1993).

[Telnov 1995] V. Telnov, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355**, 3 (1995).

[Telnov priv] V. Telnov, private communication.

[Treacy 1969] E.B. Treacy, *IEEE J. Quantum Electron.* **QE–5**, 454 (1969).

[Watanbe 1995] I. Watanabe, talk presented at Workshop on Physics and Experiments with Linear e^+e^- Colliders, Appi, Japan, (1995).

[White 1992] W.E. White, J. Hunter, L. Van Woerkum, T. Dimitre, and M.D. Perry, *Opt. Lett.* **17**, 1067 (1992).

[Xie 1995] M. Xie, K.-J. Kim, and A.M. Sessler, *Nucl. Instr. and Methods* **355**, 163 (1995).

[Yehudai 1990] E. Yehudai, *Phys. Rev.* **D41**, 33 (1990).

[Yehudai 1991] E. Yehudai, *Phys. Rev.* **D44**, 3434 (1991).

[Yokoya, 1986] K. Yokoya, KEK–Report 85–9, October 1985; also *Nucl. Instr. and Methods* **A251**, 1 (1986).

[Yokoya 1996] *User's Manual of CAIN 2.0*, (KEK publication, April, 1996).

Contributors

- Karl van Bibber
- Swapan Chattopadhyay
- Bill Fawley
- Dick Helm
- Tim Houck
- John Irwin
- Kwang-Je Kim
- Dan Klem
- David Meyerhofer
- Hitoshi Murayama
- Micheal Perry
- Mike Ronan
- Andy Sessler
- Jim Spencer
- Tohru Takahashi
- Valery Telnov
- Achim Weidemann
- Glen Westenskow
- Ming Xie
- Kaoru Yokoya
- Sasha Zholents

C

Ground Motion: Theory and Measurement

Contents

C.1 Introduction

One of the major challenges for the NLC will be to keep the beams stable pulse-to-pulse so the collision luminosity is not degraded. Because the beam sizes are so small, motions of the quadrupole magnets on the nanometer scale can cause significant orbit changes. Hence, the quadrupole vibration caused by the ambient ground motion is a concern as is vibration due to man-made sources, such as cooling systems.

The sensitivity to vibration depends on both the time scale and correlation of the motion. Motion slower than about 0.1 Hz will be heavily suppressed by the trajectory feedback systems in the NLC and so it is less of a concern. Motion that is random will generally be much more of a problem than motion which is correlated. For example, the tolerance on the vertical rms quadrupole motion in the linacs is 6 nm if the motion is random quad-to-quad, but orders of magnitude larger if it is correlated over distances longer than the betatron oscillation lengths of the beams.

At frequencies above 0.1 Hz, the ambient rms ground motion is generally a few hundred nm or more, but the motion is correlated over long distances. The exact nature of this correlation has a big effect on whether the resulting beam motion in an accelerator will be tolerable or not. In the past year at SLAC, we have developed both the theoretical understanding and the measurement capability to address this question for the NLC. In this Appendix, we present the results of this learning process, including measurements of both the spatial and spectral characteristics of vertical ground motion in the SLAC linac tunnel.

To begin, we develop the theoretical framework that is needed to interpret ground motion measurements and to predict its effect on beam motion. Next, we describe some of the instruments that are used to measure ground motion, including their resolutions. In the last section, we discuss the measurements that were made at SLAC, and present a model to describe the data that is based on the dynamics of wave motion in the ground. We then derive a general equation that will govern the ground-motion effect on luminosity in the NLC in the framework of this model. However, the implications for the NLC are described separately for each region of the machine in the chapters that discuss these regions.

C.2 Theory

In this section we derive a general equation that describes the effect of ground motion on the separation of the beams at the interaction point (IP) of a collider. Specifically, we will show that the average square vertical displacement of the left-incoming from the right-incoming beam centroid can be written as an integral

$$\left\langle \Delta y_{IP}^2 \right\rangle_t = \int_0^\infty P(\omega, k) G(k) F(\omega) \frac{dk}{2\pi} \frac{d\omega}{2\pi} \tag{C.1}$$

where $<>_t$ indicates the average over time, $P(\omega, k)$ is a function describing the relevant aspects of the ground motion, $G(k)$ is a function describing the response of the lattice, and $F(\omega)$ describes the behavior of the feedback system. The definition, meaning, and properties of each of these functions are described in the next sections.

C.2.1 Ground Motion Spectra and Related Functions

The ground displacement along the accelerator may be described by a function of two variables $y(t,s)$, where s denotes position along the beam line and t denotes time. To minimize any possible ambiguity, imagine that at a point to be designated as $s = 0$, a horizontal line is constructed parallel to the machine orientation and perpendicular to a radius through the origin of the earth. The variable s will indicate the distance from the origin measured along this straight line. For simplicity we assume the radius of the earth at the origin of this line is chosen so the average displacement $< y(t,0) >_t = 0$, and that $< y(t,s) >_t = 0$ for all s.

The Time Power Spectrum $P(\omega)$

A simple first question one might ask is "How large a ground displacement, at the point s, am I likely to experience in time τ starting from some unknown time t?" The answer can be given in a statistical sense as

$$\left\langle \Delta y(\tau;s)^2 \right\rangle = \left\langle [y(t+\tau,s) - y(t,s)]^2 \right\rangle \quad . \tag{C.2}$$

Assuming uniformity along the horizontal line, this should be independent of s, but for a local source this could conceivably be s-dependent. Squaring and expanding Eq. C.2:

$$\left\langle \Delta y(\tau;s)^2 \right\rangle = 2\left[\left\langle y(t,s)^2 \right\rangle_t - \left\langle y(t+\tau,s)y(t,s) \right\rangle_t \right] \quad . \tag{C.3}$$

The last term will be an important function for us, and we will define $P(\omega;s)$ as its Fourier transform, namely:

$$\left\langle y(t+\tau,s)y(t,s) \right\rangle_t = \int_0^\infty P(\omega;s)\cos(\omega\tau)\frac{d\omega}{2\pi} \quad . \tag{C.4}$$

The $\cos(\omega\tau)$ is appropriate because the left-hand side is unchanged by the substitution of $-\tau$ for τ (followed by the substitution of $t = t' + \tau$ and noting that the average over t' is the same as the average over t). Next we will argue that $P(\omega) \geq 0$ for all ω.

Suppose that $y(t,s)$ has been Fourier decomposed. The time average of two terms with unequal frequencies will be zero, leaving only the sum of Fourier terms times themselves shifted by the time τ. That time average gives a $\cos(\omega\tau)$ times the amplitude of the term squared. The $P(\omega;s)$ must therefore be a constant times the square of a Fourier amplitude, and hence always positive.

Setting $\tau = 0$, we have the relationship

$$\left\langle y^2(t,s) \right\rangle_t = \int_0^\infty P(\omega;s)\frac{d\omega}{2\pi} \quad . \tag{C.5}$$

The quantity on the left-hand side is clearly positive for all t. According to the above paragraph, we can identify $P(\omega;s)$ as the contribution at frequency ω to the time averaged square of $y(t,s)$.

We will now make these statements more rigorous. Assume that we are averaging over a large time interval $[0,T]$. In this time interval, the function $y(t,s)$, for any fixed t, can, by imposing periodic boundary conditions, be expanded in

a Fourier series

$$\begin{aligned} y(t,s) &= \sum_{n=1}^{\infty} \left[a_n^T(s)\sin(\omega_n t) + b_n^T(s)\cos(\omega_n t)\right] \\ &= \sum_{n=1}^{\infty} c_n^T(s)\sin\left(\omega_n t + \psi_n^T(s)\right) \qquad \text{(C.6)} \\ \text{where} \qquad \omega_n &= \tfrac{2\pi n}{T} \quad . \end{aligned}$$

Forming the product $y(t+\tau,s)y(t,s)$ and taking the time average (and using the orthogonality relationship of the sin and cos functions) we get

$$\langle y(t+\tau,s)y(t,s)\rangle = \sum_{n=1}^{\infty} \frac{1}{2} c_n^T(s)^2 \cos(\omega_n \tau) \quad . \qquad \text{(C.7)}$$

As T gets large, this can be converted to an integral by using the relationship $dn/d\omega_n = T/2\pi$:

$$\langle y(s,t+\tau)y(s,t)\rangle = \int_0^{\infty} \frac{1}{2} \lim_{T\to\infty} \left[c_\omega^T(s)^2 T\right]\cos(\omega\tau)\frac{d\omega}{2\pi} \qquad \text{(C.8)}$$

and we can identify

$$P(\omega,s) \equiv \frac{1}{2} \lim_{T\to\infty} \left[c_\omega^T(s)^2 T\right] \qquad \text{(C.9)}$$

which is clearly a positive function for all ω. Since the left-hand side of Eq. C.8 is well defined, the product $c_\omega^T(s)^2 T$ must be independent of T for large enough T. This is equivalent to saying, that as T doubles, and the number of states in a given frequency range doubles, it is the energy of the states in that region that remains constant. For this to happen, c_n, the amplitude of a state in the shorter time interval, must decrease as $\sqrt{2}$ in the doubled time interval. (The sum of the squares of two modes equals the square of the previously single mode.)

Thus $P(\omega;s)$ can be interpreted as the power spectrum of $y(t,s)$ at s. Since we will be concerned primarily with seismic sources in this appendix, we will drop the s argument. We note that this s independence implies that the $c_n(s)$ is also independent of s. Only the phase, $\psi_n(s)$, of the motion at this frequency can retain an s dependence.

Since the left-hand side of Eq. C.5 is well-defined, and by experience known to be finite, the integral of the right-hand side should exist. However the left-hand side is quite large, since the daily changes in the earth's radius at any given location are about ± 40 cm. In other words, though $P(\omega;s)$ is integrable, it is expected to be huge for small ω. We expect that these low-frequency motions are coherent, that is to say, they are predominantly composed of long spatial wavelength motions. For these motions, the beam lines move as a whole, and we are led intuitively to believe they will not be a problem. Also there are beam-based feedback systems that receive information on beam position at 180 or 120 Hz, that will suppress the influence of low-frequency motions.

However at 0.15 Hz there is a micro-seismic peak due to ocean waves, with an amplitude as large as a micrometer, much larger than the beam spot size at the IP. To determine the impact of such motions, we must determine the wavelength decomposition of the ground motion as well as the frequency decomposition.

The Wavelength Power Spectrum $P(k)$

Proceeding as in the previous section, one might ask: "At any given moment how much are the vertical placements at the point s and $s+\ell$ likely to differ?" Again, the answer can be given in a statistical sense as

$$\left\langle \Delta y(\ell;t)^2 \right\rangle = \left\langle \left[y(t,s+\ell) - y(t,s)\right]^2 \right\rangle_s \quad . \qquad \text{(C.10)}$$

Proceeding exactly as in the preceding paragraph would lead to

$$\langle y(t,s+\ell)y(t,s)\rangle_s = \int_0^\infty P(k;t)\cos(k\ell)\frac{dk}{2\pi} \tag{C.11}$$

where

$$P(k;t) \equiv \frac{1}{2}\lim_{L\to\infty}\left[c_k^L(t)^2 L\right] \tag{C.12}$$

when we analyze the motion at a fixed time t along an interval in s of length L, imposing periodic boundary conditions. $P(k;t)$ gives the wavelength power decomposition of $y(t,s)$ at time t. If the distance s is long enough, and external conditions are not changing with t, we would expect this quantity to be independent of t, at least for short intervals of time, and we drop the explicit time-dependence, $P(k)$.

As a practical matter, when measurements are made to determine $P(k)$, the average over all s is not taken. Two seismometers are placed at some s separated by a distance ℓ, and it is assumed that the average over s is equal to the average over t, for some time interval T. Thus in practice we assume

$$\langle y(t,s+\ell)y(t,s)\rangle_s = \langle y(t,s+\ell)y(t,s)\rangle_t \quad . \tag{C.13}$$

The 2D Power Spectrum $P(\omega,k)$

Because the feedback systems are operating in time (getting a piece of information at 180 or 120 Hz) the frequency of any disturbance will be of great importance. On the other hand the response of the lattice will depend on the wavelength of the disturbance, so the wavelength decomposition is of great importance. One is lead to the undeniable conclusion that one must have both, namely at each frequency one needs to know the wavelength decomposition of the motion at that frequency. One must find the answer to question, "Given two points separated by a distance ℓ how much does their relative vertical displacement change in time τ?" This quantity is

$$\left\langle \Delta y(\tau,\ell)^2 \right\rangle = \left\langle \left[\{y(t+\tau,s+\ell)-y(t+\tau,s)\}-\{y(t,s+\ell)-y(t,s)\}\right]^2 \right\rangle_{t,s} \quad . \tag{C.14}$$

Squaring the expression within the square bracket, there are two new terms: $< y(t+\tau,s+\ell)y(t,s) >_{t,s}$ and $< y(t+\tau,s)y(t,s+\ell) >_{t,s}$. The sum of these is a symmetric function of τ and ℓ, and as above, we can define its Fourier transform through

$$\begin{aligned} &\tfrac{1}{2}\left[\langle y(t+\tau,s+\ell)y(t,s)\rangle_{t,s} + \langle y(t+\tau,s)y(t,s+\ell)\rangle_{t,s}\right] \\ &\qquad = \int_0^\infty\int_0^\infty P(\omega,k)\cos(\omega\tau)\cos(k,\ell)\tfrac{dk}{2\pi}\tfrac{d\omega}{2\pi} \quad . \end{aligned} \tag{C.15}$$

As with P(ω) and P(k), P(ω,k) ≥ 0 for all ω and k, and it is also a power spectrum.

Assume that we are averaging over a large time interval $[0,T]$ and space interval $[0,L]$. In this time and space interval the function $y(t,s)$ can be expanded in a double Fourier series

$$\begin{aligned} y(s,t) = \quad & \textstyle\sum_{m=1}^\infty\sum_{n=1}^\infty \left[a_{n,m}^{T,L}\sin(\omega_n t)\sin(k_m s) + b_{n,m}^{T,L}\sin(\omega_n t)\cos(k_m s) + \right. \\ & \left. d_{n,m}^{T,L}\cos(\omega_n t)\sin(k_m s) + e_{n,m}^{T,L}\cos(\omega_n t)\cos(k_m s)\right] \\ & \text{where } \omega_n = \tfrac{2\pi n}{T} \text{ and } k_n = \tfrac{2\pi m}{L} \quad . \end{aligned} \tag{C.16}$$

Forming the products $y(t+\tau, s+\ell)y(t,s)$ and $y(t+\tau, s)y(t, s+\ell)$, and taking the time average (and using the orthogonality relationship of the sin and cos functions) we get

$$\begin{aligned} &\tfrac{1}{2}\left[\langle y(t+\tau, s+\ell)y(t,s)\rangle_{t,s} + \langle y(t+\tau, s)y(t, s+\ell)\rangle_{t,s}\right] \\ &\quad = \textstyle\sum_{n,m=1}^{\infty} \tfrac{1}{4} c_{n,m}^{T,L^2} \cos(\omega_n \tau)\cos(k_m \ell) \\ &\text{where } c_{n,m}^{T,L^2} = a_{n,m}^{T,L^2} + b_{n,m}^{T,L^2} + d_{n,m}^{T,L^2} + e_{n,m}^{T,L^2} \quad . \end{aligned} \tag{C.17}$$

As T and L get large this can be converted to an integral by using the relationships $dn/d\omega_n = T/2\pi$ and $dm/dk_m = L/2\pi$ and we can identify

$$P(\omega, k) = \frac{1}{4} \lim_{T,L \to \infty} \left[c_{\omega,k}^{T,L^2} TL\right] \quad . \tag{C.18}$$

By taking $\ell = 0$ or $\tau = 0$ in the defining expression for $P(\omega, k)$, we have the following relationships among the power spectrums so far defined:

$$P(k) = \int_0^\infty P(\omega, k)\, \frac{d\omega}{2\pi} \quad \text{and} \quad P(\omega) = \int_0^\infty P(\omega, k)\, \frac{dk}{2\pi} \quad . \tag{C.19}$$

The Measure $\mu(\omega, k)$

The spectrum $P(\omega)$ is observed to vary according to weather conditions and can change dramatically if there is an earthquake. Excluding earthquakes, for which we have little data, the spectrum $P(\omega)$ may change by factors of up to ten, however for any given ω the distribution of wavelengths at that frequency appears to be quite constant. This has led us to define a function $\mu(\omega, k)$:

$$\mu(\omega, k) = \frac{P(\omega, k)}{P(\omega)} \quad . \tag{C.20}$$

The function $\mu(\omega, k)$ is a probability measure in the variable k, since $\int_0^\infty \mu(\omega, k)\frac{dk}{2\pi} = 1$ for any ω. In words, $\mu(\omega, k)\, dk/2\pi$ gives the probability that a ground disturbance of frequency ω will contain wavelengths in the interval dk about wave number $k = 2\pi/\lambda$.

If at frequency ω one has simply a wave of wavelength $\lambda(\omega)$ traveling to the right or left along the direction s at velocity $v(\omega) = f\lambda = \omega/k$, then

$$\mu(\omega, k) = 2\pi\delta(k - k_0(\omega)) \quad \text{where} \quad k_0(\omega) = \frac{2\pi}{\lambda(\omega)} = \frac{\omega}{v(\omega)} \tag{C.21}$$

and $\delta(k - k_0)$ denotes the Dirac delta function.

If at frequency ω one has only waves of wavelength $\lambda(\omega)$ traveling back and forth at some distinct angle θ to the direction s, then the wavelength of the disturbance along s will be $\lambda_s = \lambda/\cos\theta$ and

$$\mu(\omega, k) = 2\pi\delta(k - k_0(\omega)\cos\theta) \quad \text{where} \quad k_0(\omega) = \frac{2\pi}{\lambda(\omega)} = \frac{\omega}{v(\omega)} \quad . \tag{C.22}$$

If at frequency ω one has only waves of wavelength $\lambda(\omega)$ traveling at all angles θ to the direction s, then there will be a distribution of wavelengths along s. There will be none with $k > k_0$. Since for any given angle the wave number along s is $k = k_0 \cos\theta$, and the distribution is presumed uniform in $0 < \theta < \pi/2$, the fraction at any θ in interval $d\theta$ will be

$$\frac{2}{\pi} d\theta = \frac{2}{\pi} dk \frac{d\theta}{dk} = \frac{2}{\pi} \frac{dk}{k_0 \sin\theta} = \frac{2}{\pi} \frac{dk}{\sqrt{k_0^2 - k^2}} \tag{C.23}$$

and we have

$$\mu(\omega, k) = \frac{4}{\sqrt{k_0^2(\omega) - k^2}} \quad \text{for} \quad k < k_0 \quad . \tag{C.24}$$

It is possible that at any frequency there could be contributions from several different types of waves, for example surface waves, compression waves or Raleigh waves. And there is the possibility for motion due to stochastic processes that would have an entirely different kind of behavior. For example, much discussed in the literature is a relationship called the ATL law [Baklakov 1991]. If this type of motion is present then we will show later that

$$\mu(\omega, k) = \frac{A(\omega)}{k^2} \quad \text{as} \quad k \to \infty \quad . \tag{C.25}$$

The main task of ground motion measurements will be the determination of $\mu(\omega, k)$.

The Ratio $R(\omega, \ell)$

$P(\omega)$ gives the spectral decomposition of the rms motion of a single point. According to our definitions above

$$\left\langle y^2 \right\rangle = \int P(\omega) \frac{d\omega}{2\pi} \quad . \tag{C.26}$$

The power spectrum of the difference measurement of two points as indicated in Eq. C.10 can be written

$$\begin{aligned} \left\langle \Delta y(\ell)^2 \right\rangle &= 2 \int P(k) \left[1 - \cos(k\ell)\right] \tfrac{dk}{2\pi} \\ &= 2 \int \int P(\omega, k) \left[1 - \cos(k\ell)\right] \tfrac{dk}{2\pi} \tfrac{d\omega}{2\pi} \\ &= 2 \int P(\omega) \int \mu(\omega, k) \left[1 - \cos(k\ell)\right] \tfrac{dk}{2\pi} \tfrac{d\omega}{2\pi} \\ &\equiv 2 \int P(\omega) R(\omega, \ell) \tfrac{d\omega}{2\pi} \end{aligned} \tag{C.27}$$

from which we can conclude that the power spectrum for the difference of two points separated by distance ℓ is related to the power spectrum of single point by the ratio $2\ R(\omega, \ell)$ with

$$R(\omega, \ell) = \int \mu(\omega, k) \left[1 - \cos(k\ell)\right] \frac{dk}{2\pi} \quad . \tag{C.28}$$

We are guaranteed that $R(\omega, \ell) \leq 2$.

$R(\omega, \ell)$ can be determined directly in experiments, so we must try to extract the measure $\mu(\omega, k)$ given the ratio $R(\omega, \ell)$ for many ℓ. If one had data for many ℓ one could simply take a Fourier transform to find $\mu(\omega, k)$. Since we only have data at a small number of ℓ values, we will proceed differently by concentrating on the asymptotic behavior of R.

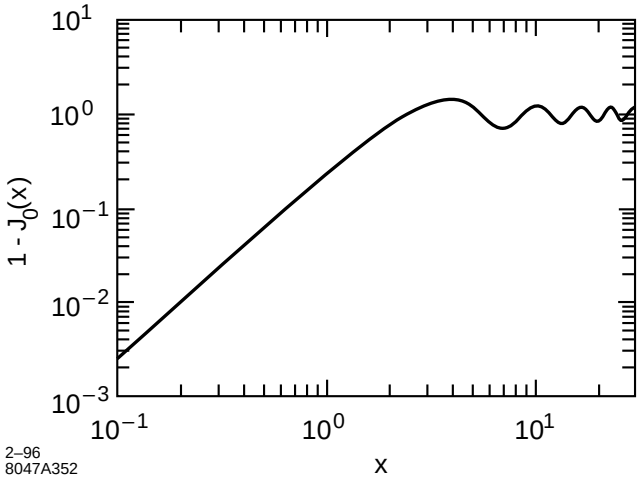


Figure C-1. *A log-log plot of the function $y = 1 - J_0(x)$. Note the asymptotic x^2 behavior for small x and the position $x = 2.4$ of the intercept at which $R = 1$ ($J_0 = 0$). The first maximum is about 1.44 at $x = 3.8$.*

The Ratio $R(\omega, \ell)$ as $\ell \to 0$

If the coherence of ground motion is to be helpful in suppressing its effect on beam motion, then we will be interested to know when $R(\omega, \ell)$ is small. This typically occurs as ℓ becomes small, so the asymptotic behavior for small ℓ is of special importance.

For the measure given by Eq. C.24, corresponding to waves of a single wavelength coming from all directions, then the integral of Eq. C.28 gives

$$R(\omega, \ell) = 1 - J_0\left(k_0(\omega)\ell\right) \tag{C.29}$$

where $J_0(k_0\ell)$ is the zeroth-order Bessel function. In this case, $1 - J_0(k_0 l) \to (k_0\ell/2)^2$ as $\ell \to 0$.

This asymptotic quadratic behavior is quite general for $R(\omega, \ell)$ as can be seen by looking at Eq. C.28. Suppose the measure μ has the property that

$$\overline{k^2(\omega)} \equiv \int k^2 \mu(\omega, k) \frac{dk}{2\pi} < \infty \quad . \tag{C.30}$$

Then $1 - \cos(k\ell)$ in the definition of R can be expanded in a power series with the result that

$$R(\omega, \ell) \to \frac{1}{2}\overline{k^2(\omega)}\ell^2 \quad \text{as} \quad \ell \to 0 \quad . \tag{C.31}$$

This predicts a clear l dependence for very general circumstances, and can be a strong consistency check on the data.

There can be other kinds of asymptotic behavior as $l \to 0$ if the second moment of μ does not exist. One such case of interest is when $\mu(\omega, k)$ has the form $A(\omega)/k^2$. In this case the condition of Eq. C.30 is not fulfilled, and the conclusion of Eq. C.31 is not valid. We will see in the next section that then $R(\omega, \ell)$ will be proportional to ℓ as $\ell \to 0$, which we refer to as the ATL law behavior.

Finding the ATL Law in Ground Motion Measurements

If at a given frequency ω there is a small ATL law component together with a dominant wave motion, then its signature will be that as ℓ becomes small one first observes $R(\omega,\ell) \to \alpha\ell^2$, and then this behavior shifts to $R(\omega,\ell) \to \beta\ell$. In other words, on a log-log plot of $R(\omega,\ell)$ as a function of ℓ one should observe a change from a slope of 2 to a slope of 1. If only the ℓ^2 behavior is seen, one can find an upper bound on the strength of the ATL coefficient. Let us suppose that this behavior is seen and for small ℓ. Let us take

$$R(\omega,\ell) = \alpha\ell^2 + \beta\ell \quad . \tag{C.32}$$

The intersection of the two straight-line behaviors will occur at $\ell = \ell_x$ with $\ell_x = \beta/\alpha$. The value of R at this intersection will be $R(\omega,\ell) = 2\beta^2/\alpha$, twice the value of the extension of either straight line. If no departure from an ℓ^2 behavior can be seen, we can conclude that $\ell_x(=\beta/\alpha) < \ell_{min}$, where ℓ_{min} is the smallest ℓ for which data is available. Hence $\beta < \alpha\ell_{min}$. We will use this relationship to establish upper bounds on the magnitude of the ATL law coefficient. Now we show how this asymptotic behavior is related to the measure $\mu(\omega,k)$.

Suppose that $\mu(\omega,k)$ has a $1/k^2$ tail. This is allowed because dk/k^2 is integrable as $k \to \infty$. Let us decompose the measure into two parts, one which contains the $1/k^2$ tail, and another which possesses a second moment. We write $\mu = \mu_R + \mu_T$ and take

$$\mu_T(\omega,k) = \frac{a(\omega)}{k^2}\left[1 - e^{-\frac{k^2}{2k_1^2}}\right] \tag{C.33}$$

where k_1 is a cut-off parameter.

The measure μ_T contains the $1/k^2$ tail and is integrable at both large and small k. The integral of the measure (integrate by parts) gives the fraction of the measure which is in the tail:

$$\mu_T(\omega) \equiv \int_0^\infty \mu_T(\omega,k)\frac{dk}{2\pi} = \frac{1}{2\sqrt{2\pi}}\frac{a(\omega)}{k_1(\omega)} \quad . \tag{C.34}$$

$\mu_R = \mu - \mu_T$ is supposed to have a finite second moment. Hence for small ℓ

$$\begin{aligned} R(\omega,\ell) &= \int(\mu_R+\mu_T)\left[1-\cos(k\ell)\right]\tfrac{dk}{2\pi} \\ &\to \tfrac{\ell^2}{2}\int \mu_R k^2 \tfrac{dk}{2\pi} + \int \mu_T\left[1-\cos(k\ell)\right]\tfrac{dk}{2\pi} \\ &= \tfrac{\ell^2}{2}\left\langle k^2\right\rangle_R + \tfrac{a(\omega)}{4}\ell - \tfrac{\ell^2}{4}\tfrac{a(\omega)k_1}{\sqrt{2\pi}} \\ &= \tfrac{\ell^2}{2}\left[\left\langle k^2\right\rangle_R - \mu_T(\omega)k_1^2\right] + \tfrac{a(\omega)}{4}\ell \quad . \end{aligned} \tag{C.35}$$

The coefficient A in the ATL law results in a spectral coefficient $4A/(k^2\omega^2)$. Hence the relation between A and $a(\omega)$ is given by

$$A = \omega^2 P(\omega)\frac{a(\omega)}{4} \quad . \tag{C.36}$$

Typically A is thought to be in the range of $10^{-4} < A < 10^{-6}\,\mu\mathrm{m}^2/\,\mathrm{m}/\,\mathrm{s}$. If there is no evidence of linear ℓ behavior, we can place a limit on A of $\omega^2 P(\omega)\beta < \omega^2 P(\omega)\alpha\ell_{\min} = \omega^2 P(\omega)\alpha\ell_{\min}$ where α is the coefficient of the ℓ^2 behavior and $\ell_{\min}$ is the smallest ℓ in the data set.

We note in closing that mathematically other behaviors at small ℓ are allowed, of the form ℓ^ϵ with $0 < \epsilon \leq 2$, corresponding to tail behaviors of $1/k^{1+\epsilon}$.

C.2.2 The Lattice Response Function $G(k)$

Sum Rules for Optical Functions

We will suppose that we have a lattice of elements with the quadrupole strength, $\kappa(s)$, given as a function of s. This function is defined by

$$\kappa(s) = \frac{B_T(s)}{a(s)(B\rho)} \tag{C.37}$$

where $B_T(s)$ is the quadrupole pole-tip field at s, and $a(s)$ is the radius of the pole tip. $B\rho = p/q$ is the particle momentum divided by the charge.

We will assume that there is some ideal line along which the magnets are arranged, and the displacement is given as a function of time by $y(t,s)$ with $< y(t,s) >_t = 0$. Then the displacement of the beam at the end of the beam line, which we will denote by $s = 0$, is

$$y_B(t) = \int_0^L \kappa(s) y(t,s) R_{34}(s) ds + R_{33}(L) y(t,L) \quad . \tag{C.38}$$

The assumption here is that the beam enters the beam line at a displacement equal to the ground level. It will be the function of the upstream beam line to carry out this task. Any failure to do this can be accounted for by calculating the deviation due to the upstream beam line and multiplying by $R_{33}(L)$. We will define a quantity $\Delta y_B(t)$ which is the difference $y_B(t) - y(t,0)$. Then

$$\Delta y_B(t) = \int_0^L \kappa(s) y(t,s) R_{34}(s) ds + R_{33}(L) y(t,L) - y(t,0) \quad . \tag{C.39}$$

If one translates the beam line by inserting a $y(t,s)$ which is independent of s, then $\Delta y_B(t)$ should be zero. According to Eq. C.39 this requires that

$$0 = \int_0^L \kappa(s) R_{34}(s) ds + R_{33}(L) - 1 \quad . \tag{C.40}$$

This is true, because $R_{34}(s)$ from $s \to 0$ is the same as the R_{34} from $0 \to s$. The latter is a sine-like ray coming from the origin, which we will denote by $\tilde{y}$ and has $\tilde{y}'(0) = -1$. Because of the equations of motion $\kappa(s)\tilde{y}(s) = \tilde{y}''(s)$, the integral in Eq. C.40 is $\tilde{y}'(L) - \tilde{y}'(0)$. The derivative of the R_{34} function coming from the origin is just -1 times the R_{44} function coming from the origin, which equals the -1 times the R_{33} function going from s to the origin. Hence Eq. C.40 is found to be valid, as it must.

If one inserts $y(t,s) = \alpha s$ into Eq. C.39, then $\Delta y_B(t)$ should again be zero if L is $n\pi$ from the origin. (If the phase change from 0 to L equals $n\pi$, the slope of the incoming beam does not affect the position at the IP). This implies that

$$0 = \int_0^L s\kappa(s) R_{34}(s) ds + L R_{33}(L) \quad . \tag{C.41}$$

Inserting $\tilde{y}(s)$ for R_{34}, substituting $\kappa(s)\tilde{y}(s) = \tilde{y}''(s)$ as above, and integrating by parts, and using the fact that $\tilde{y}(L) = 0$, one can confirm Eq. C.41.

These intuitively valid relationships would imply that it is not the magnitude of a displacement, nor even its slope that will cause beams not to collide, but it is the curvature of the earth's displacement that is important.

We have gone into some detail to describe and confirm these optical rules, because often approximations used for optical functions do not satisfy these conditions, and numerically incorrect results are obtained. A high degree of suppression of motion will only be obtained numerically when these sum rules are obeyed.

Paired Lattices

For lattices that have the property that for every $-I$ section there is an identical $-I$ section adjacent to it, which we designate as a "paired lattice," we have the result that if the displacement function has constant curvature the effects of the two adjacent sections will cancel one another. This leads to the conclusion that for paired lattices it is not the displacement, the slope, or the curvature, but the derivative of the curvature that causes a displacement of the beam.

Definition of $G(k)$

We will first write the beam displacement from the ground motion at the IP as a sum,

$$\Delta y_B(t) = \sum \mu_i y(t, s_i) \tag{C.42}$$

where the μ_i are dimensionless numbers equal to $\kappa(s_i)R_{34}(s_i)\Delta s_i$ or $R_{33}(L)$ or -1. Now we square this to get

$$\begin{aligned} < \Delta y_B(t)^2 > \quad &= \textstyle\sum_{i,j} \mu_i\mu_j < y(t,s_i)y(t,s_j) >_t \\ &= \textstyle\sum_{i,j} \mu_i\mu_j \int_0^\infty P(k)\cos\left[k(s_i - s_j)\right]\frac{dk}{2\pi} \\ &= \textstyle\int_0^\infty P(k)\frac{dk}{2\pi}\sum_{i,j}\mu_i\mu_j\cos\left[k(s_i-s_j)\right] \\ &= \textstyle\int_0^\infty P(k)\frac{dk}{2\pi}\left[\left(\sum_i \mu_i\cos ks_i\right)^2 + \left(\sum_i \mu_i \sin ks_i\right)^2\right] \\ &\equiv \textstyle\int_0^\infty P(k)G(k)\frac{dk}{2\pi} \quad . \end{aligned} \tag{C.43}$$

The lattice properties are completely contained in the sum, and therefore we have defined a lattice response function $G(k)$ by

$$G(k) = \left(\sum_i \mu_i \cos ks_i\right)^2 + \left(\sum_i \mu_i \sin ks_i\right)^2 \quad . \tag{C.44}$$

It is precisely Δy_B^2 for a cosine wave plus Δy_B^2 for a sine wave.

We note that in the special case were we are looking at the displacement of the two beams from one another at the IP, then if we assume that the NLC is symmetric, the cosine wave gives no displacement of the two beams, and the sine wave effect is doubled. Hence, when considering the effect on luminosity due to ground motion in both halves of a symmetric NLC, we have

$$G(k) = 4\left(\sum_i \mu_i \sin ks_i\right)^2 \tag{C.45}$$

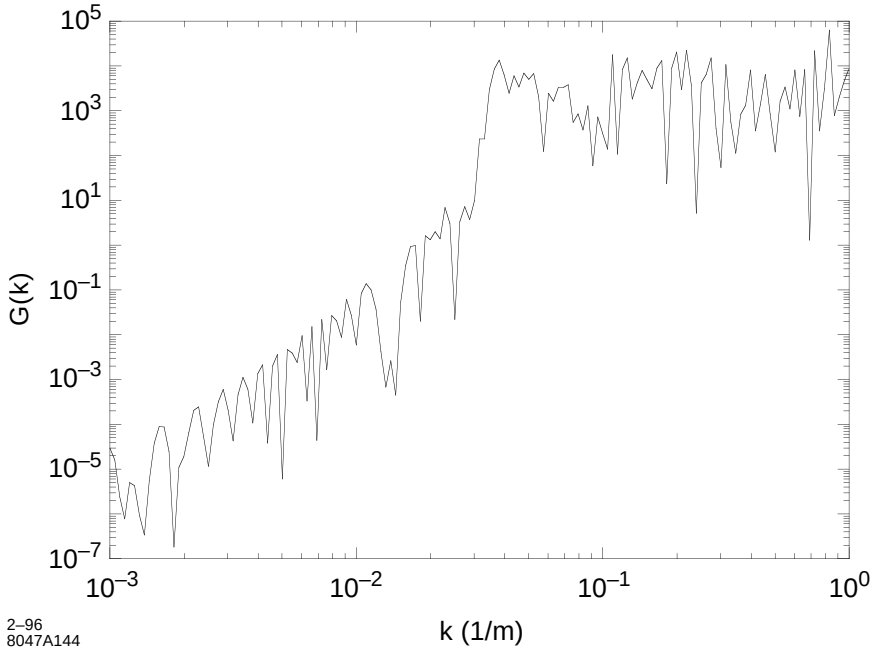


Figure C-2. *NLC-IIb linac lattice response function assuming symmetric linacs (i.e., Eq. C.45) that are separated by 10 km.*

where the sum is only over one half of the machine.

Behavior of $G(k)$ as $k \to 0$

As $k \to 0$, $\cos(ks_i) \to 1$. As shown above the constant term can give no net displacement. The next term is $1/2k^2s_i^2$. This term has a constant curvature, so for a paired lattice it will also be zero. The first term surviving from the cosine sum squared will go as k^8 in this case.

As $k \to 0$, $\sin(ks_i) \to ks_i$. As shown above the constant slope term also gives no net displacement. The first term surviving from the sine sum squared will go as k^6.

Some examples of $G(k)$ for various lattice sections are shown in Figures C-2, C-3, and C-4.

C.2.3 Combining Spatial and Frequency Dependence

$F(\omega)$

The particulars of a feedback response function will vary according to the exact feedback algorithm employed. In each case however there is a function $F(\omega)$ which will give the effect of the feedback for each input frequencies. As a result Eq. C.43 can be generalized, in the presence of feedback, to be

$$< \Delta y_B(t)^2 >_t = \int P(\omega, k) G(k) F(\omega) \frac{dk}{2\pi} \frac{d\omega}{2\pi} \quad . \tag{C.46}$$

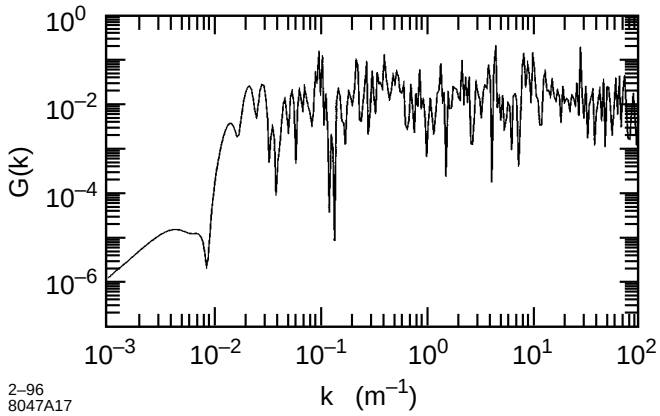


Figure C-3. *Collimation system lattice response function.*

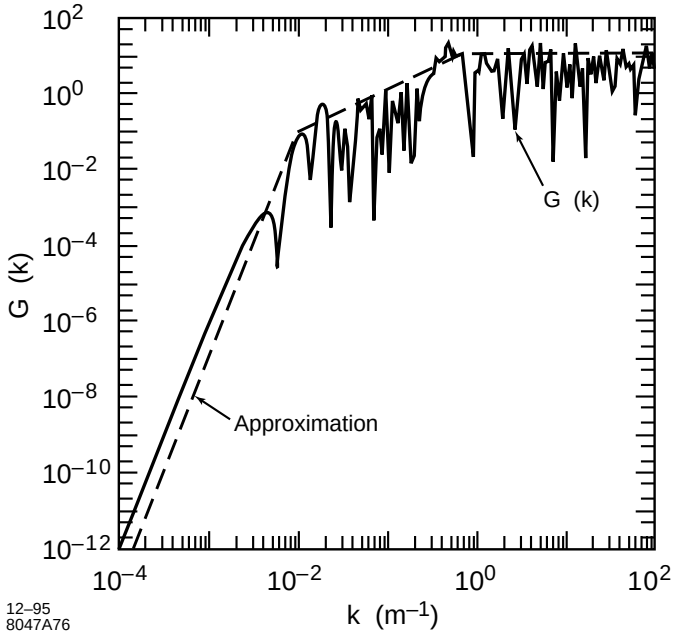


Figure C-4. *Final-focus lattice response function.*

$G_\mu(\omega)$

Once the measure $\mu(\omega,k)$ has been determined, $P(\omega,k)$ in Eq. C.46 can be replaced by $P(\omega,k) = P(\omega)\mu(\omega,k)$ and the integral over k in Eq. C.46 can be carried out. The result of doing the k integral we will denote by

$$G_\mu(\omega) \equiv \int_0^\infty \mu(\omega,k)G(k)\frac{dk}{2\pi} \quad . \tag{C.47}$$

This gives the response of the lattice to the combination of wavelengths present in the ground motion at the frequency ω. Doing this integral first, Eq. C.46 becomes

$$\left\langle \Delta y_B(t)^2 \right\rangle_t = \int_0^\infty P(\omega)G_\mu(\omega)F(\omega)\frac{d\omega}{2\pi} \quad . \tag{C.48}$$

C.3 Ground Motion Measurement Devices

C.3.1 Ground Motion Levels

Seismic ground noise varies greatly with time and place. Down to about 1 Hz, a large variety of instrument types can resolve ground noise at noisy sites where cultural noise dominates the spectrum. At the quietest sites though, only the most sophisticated instruments have broad enough bandwidth and a low enough noise floor to accurately measure ground motion from 0.1 Hz to 100 Hz. An informal ground noise standard has been developed at the USGS for bench marking seismic instruments. The power spectral densities for ground noise at some of the most quiet sites in North America have been averaged into a low-noise model which is now included in the instrument noise-floor plots of the most sensitive commercial instruments. Figure C-5 shows the acceleration and displacement amplitude spectra for this low-noise model.

These plots show that at 1 Hz the seismic instrument must detect nano `g's of acceleration per $\sqrt{\text{Hz}}$ which corresponds to 1/10ths of nanometers of ground displacement per $\sqrt{\text{Hz}}$. Ground motion at most accelerator sites is 10 to 100 times larger than the low-noise model for frequencies $\geq$1 Hz where cultural noise dominates.

C.3.2 Accelerometers

The simplest, least expensive detectors are piezoelectric crystals coupled to sensitive charge amplifiers. These instruments are high-sensitivity versions of industrial machine vibration detectors. Their frequency response is nearly flat up to the first resonance of the crystal (100 Hz–1000 Hz). Their noise floor is dominated by electrical noise in the high-gain pre-amp. Figure C-6 shows information from the data sheet for the Wilcoxon Model 731 piezoelectric seismic accelerometer.

At 1 Hz, noise in the Model 731 is equivalent to an acceleration spectral density of 50 ng/$\sqrt{\text{Hz}}$ or a displacement density of 13 nm/$\sqrt{\text{Hz}}$. For noisy sites this detector might show ground noise signals above self-noise at frequencies above 5 Hz, but it does not reach the low-noise model noise floor at any frequency. Its most useful application is the direct measurement of local sources of mechanical noise such as cooling system flow turbulence and pulsed power conversion equipment.

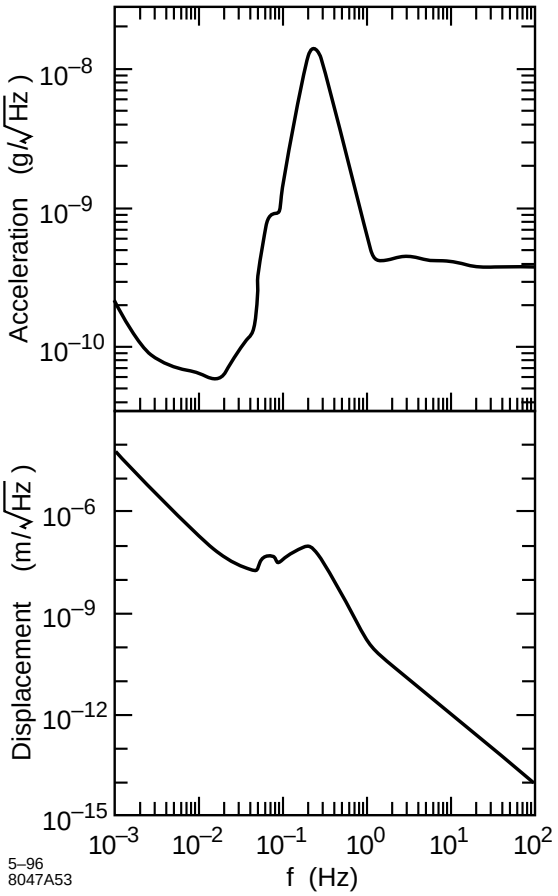


Figure C-5. *USGS low-noise model.*

C.3.3 Geophones

Moving coil velocity meters or geophones transduce earth motion by electromagnetic induction. Figure C-7 shows a simple geophone model. They contain a permanent magnet and a moving coil attached to a spring suspended mass. A voltage is produced across the coil that equals the transduction (G) times the relative velocity ($\dot{x} - \dot{x}_g$) of the mass measured with respect to the permanent magnet which is fixed in the housing of the transducer.

A typical geophone suitable for microseismic ground motion studies is the L4C Geophone made by Mark Products. The 1-kg-proof mass and suspension system has a 1-Hz resonant frequency. In addition to its intrinsic internal mechanical damping, this instrument is usually brought to a .707 critically damped response by shunting the output by a damping resistor $R_s = 8905\Omega$. The Laplace transform transfer function of output voltage *vs.* input ground velocity is:

$$\frac{V_{\rm out}(s)}{\dot{x}_g(s)} = \frac{R_s}{R_c + R_s} G \left(\frac{s^2}{s^2 + 2\omega_0\zeta s + \omega_0^2} \right) \quad \text{where} \quad 2\omega_0\zeta = \left(b + \frac{G^2}{R_c + R_s} \right) / M \quad . \tag{C.49}$$

The amplitude and phase of the complex function $V_{\rm out}(i\omega)/\dot{x}_g(i\omega)$ are plotted in Figure C-8. Above resonance ($s >> i\omega_0$), the transducer conversion factor from Table C-1 for a damped L4C geophone is:

Specifications

Sensitivity, ±10%, 25° C	10 V/g
Accleration Range	0.5 g peak
Amplitude Nonlinearity	1 %
Frequency Response	
± 10 %	0.10 – 300 Hz
± 3 dB	0.05 – 500 Hz
Resonance Frequency, mounted, nominal	950 Hz
Transverse Sensitivity, max	1 % of axial

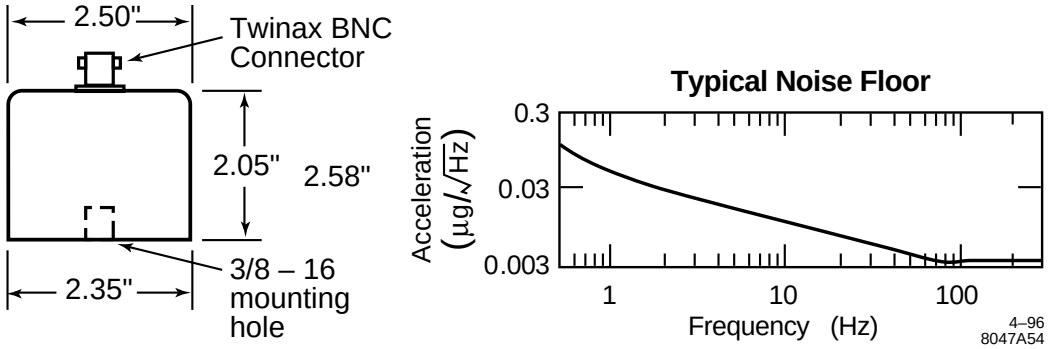


Figure C-6. *Wilcoxon Model 731 accelerometer specifications.*

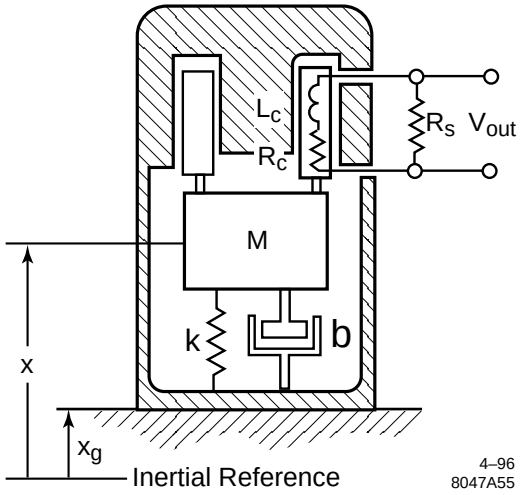


Figure C-7. *Geophone Model.*

Transduction	G	276.4	V s/m
Natural Frequency	ω_0	2π	rad/s
Coil Resistance	R_c	5500	Ω
Shunt Resistance	R_s	8905	Ω
Damping	ζ	.707	

Table C-1. *L4C parameters.*

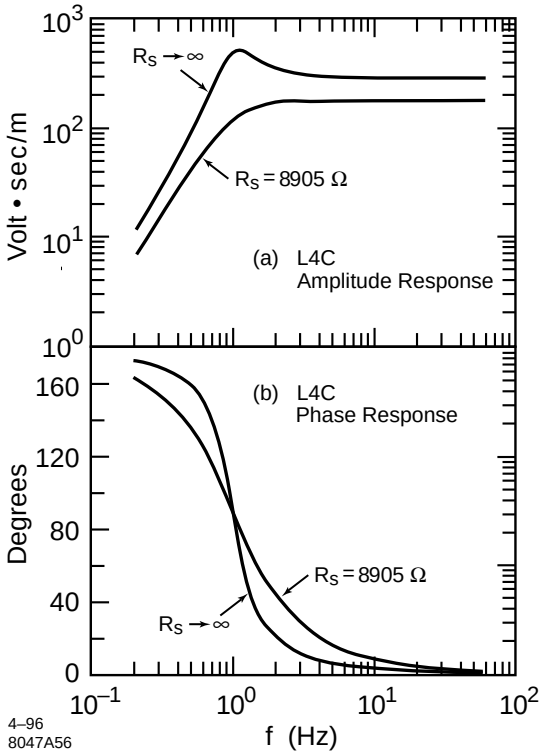


Figure C-8. *L4C transfer function.*

$$\frac{V_{\rm out}}{\dot{x}_g} = \frac{8905\ \Omega}{5500\Omega + 8905\Omega}(276.4\,{\rm V\,s/\,m}) = 170.9\,{\rm V\,s/\,m} \quad . \tag{C.50}$$

One-nm rms motion at 10 Hz, for example, would produce a 11-μV signal at the geophone's output if a 8905-Ω external damping resistor is used.

C.3.4 Sensor Internal Noise Sources

All inertial instruments sense the relative motion of a proof mass with respect to ground. There are three major sources of instrument noise: Brownian motion of the proof mass which is in thermal equilibrium with the outside world through various damping mechanisms [Gabrielson 1993], Johnson noise in the sensor output circuit and electrical current and voltage noise in the input amplifier [Rodgers 1992, Riedesel 1990]. Thermal noise in the mechanical proof mass/suspension system depends on the damping ratio ζ and the ratio of resonant frequency to proof mass. The power spectral density of acceleration noise above resonance is

$$S_{\rm mech} = 16\pi kT\frac{\zeta f_0}{M} \quad ({\rm m/s^2})^2/\,{\rm Hz} \quad . \tag{C.51}$$

For a .707 critically-damped 1-Hz oscillator with a 1-kg mass at room temperature, $S_{\rm mech}$ is $1.46 \times 10^{-19}({\rm m/\,s^2})^2/\,{\rm Hz}$. Integrating the corresponding displacement density above one Hz yields 6-

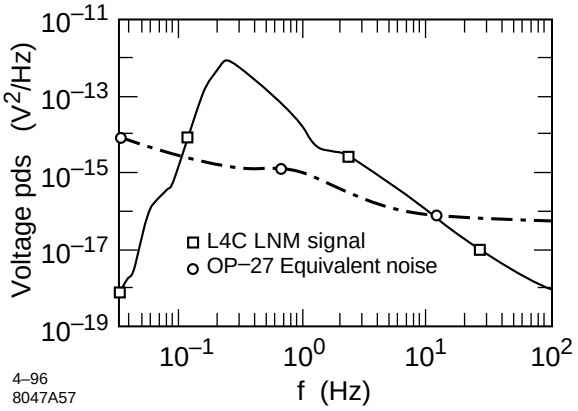


Figure C-9. *L4C low-noise model signal and amplifier noise.*

picometer rms motion. For comparison, the low-noise model integrates to 60-picometer rms above 1 Hz. While thermal mechanical noise is not a serious limit for large proof masses with critical damping, it does set a limit for smaller sensors with proof masses measured in grams.

Johnson noise in the sensor output circuit generates a white-noise voltage signal:

$$S_{\text{johnson}} = 4kTR \quad (\text{V}^2/\,\text{Hz}) \quad . \tag{C.52}$$

A .707 critically-damped L4C has a total parallel coil+damping resistance of 3.4 kΩ which generates a white-noise spectral density of $5.63 \times 10^{-17}\ \text{V}^2/\,\text{Hz}$. When integrated over frequencies above one Hz, this contributes noise equivalent to 7-picometer rms motion.

The third and largest source of electronic noise comes from the various amplifiers used with the sensor. Figure C-9 [Rodgers 1992] shows the performance of the L4C with the OP-27 amplifier. The amplifier noise is below the low-noise model signal from about 0.1 Hz to 10 Hz.

C.3.5 Feedback Seismographs

The most sensitive, low-noise, broad-band modern seismographs employ electromagnetic feedback to stabilize the proof mass to the instrument housing. A model for a feedback seismograph is shown in Figure C-10. Proof-mass motion relative to the instrument housing is detected by a capacitive displacement transducer. This signal is used as the instrument output as well as for feedback control. The feedback elements produce a response equivalent to an oscillator with damping.

The Streckeisen STS-2 is a portable instrument employing feedback to detect seismic signals at the low-noise model level or better from .01 Hz to 10 Hz. Over a wide range of frequencies, the output voltage is proportional to ground velocity $\dot{x}_g$. The instrument contains three identical independent orthogonal seismographs oriented 45° to vertical (Figure C-11). Outputs for vertical and horizontal motions are appropriate vector sums of the three seismograph signals.

To fully model the instrument's response at frequencies around 50 Hz, it is necessary to include the dynamics of the housing's three support legs. The instrument response is the product of the feedback seismograph response and the

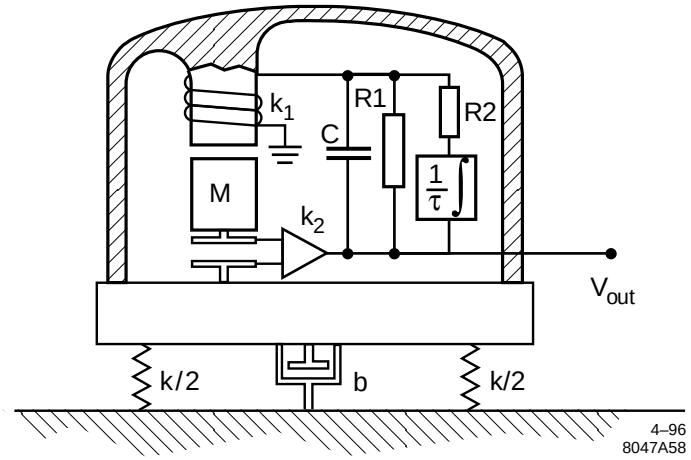


Figure C-10. *Feedback seismograph model.*

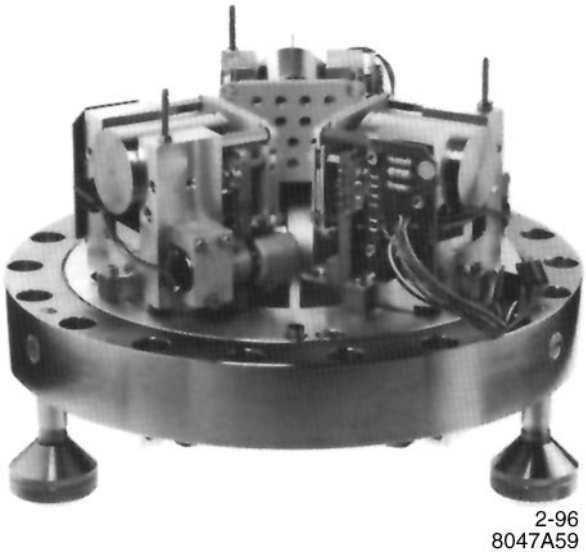


Figure C-11. *STS-2 with its cover removed.*

Proof mass	M	0.3	kg
Transduction	k_1	50	Newtons/amp
Displacement transducer	k_2	8×10^4	V/m
Integration time	τ	80	s
Feedback resistor	R_1	1.7×10^6	Ω
Integrator resistor	R_2	0.6×10^6	Ω
Derivative capacitor	C	7.8×10^{-6}	farads
Support frequency	ω_0	$2\pi(50)$	rad/s
Support damping	ζ	0.2	

Table C-2. *STS-2 parameters.*

response of the instrument's support structure. The Laplace transform transfer function for the STS-2 is

$$\frac{V_{\rm out}(s)}{\dot{x}_g(s)} = \left(\frac{\omega_0^2 + 2\omega_0\zeta s}{s^2 + 2\omega_0\zeta s + \omega_0^2}\right)\left(\frac{Ms^2/k_1C}{\frac{M}{k_1k_2C}s^3 + s^2 + \frac{1}{R_1C}s + \frac{1}{R_2C\tau}}\right) \tag{C.53}$$

and its parameters are listed in Table C-2. The amplitude and phase of this transfer function evaluated with $s = i\omega$ are shown in Figure C-12.

C.3.6 Other New Sensors

New methods of measuring inertial forces have recently started development. Optical strain sensor in the form of glass fiber interferometers wrapped around elastic proof-masses have recently been proposed as the basis of a sensitive accelerometer [Gardner 1987]. Field effect tunneling currents are sensitive to Å changes in the tunneling gap. Integrated circuit seismographs have recently been built using tunneling current to monitor proof-mass motion [Kenny 1994]. Because of their small size these instruments are subject to thermal noise but in the future larger transducers using this principle may be developed.

Over the past decade the needs of integrated circuit manufacturing and optical fiber communication have made commercial transducers and actuators available with sub-nanometer resolution. Capacitance position measuring sensors can be purchased with a better than 0.01-nm/$\sqrt{\rm Hz}$ noise floor and 50-μm range.

C.4 SLAC Ground Motion Measurements and Analysis

As discussed in Section C.2, one needs to know both the ground motion spectrum and correlations to compute the its effect on beam motion in the NLC. While much data exists on ground motion spectra, little is available on its spatial characteristics. One exception is a set of correlation measurements between points separated by 0 to 3 km in the LEP tunnel [Juravlev 1993]. However, the limited resolution of the seismometers used for these measurements does not allow one to fully examine the quantities of interest. To see how typical the CERN results are, and to better understand the dynamics of ground motion, we did a similar set of correlation measurements in the SLAC linac tunnel using higher-resolution seismometers. Because of time constraints, only vertical ground motion was studied in detail. In this section, we describe these measurements, the spectral and spatial analysis of the data, and how the results can be used to compute the effect on beam motion in the NLC.

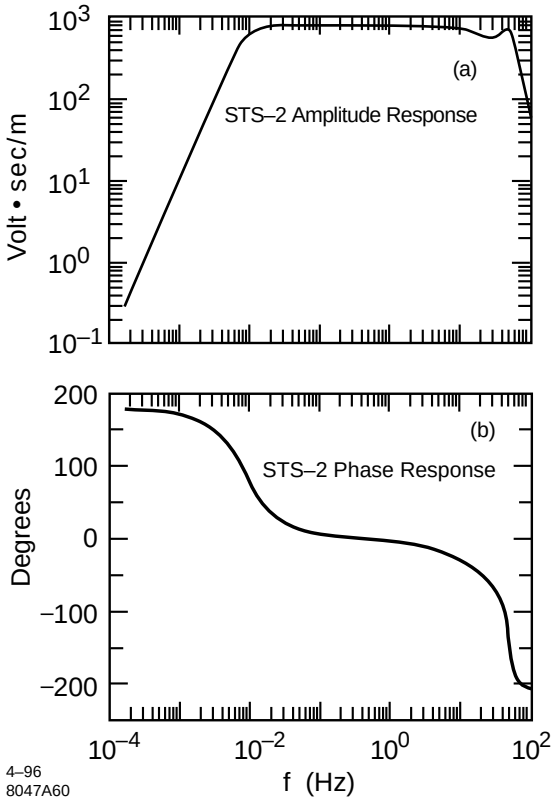


Figure C-12. *STS-2 frequency response.*

C.4.1 STS-2 Measurements and Data Analysis

The seismic measurements were made in the SLAC linac tunnel during a period when the magnet and accelerator cooling water was shut off. Two Streckeisen STS-2 seismometers were borrowed for this purpose and were placed on the tunnel floor with Styrofoam boxes around them for thermal insulation and to shield them from air currents (see Section C.3.5 for a description of these devices). The basic measurement consisted of simultaneously digitizing the vertical ground motion signals from the two seismometers at an effective rate of 128 Hz for about a two-hour period using a PC-based data acquisition system. The actual acquisition rate was 6.4 kHz but the digitized results were decimated by 50 to reduce the data volume to a manageable size. Before decimation, the data were convoluted with a Gaussian function (3 dB point = 96 Hz) to filter out high frequency components.

Over a period of a few weeks, measurements were taken with one seismometer at the end of linac Sector 4, and the other downstream from it by distances (Δz) of 0, 6, 12, 24, 50, 100, 200, 400, 1000, and 2000 m. After any change in position, the STS-2s were allowed to stabilize for at least 10 hours before data were recorded. Typically, the configuration was kept fixed for one or two days, and three measurements were made each day beginning at 10 PM, 2 AM and 6 AM. The results presented here are based on the 2 AM data.

After completing a measurement, the data were divided into 50 128-s samples and the complex Fourier transform of each sample was computed with a cosine-like windowing function. Since the response function of the seismometers is fairly flat in the frequency (f) range of interest, no corrections were applied. However, the seismometers measure

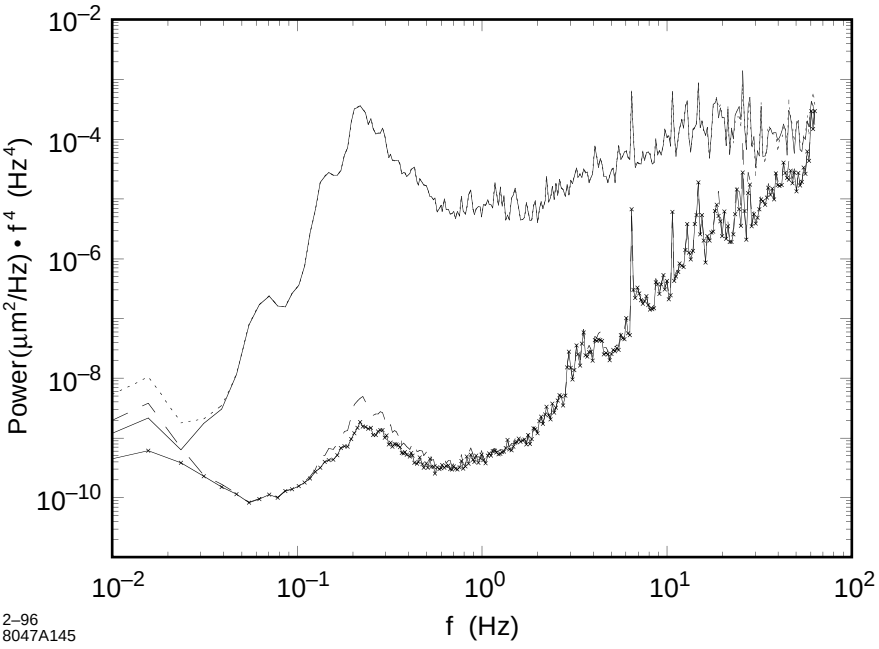


Figure C-13. *Vertical power spectra (weighted by f^4) measured by two STS-2 seismometers placed side-by-side (solid and dotted lines). Also shown are the difference (dashed line) and corrected difference (line with crosses) power spectra.*

velocity, so the Fourier results were divided by $2\pi fi$ to obtain the vertical displacement components, $A_1(f)$ and $A_2(f)$, of the ground motion at the two STS-2 locations. From these, the average of the 50 power spectra were computed, $< A_1^2(f) >$ and $< A_2^2(f) >$, as well as the cross power $< A_1(f)A_2^*(f) >$, from which the correlation was obtained,

$$\text{Correlation} \equiv Re\left\langle A_1(f)A_2^*(f)\right\rangle / \sqrt{\left\langle A_1^2(f)\right\rangle \left\langle A_2^2(f)\right\rangle} \quad . \tag{C.54}$$

Also, the difference spectrum, $\left\langle \left(A_1(f) - A_2(f)\right)^2 \right\rangle$, was calculated.

C.4.2 Measurement Resolution

The performance of the measurement system is best characterized by the data that were taken with the seismometers placed side-by-side ($\Delta z = 0$). Figure C-13 shows the two power spectra for this configuration together with the difference spectrum, and the quantity,

$$\text{Corrected Difference} \equiv 2(1 - \text{Correlation})\sqrt{\left\langle A_1^2(f)\right\rangle \left\langle A_2^2(f)\right\rangle} \tag{C.55}$$

which would be the difference spectrum if the calibration of the two seismometers were identical and equal to the geometric mean of the actual calibrations (note that the spectra are weighted by f^4 to make the plot more compact). One sees that the difference spectrum is generally close to the corrected spectrum except at low frequency where the deviation corresponds to a calibration difference of about a factor of two. Also, the curves deviate near the micro-seismic peak (f = .2 Hz) where the corresponding calibration difference is only a few parts in a thousand. While this latter deviation is consistent with the expected accuracy of the seismometers, the reason for large low frequency

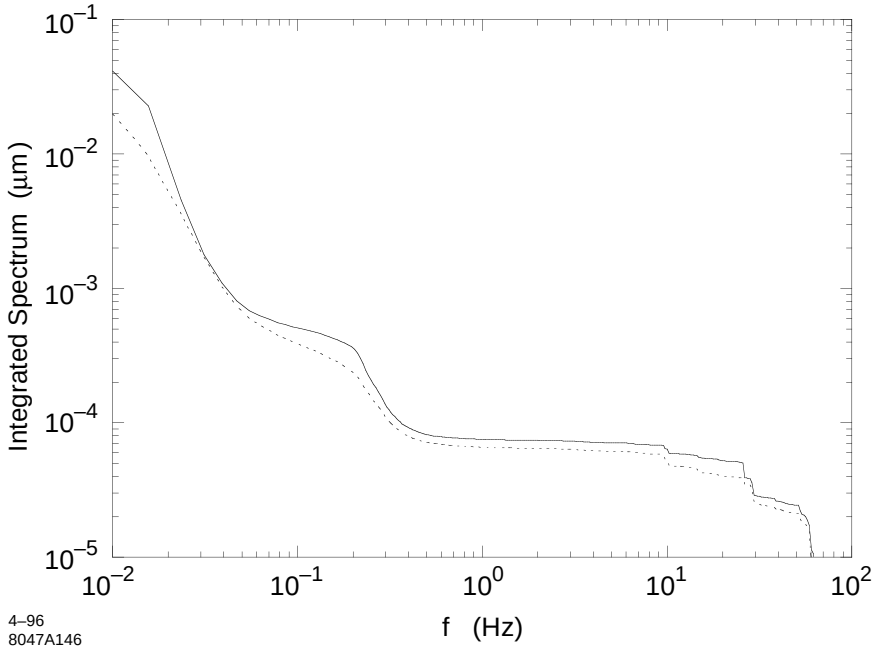


Figure C-14. *Integrated difference (solid line) and corrected difference (dotted line) power spectra computed from the difference spectra in Figure C-13.*

difference is not clear. However it occurs in a region where power spectrum contributes little to the integrated motion so it does not have a big effect on the results.

The difference spectrum in Figure C-13 is roughly that expected from the noise characteristics of the seismometers. To better characterize this noise floor, we computed the integral of the difference spectrum which yields the vertical rms ground motion difference that would be measured if only frequencies above the frequency plotted contributed to the motion. Figure C-14 shows this integrated spectrum together with that for the corrected difference spectrum. Thus, the STS-2s can discern ground motion differences of about 0.5 nm for $f > 0.1$ Hz, which is adequate for most NLC studies. As a measure of how sensitive these devices are, we note that this difference level increases by about 100 if the Styrofoam boxes are not used to isolate the seismometers.

C.4.3 Spatial Properties of the Vertical Ground Motion

The data that were taken with the two seismometers separated by various distances were analyzed in the same manner as described above. One sees that the difference spectra rise as the seismometer separation increases. The differences in these cases are best characterized by the correlation spectra since these ratios tend to be more stable than the power spectra, which can vary by as much as ± 10 dB day-to-day. Also, they are more useful for understanding the dynamics of ground motion and for computing its effect on beam transport.

As examples of correlation measurements, Figures C-15 and C-16 show the results obtained with the seismometers separated by 100 m and 1000 m, respectively. As with the CERN results, the correlation is nearly unity at low frequencies (the dip near .01 Hz is due to the electronic noise of the seismometers). Just how the correlation varies with both frequency and separation is needed to compute the effect of ground motion on the beams in the NLC. In

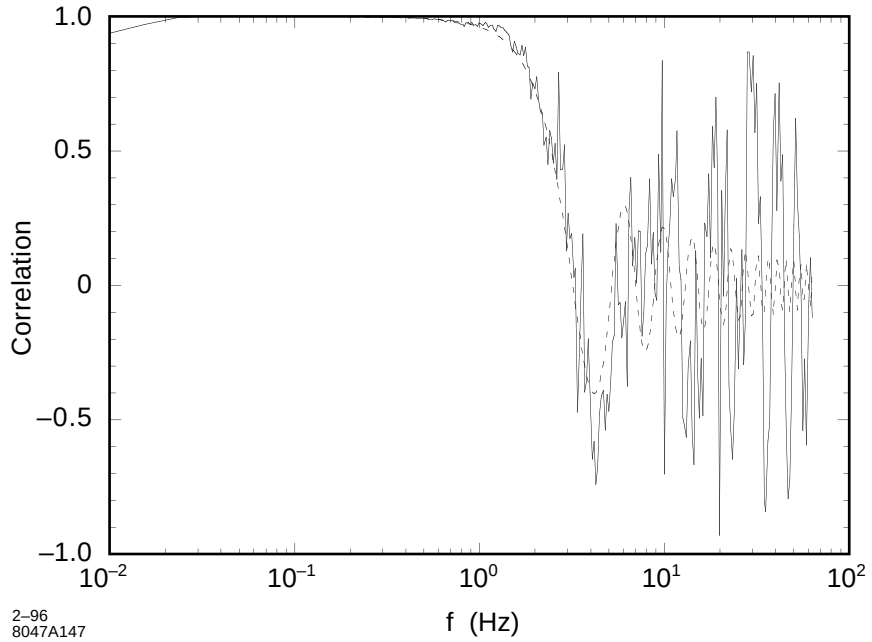


Figure C-15. *Correlation spectrum (solid line) measured with the seismometers separated by 100 m. The dashed line is described in the text.*

examining the data, it was found that it can be reasonably well described by a model in which the ground motion consists of horizontally traveling waves that are isotropically distributed in direction. This model predicts that the correlation equals $J_0(2\pi f \Delta z/v)$ where J_0 is the zero order Bessel function, f is the frequency of the waves, Δz is the separation of the seismometers, and v is the wave phase velocity, which we find depends on f (*i.e.*, the ground is dispersive).

We determined v as a function of f by dividing the data into narrow frequency bins and then fitting the correlation *vs.* Δz curves for each set to the expected Bessel function dependence with velocity as the only free parameter. For this procedure, the log of ($1-$ correlation) was actually fit to better match how the data deviates from unity, which is particularly important for modeling ground motion (note, $1-$ correlation $= R$ in Section C.2). As an example, Figure C-17 shows the $0.8 < f < 0.9$ Hz data, which have been averaged together in this plot, and the fit function, $1 - J_0(2\pi f \Delta z/v)$, which is evaluated at the fit velocity, 1510 m/s, and the mean frequency, 0.85 Hz.

The velocity fits were done only for the data in the $0.15 < f < 12$ Hz range. At higher frequencies, the correlations are small even at the lowest values of Δz so the velocity dependence is hard to discern. At lower frequencies, the decreasing ratio of ground motion to seismometer noise limits the minimum value of ($1-$ correlation) that can be measured.

The values of velocity that result from the fits are plotted in Figure C-18 as a function of f. To characterize the dispersion, the data were empirically fit: an exponential function provides a good match and yields,

$$v(\text{ m/ s}) = 450 + 1900\exp(-f(\text{ Hz})/2.0) \tag{C.56}$$

which is shown as a dashed line in Figure C-18. To check our results, we have searched geophysics journals for other measurements of dispersion in this frequency range. Although we have not yet found direct measurements, theoretical predictions of the dispersive properties of surface waves at lower frequencies would suggest that our results

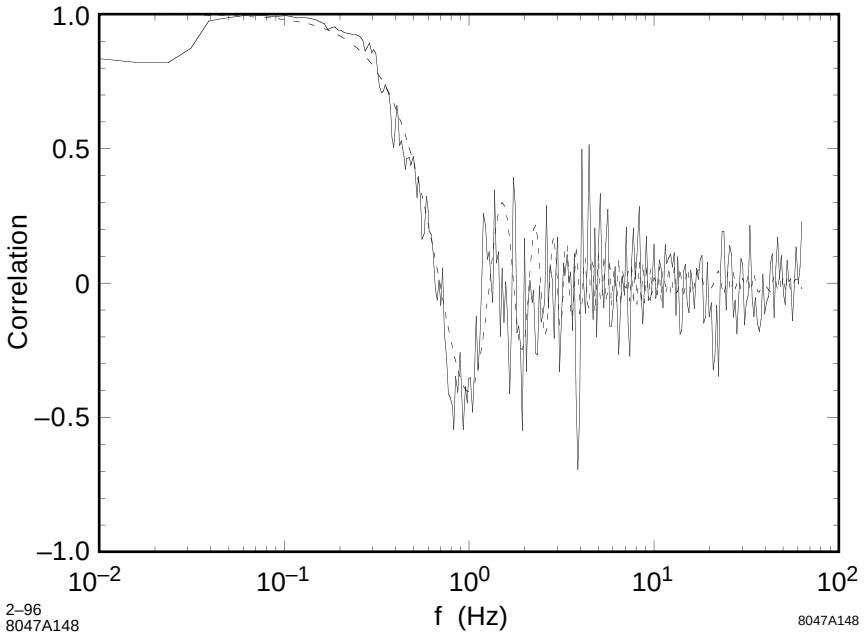


Figure C-16. *Correlation spectrum (solid line) measured with the seismometers separated by 1000 m. The dashed line is described in the text.*

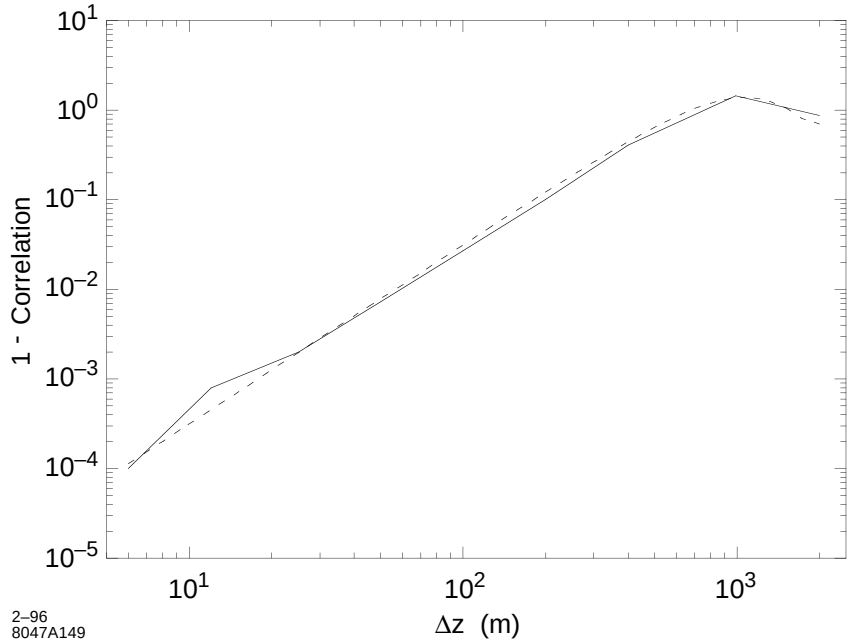


Figure C-17. *Dependence of (*$1-$ *correlation) on the separation (*Δz*) of the seismometers for* $0.8 < f < 0.9$ *Hz. The data in this frequency range has been averaged (solid line). Also shown is the fit to the data (dashed line) which is described in the text.*

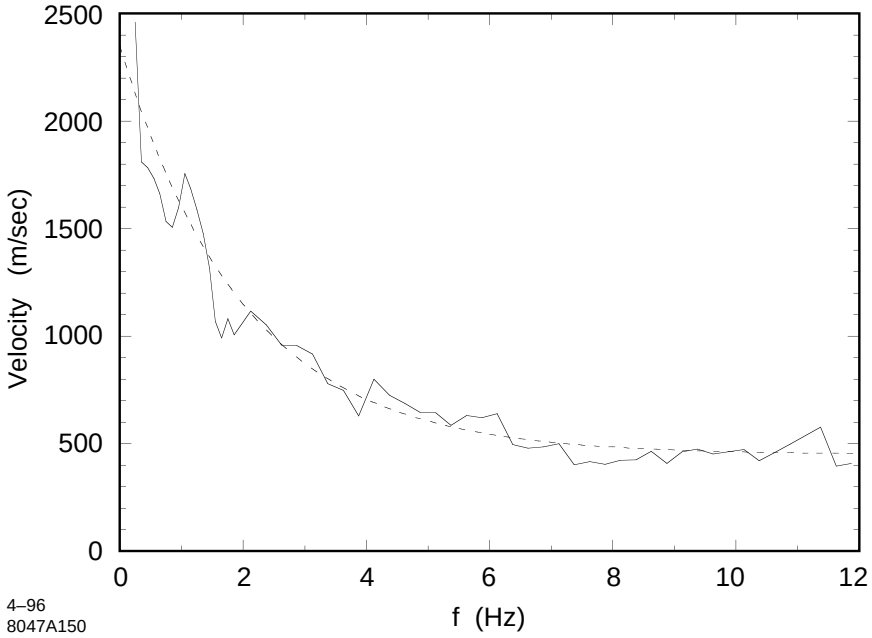


Figure C-18. *Velocity derived from fits to the correlation data (solid line). The dashed line is an empirical fit to the data:* $v = 450 + 1900\exp(-f/2.0)$.

are not unreasonable [Bolt 1972]. Also, a simple analysis of the CERN data, using only the values of Δz at which the correlations first cross zero, yields a similar dispersion relation although the velocities are about 30% larger.

Using the measured dispersion function, the correlation versus frequency measurements were compared to predictions. The dashed lines in Figures C-15 and C-16 are examples. In regions where the correlation is high, the match is good as expected. The fact that the predictions do not match the data in regions of low correlation is less important since it is the Δz dependence of transition from high to low correlation that needs to be well modeled to characterize the spatial properties of the ground motion at any frequency.

So far we have considered only the real part of the measured correlations. If the ground motion is wave-like, and the waves are isotropic in direction, then the imaginary part of the correlations should be zero. What is observed, however, is roughly a $J_1(2\pi f \Delta z/v)$ dependence for frequencies below about 1 Hz. This dependence corresponds to a $(1+\cos(\phi))^2$ distribution of incoming waves where ϕ is the angle with respect to the linac axis. This result is not too surprising given that the SLAC linac is oriented nearly perpendicular to the Pacific coastline, and most seismic activity in this frequency range originates from ocean waves. If we redo our analysis of the real component of the correlations assuming a $(1+cos(\phi))^2$ distribution, the velocities that result are systematically about 10% larger, and the goodness-of-fits are about the same. Since only the real part of the correlation is needed for the beam motion predictions, and the J_0 dependence adequately matches the data, we continue to use this simple functional dependence and the corresponding dispersion results.

C.4.4 Application to the NLC

With the correlation function described above, the effect of ground motion on beam transport in the NLC can be derived using the formalism given in Section C.2. In particular, the luminosity reduction, $\Delta\mathcal{L}/\mathcal{L}$, due to vertical

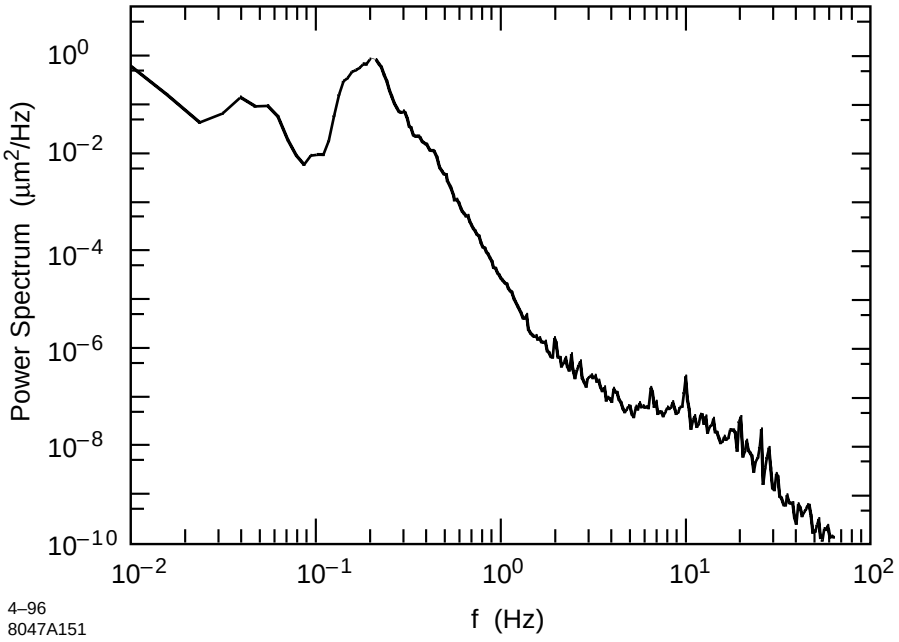


Figure C-19. *Average of 10 power spectra of vertical ground motion that were measured on different days in Sector 4 of the SLAC linac at 2 AM.*

ground motion with frequencies $> f_0$ is

$$\Delta\mathcal{L}/\mathcal{L}(f_0) = \frac{\left\langle \Delta^2_{\rm IP} \right\rangle_{f_0}}{4\sigma^2_{y,\rm IP}} = \frac{1}{2\pi\sigma^2_{y,\rm IP}} \int_{f_0}^{\infty} P(f) \int_0^{k_0} \frac{G(k)}{\sqrt{k_0^2 - k^2}} dk\ df \tag{C.57}$$

where: $\sigma_{y,\rm IP}$ = rms vertical beam size at the IP, $k_0 = 2\pi f/v(f)$ with $v(f)$ from Eq. C.56, $P(f)$ = power spectrum, and $G(k)$ = lattice response function, which is defined in Section C.2.2. The effect of trajectory feedback loops in the NLC can be included by factoring into the integral the frequency response function of the feedback system.

To a good approximation, one can split the function $G(k)$ for the NLC into a sum over its values in different regions of the machine (this generally yields a worst-case estimate). Only one half of the NLC needs to be considered since $G(k)$ for the opposing half is simply related by symmetry. In this approximation, $G(k)$ for a region is computed for the case where the beam initially follows the ground motion, and the difference in the beam and ground motion that develops by the end of the region remains unchanged relative to the beam size as the beam propagates to the IP. Using the $G(k)$ functions given in Section C.2.2 and the power spectrum that was measured in the SLAC linac tunnel (Figure C-19) and elsewhere, the luminosity losses due to ground motion in most regions of the NLC have been computed.

To be conservative, estimates have also been made which account for the limits that the electronic noise of the seismometers place on our measurement of correlation. As a worst case, the electronic noise contribution to the data is assumed to be ground motion that is uncorrelated, quadrupole-to-quadrupole. That is, at each frequency it is assumed that a certain fraction of the measured power is due to waves, and the remainder is due to uncorrelated motion. Combined, the two sources account for the measured power and correlation spectra except for $\Delta z = 0$ where the predicted correlation is unity but measured is not since the uncorrelated component is actually noise.

The noise power was obtained from the $\Delta z = 0$ difference data and is plotted in Figure C-20 as a fraction ($\equiv \rho_n$) of the average total power (note that ρ_n is the minimum value of (1 - correlation) measurable if the ground motion

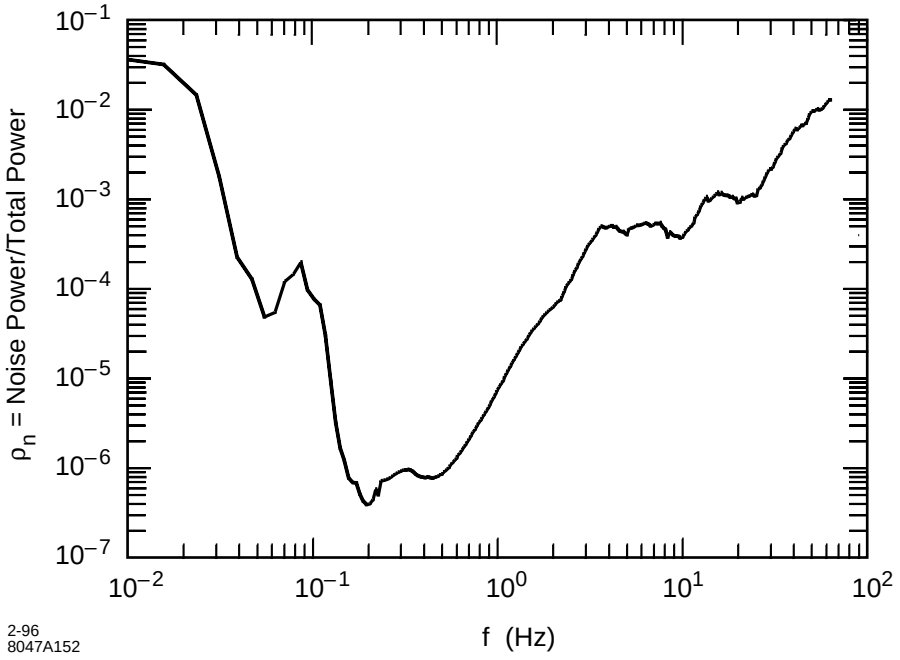


Figure C-20. *Fraction (ρ_n) of the power in Figure C-19 that is due to STS-2 electronic noise.*

spectrum equals that in Figure C-19). Including the noise as uncorrelated motion, the expression for the luminosity reduction becomes

$$\Delta\mathcal{L}/\mathcal{L}(f_0) = \frac{1}{2\pi\sigma^2_{y,\mathrm{IP}}}\int_{f_0}^{\infty} P(f)\left[\int_0^{k_0}\frac{(1-\rho_n)G(k)}{\sqrt{k_0^2-k^2}}dk + \frac{\pi\rho_n G(\infty)}{2}\right]df \quad . \tag{C.58}$$

To predict the luminosity loss due to lower-frequency ($f < .01$ Hz) ground motion requires some assumption about the spatial characteristics of the motion since we do not have measurements in this regime. There is evidence at long time scales (days to years) that ground motion has an "ATL" behavior where the variance of relative displacements grow linearly with time and separation (*i.e.*, $\langle\Delta y^2\rangle = A \times T \times L$ where A is a constant) [Baklakov 1991]. The frequency scale at which the motion changes from predominately wave-like to ATL-like is not known. However, as a worst case, one can assume that this occurs at 0.01 Hz and use the value of A obtained from our data as an upper limit. To compute A, we note that the power spectrum for ATL-like motion is

$$P(\omega, k) = 4A/\omega^2 k^2 \tag{C.59}$$

so the rms squared difference of the motion for $f > f_0$ at two locations separated by Δz is

$$\langle\Delta y^2\rangle = \int_{2\pi f_0}^{\infty}\int_0^{\infty}\frac{4A}{\omega^2 k^2}2\left(1-\cos(k\Delta z)\right)\frac{dk}{2\pi}\frac{d\omega}{2\pi} = \frac{A\Delta z}{2\pi^2 f_0} \tag{C.60}$$

For $\Delta z = 200\,\mathrm{m}$ and $f_0 = .02\,\mathrm{Hz}$, we measure a 10-nm-rms vertical deviation, yielding $A = 2\times10^{-7}\,\mu\mathrm{m}^2/\,\mathrm{m}/\,\mathrm{s}$, which is on the low end of that measured on long time scales (another method to compute an upper limit on A is discussed in Section C.2.1). Using this value in the ATL power spectrum above, one can thus make worst-case estimates of the low frequency contribution to luminosity loss assuming that A is indeed frequency-independent.

The results on luminosity loss due to ground motion in the various regions of the NLC are presented in the chapters that discuss these regions. It should be noted that the general formalism presented here can also be used to estimate beam-emittance growth that depends quadratically on quadrupole position, such as that due to dispersion.

References

[Baklakov 1991] B.A. Baklakov, P.K. Lebedev, V.V. Parkhomchuk, A.A. Sery, A.I. Sleptsov and V.D. Shiltsev, "Investigation of Seismic Vibrations and Relative Displacement of Linear Collider VLEPP Elements", *Proc. 1991 Part. Accel. Conf.*, San Francisco, CA, 3273 (1991).

[Bolt 1972] B.A. Bolt (editor), "Seismology: Surface Waves and Earth Oscillations", *Methods in Computational Physics*, Vol. 11 (Academic Press, NY, 1972).

[Gardner 1987] D.L. Gardner and S.L. Garret, "Fiber Optic Seismic Sensor", SPIE Vol. 838, *Fiber Optic and Laser Sensors V*, 271–277 (1987),

[Gabrielson 1993] T.B. Gabrielson, "Mechanical-Thermal Noise in Micromachined Acoustic and Vibration Sensors", *IEEE Trans. Electron. Devices* **40**, 903–909 (1993).

[Juravlev 1993] V.M. Juravlev *et al.*, "Investigations of Power and Spatial Correlation Characteristics of Seismic Vibrations in the CERN LEP Tunnel for Linear Collider Studies", CERN Note CERN–SL/93–53 and CLIC–Note 217 (December 1993).

[Kenny 1994] T.W. Kenny, W.J. Kaiser, H.K. Rockstad, J.K. Reynolds, J.A. Podosek, and E.C. Vote, "Wide-Bandwidth Electromechanical Actuators for Tunneling Displacement Transducers", *J. Micro. Elec. Sys.* **3**, 97–103 (1994).

[Riedesel 1990] M.A. Riedesel, R.D. Moore, and J.A. Orcutt, "Limits of Sensitivity of Inertial Seismometers with Velocity Transducers and Electronic Amplifiers", *Bull. Seismological Soc. of Am.* **80**, 1725–1752 (Dec 1990).

[Rodgers 1992] P. Rodgers, "Frequency Limits for Seismometers as Determined from Signal-to-Noise Ratios", *Bull. Seismological Soc. of Am.* **82**, 1071–1098 (April 1992).

Contributors

- Chris Adolphsen
- Gordon Bowden
- John Irwin

D

Beam-based Feedback: Theory and Implementation

Contents

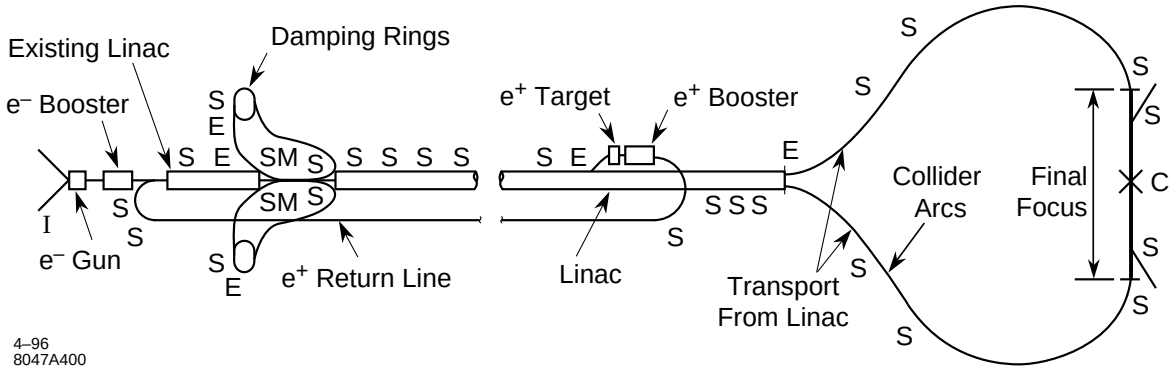


Figure D-1. *SLC schematic with feedback locations. S=steering; E=energy; I=intensity/gun; C=maintains beam collisions; M=minimization*

D.1 Introduction

In planning for feedbacks in the NLC, it is useful to study the existing SLC feedback system as a model. While the control system architecture for the NLC is likely to be different from the SLC, the basic feedback algorithms should be similar. The SLC feedback system is the first of its kind for accelerators. It is a generalized, database-driven system that applies linear closed-loop control. The mathematical design goal is to minimize the RMS of the beam jitter by applying state space techniques. In practice in the SLC, the feedback serves to stabilize the beam and decouple different areas of the machine, facilitating smoother startup after machine outages and easier machine tuning. This system has been essential to the operation of the SLC, and it is assumed that a similar system will be equally necessary for the NLC.

D.2 Planned NLC Feedbacks

For the NLC, feedback loops are planned for each major area of the machine, similar to the SLC feedbacks. At the guns, intensity and timing parameters will be controlled. In addition, feedback controls for each of the laser wire scanners are expected. Energy and steering (position and angle) parameters will be regulated for damping ring injection and extraction. For each linac, five diagnostic sections will each have energy and steering controls. For each final focus, feedbacks will stabilize the beam steering at two locations. Finally, an interaction-point feedback loop will regulate the beam-beam deflections to keep the beams in collision.

The feedbacks are planned to run at the full beam rate of 120 or 180 Hz. The steering loops control the average trajectory of the bunch train rather than individual bunches, so the "Q" BPMs are used, which measure the average of the train. Dipole corrector magnets are used, for which the feedback controls the magnetic field by setting a DAC (digital to analog converter) which alters the current from a power supply. Correctors and other actuators need to respond (to make 90 percent of a requested change) in a single 180-Hz period.

If SLC experience is taken as a model, many additional feedbacks are likely to be added to those originally planned. For the SLC, the original eight loops have been expanded to over 50 control loops, many of which are shown in Figure D-1.

D.3 Feedback System Design

The SLC feedback system is designed in a generalized, database-driven fashion, which contributes greatly to the flexibility of the system [Hendrickson 1995]. The system uses standard control-system hardware. As a result of this design, unplanned control loops can often be added with only database work, without requiring hardware or software changes. In addition to the beam position monitors and correctors used for steering control, the feedback system is capable of measuring and controlling a wide variety of devices. For example, it is equally trivial to add a steering loop in the linac as it is to stabilize the laser gun timing in the injector. Special-purpose extensions to the linear feedback system have been added to accommodate non-linear cases, such as optimization feedbacks in which the measurement responds parabolically to actuator movement. The system also provides built-in diagnostic and analysis capabilities, and the many sample-only monitoring loops provide a wealth of diagnostic information. These design features have been key to the success of the SLC system, and the NLC design should be equally flexible and extensible in order to support unplanned controls needs.

The feedback control algorithm is based on state space formalism, with an LQG (Linear Quadratic Gaussian) controller. Matrices are designed and calculated offline in advance, with inputs including a model of the beam transport and the expected beam and BPM noise characteristics. The mathematical design minimizes the rms of the beam states over time, given the noise inputs [Himel 1991]. By modifying the input beam noise design assumptions, it is possible to tune the feedback performance response characteristics. The initially-proposed feedback algorithm does not adapt to modeling or noise spectrum changes, consistent with the current SLC design. In the future, adaptive methods may provide improved performance. The beam transport characteristics within a single loop may either be obtained from the accelerator model, or be measured by an invasive online beam-based calibration procedure.

An extension to the basic feedback system, “cascade” is designed to allow multiple linac loops to communicate with each other, avoiding overcorrection problems when a perturbation is induced upstream of the chain of feedback loops. With cascade, each feedback loop receives the calculated positions and angles from the next upstream loop, mathematically transports them to the downstream location, and subtracts them from the states calculated using the local BPMs. The resulting adjusted states are then used for the local feedback corrections, so that each loop should correct only perturbations which were not seen by an upstream loop.

D.4 Performance Questions

There are a variety of performance questions which have been investigated in the SLC system, in particular with the chain of linac steering loops. These include concerns about the stability of the SLC model, speed of the steering magnets, and functions of the cascade system. In the SLC, gain factors can be used to slow the feedback response and reduce sensitivity to suspected feedback imperfections. Several types of imperfections are discussed in the following sections. Where quicker feedback response is not needed, the lower gain factors have been successful in improving the stability of the feedback system.

D.5 Adaptive Cascade

Problems were observed with the cascade system, which is intended to allow each of a string of feedback loops to correct only the perturbations initiated immediately upstream of it. The cascade system relies on a linear beam

transport which is independent of the source of the perturbation. The adaptive feature of the cascade system enables each feedback loop to learn the transport from the upstream loop, using the beam jitter to calculate the beam transport. The adaptation has the assumption that perturbations immediately upstream of a feedback loop are uncorrelated with upstream perturbations. In the current SLC cascade design, each loop obtains beam information only from the adjacent upstream loop; the assumption is that if the loop upstream saw a perturbation, either it or any loops further upstream will eventually fix it.

In several tests, it was observed that the linac loops did not exhibit perfect cascade response. In particular, the feedback response to perturbations induced in the middle of the linac is different from the response to perturbations from the beginning of the linac. The SLC design assumes that the beam transport is independent of the source of a perturbation. At high currents this is not valid because of the effect of transverse wakefields which cause oscillations to propagate differently depending on their source. This is a fundamental problem with the SLC architecture, but for the low currents of the NLC, this effect should be less significant. If wakefield effects are expected to be a problem for the NLC, an alternate cascade design should be developed to provide downstream feedback loops access to beam information from all upstream loops instead of just the single nearest upstream neighbor. This would require more complicated algorithms and more communications paths than those available in the present SLC system.

Tests of cascade performance during low-current operation uncovered additional information. A test was done to confirm that the cascade system responds well when the transport is invasively measured (instead of adaptively calculated), and the resulting cascade response was between 95 and 100% effective (nearly perfect). A design flaw in the adaptive beam transport calculation, associated with exception handling for broken BPMs, was found and fixed. In the most recent test, large betatron jitter was induced upstream of the chain of linac feedback loops. The feedbacks adapted to the induced beam noise, and the resulting cascade response was between 90 and 100% effective. However, without either measurement or induced noise, the performance was much poorer. This may be explained by an additional design flaw, in which poor BPM resolution during low-current SLC running results in incorrect adaptive transport calculations. It is hoped that when this design flaw is fixed or better BPMs are available, low-current cascade adaption performance will be improved.

Further concerns include questions about the nature of the incoming beam jitter, which is the source of the adaptive transport calculations. If the incoming jitter is not dominated by betatron jitter, adaptive transport calculations would get the wrong answer. For the NLC, a conservative initial plan is to omit the cascade adaption to calculate the interloop beam transport matrices, and instead measure the transport semi-invasively by stealing beam pulses and perturbing the beam once an hour. With this method, the cascade should be able to work well at low intensities.

D.6 Rate Considerations, and Corrector Speeds

Ideally, feedbacks should operate at the full repetition rate of the machine. At the SLC, where the rate for many loops is limited to 20 Hz by CPU and other hardware constraints, aliasing problems associated with the partial sampling rate have been observed. Full-rate operation requires sufficient processing power and hardware response time.

A source of poor performance which was investigated at the SLC is the sensitivity to corrector speeds. Feedback simulations showed that, when the corrector speeds are slower than expected, the feedback performance is degraded. In particular, performance is extremely sensitive to the relative speeds of correctors within a single loop. If some correctors are slow, it is better to have them all the same speed as the slowest one. This effect is exaggerated when there are modeling errors or other imperfections in the loop design.

The correction system needs to be designed such that the response of the correctors is within one interpulse period, considering the speed of the power supply and the field propagation through the beam pipe. For the latest SLC run, a

new feedback linac steering loop was implemented which is capable of 120 Hz response. This feedback loop has been commissioned and, under some noise conditions, has decreased the RMS of the beam jitter by up to 40%.

D.7 Calibrations and Modeling

The correct functioning of a feedback loop depends on knowledge of the model and transport between the steering magnets and position measuring devices. For the SLC, this transport is derived from the online machine model. A calibration procedure is provided to check and possibly update the transport matrices. Because such a procedure is invasive, it is rarely used. For the future, it may be desirable to use a fully adaptive feedback algorithm which is capable of responding online to machine changes. This extension of the state-space formalism is currently being developed at CEBAF, and may be useful for both the SLC and the NLC.

Recent SLC studies indicate that poor modeling is not currently as significant a problem for the linac feedbacks as are slow correctors and imperfect cascade performance. However, sensitivity to the model is exacerbated by incorrect modeling of the corrector response or other errors. Furthermore, simulations show that minor modeling imperfections can have disastrous results when combined with aggressive noise designs such as notch filters.

D.8 Global Performance Characterization

In addition to studies of specific sources of feedback imperfection, measurements and simulations were performed to study the global feedback response of a series of linac loops. For the SLC, measurements were taken, comparing the response of a step function with the series of linac loops off versus on. The FFT of both sets of data were calculated; the ratio of the FFTs with feedback on and off provides a measured amplification curve. Simulations reproduced the measured results, including the effects of measured imperfections. These measurements and simulations included effects of imperfect cascade correction, imperfect modeling, slow correctors, multiple loops running at low rates, and different gain factors. Reproducing the measured SLC performance improved confidence in our ability to simulate imperfections realistically. Given some assumptions about NLC conditions and imperfections, performance was evaluated for the series of NLC linac loops.

The NLC simulations evaluate a series of five linac loops, each running at the full beam rate of 120 Hz with relatively fast correctors that can control in one machine pulse. For the NLC, slow correctors should not be a problem. Imperfections in modeling and cascade response are assumed for the NLC. We assume that modeling/calibration imperfections are comparable to SLC conditions. We assume that cascade is 85% effective. While still imperfect, this is better than the current SLC performance; wakefield effects shouldn't be a significant problem, and we assume that the cascade transport for the NLC will be measured each hour. For the “SLC” noise design, which corrects a step function with an exponential time constant of six pulses, the results of the NLC simulation were very good. Response of the system to a step function is good and the system is able to damp very well at frequencies below 6 Hz. Above 6 Hz, beam noise is amplified somewhat, with a maximum amplification of 1.5. In order to damp more strongly at low frequencies, more aggressive noise designs were considered. While the more aggressive designs damp better at low frequencies and are able to damp noise up to 10 Hz, the jitter amplification at higher frequencies is increased, up to a factor of 4.5 for one design. Furthermore, when the imperfections are considered, the more aggressive designs have a poorer response to a step function. These studies indicate that with the more-conservative “SLC” noise design, modest imperfections have a minor effect on performance, but with a more aggressive design the same imperfections become significant. More work should be done to find an optimal design, but at this point the “SLC” noise design produces acceptable results while providing a robust system which is tolerant of minor imperfections.

D.9 Summary

Initial NLC simulations indicate that acceptable feedback performance can be obtained with an "SLC"-type noise design. More work should be done to characterize feedback performance and to study possibilities for improvements. Additional beam studies and simulations should be done at low current to insure that the cascade performance under NLC conditions will be acceptable. Also, more work should be done to determine an optimal noise design using the results of recent ground motion studies. Simulations show that any feedback system will amplify incoming jitter at some frequency, but in the design of the feedback system we have some control over the frequency and magnitude of the amplification. The goal would be a design which damps at high-jitter frequencies and minimizes amplification while providing a robust feedback system.

SLC experience has shown that operational considerations are just as important as noise response. Feedback systems which decouple different areas of the machine minimize the invasiveness of tuning procedures, allowing downstream programs to continue by automatically stabilizing the beam. Operators are freed from many routine responsibilities, allowing time for more subtle tuning. The generalized design of the feedback system has allowed extension to many unplanned applications. Feedbacks are integrated with optimization packages; for example, feedback set-points are controlled to minimize beam emittance. Finally, the system improves machine reproducibility, supporting easier startup after outages and improving machine efficiency.

References

[Hendrickson 1995] L. Hendrickson *et al.*, "Fast Feedback for Linear Colliders", *Proc. 1995 Part. Acc. Conf.* (1995).

[Himel 1991] T. Himel *et al.*, "Use of Digital Control Theory State Space Formalism for Feedback at SLC", *Proc. 1991 Part. Acc. Conf.* (1991).

Contributors

- Linda Hendrickson