SLAC–PUB–5597
July 1991
(A)

LINEAR COLLIDERS*

EDITED BY R. D. RUTH

Stanford Linear Accelerator Center,
Stanford University, Stanford, CA 94309

Invited paper contributed to the 1990 DPF Summer Study on
High Energy Physics: Research Directions for the Decade,
Snowmass, CO, June 25–July 13, 1990.

---

* Work supported by Department of Energy contracts DE–AC03–76SF00515, DE–AC05–85ER40216, W–7405–ENG–48, and W–7405–ENG–36.

LINEAR COLLIDERS[*]

D.L. Burke, P. Chen, S.D. Ecklund, J. Irwin, R.B. Palmer[†]
J.M. Paterson, R.D. Ruth, J.T. Seeman, P.B. Wilson
Stanford Linear Accelerator Center, Stanford, CA 94309

W. Kozanecki, W. Schnell
CERN, Geneva 23, Switzerland

W.A. Barletta
Lawrence Livermore National Laboratory, Livermore, CA 94550

V. Granatstein
University of Maryland, College Park, MD 20742

H.-G. Moser
Max Planck Institute for Physics, 8000 Munich, Germany

R.E. Shafer
Los Alamos National Laboratory, Los Alamos, NM 87545

OUTLINE

---

[*] Work supported by Department of Energy contracts DE-AC03-76SF00515, DE-AC05-85ER40216, W-7405-ENG-48 and W-7405-ENG-36.

[†] Also at SSC Laboratory, Dallas, TX 75237.

1. Introduction

1.1 ENERGY AND LUMINOSITY GOALS

Past and present colliding beam machines have provided us with spectacular views of the interactions of quarks and leptons at center of mass energies from a few GeV up to the masses of the gauge bosons that mediate the electro-weak forces of nature. Experiments have confirmed that the Standard Model of these interactions provides an accurate description of particle physics in this energy domain. Our confirmation of this model, however, is incomplete until we have determined the nature of the breaking of the exact SU(2) X U(1) symmetry that is at its base. It is widely believed that to accomplish this goal, it will be necessary to explore particle physics at center of mass energies well above the masses of the W and Z bosons. In the regime $s >> m_Z^2$ where these particles become "light", it is expected that the true source of the symmetry-breaking will be found. Either new particles must appear, the simplest possibility being a single neutral Higgs scalar, or the weak interactions between gauge bosons must begin to become ever stronger until some new physics emerges to control the behavior of these forces.

A thorough discussion of the physics opportunities provided by e^+e^- colliders with center of mass energies at the TeV scale is given in another section of these proceedings[1]. Searches for neutral scalars and studies of their spectra, studies of the top quark, and detailed examination of the three-gauge-boson vertex can be done with excellent precision with a collider operating at a center of mass energy of 500 GeV. Such a machine would also be sensitive to some phenomena, such as the properties of neutral gauge bosons, with mass scales above 1 TeV. Data samples of 5 fb^{-1} are sufficient to begin these studies, while full exploitation of this machine will require the accumulation of 50 fb^{-1} or more. This machine should, therefore, be able to reliably deliver $\geq 3 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ shortly after turn-on, and be able to reach $\approx 3 \times 10^{33}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ after several years of operation. It is important for much of this physics program that the energy spread of the beam-beam collisions be kept below a few percent.

Additional physics may be found at still higher center of mass energies, and we consider it imperative that the design of the NLC be such that the beam energy can be increased over time to reach $E_{cm} \approx 1.5$ TeV. Such a machine will provide an energy reach approximately equivalent to that of the SSC. An e^+e^- collider is an ideal instrument for the study of the electroweak interactions at high energies. Production of new particles is easily observed, and their spectra and decay properties can be studied in detail. Deviations of, for example, WW scattering cross sections from expected values can be precisely determined. The experimental program of a 1.5 TeV collider will require data samples of $\sim 30 fb^{-1}$ for initial success, and samples of several hundred fb^{-1} will be required to fully explore the physics. Control of the energy spread of the beam-beam interactions is not so important for the physics program in this energy range, and considerable beamstrahlung can be tolerated if backgrounds from low energy photons and electron-positron pairs created in the collision can be eliminated by appropriate masking.

1.2 DESIGN PHILOSOPHY[2 – 5]

A factor of 5–10 energy increase beyond the SLC can be obtained in two ways: by increasing the collider length to 10–20 times that of the SLC (3 km), or by increasing its accelerating field to 5–10 times the SLC gradient (20 MV/m). The present consensus is that we should first increase the accelerating field by about a factor of 5 — up to about 100 MV/m. To limit the RF power required, this field should be provided by structures similar to those used in SLC but at a higher RF frequency of 10–30 GHz. At SLAC and KEK, the frequency choice for the NLC is 11.4 GHz, or four times the present SLC frequency. INP is pursuing a design with $f_{rf} = 14$ GHz; CERN is working on CLIC, a two-beam design at 30 GHz; and DESY and Darmstadt have recently teamed up to revisit 2.8 GHz, the SLAC frequency, for application to a 0.5-TeV design. At fixed RF frequency, the tradeoff between collider length and accelerating field is governed by the overall cost. A broad optimum occurs at the point where the linear costs (accelerating structure, magnets, tunnel, etc.) equal the cost of providing the RF power. The choice of technology to obtain the energy will also be governed by cost and the potential for upgrading the energy from 0.5 TeV to 1.5 TeV.

The choice of luminosity range also greatly influences the design of the linear collider. In principle, one could increase the luminosity simply by raising the repetition rate of the accelerator, but the wall-plug power increases in direct proportion. In a reasonable design, the wall-plug power should fall in the range 100–200 MW. Given this constraint, the best way to increase the luminosity is

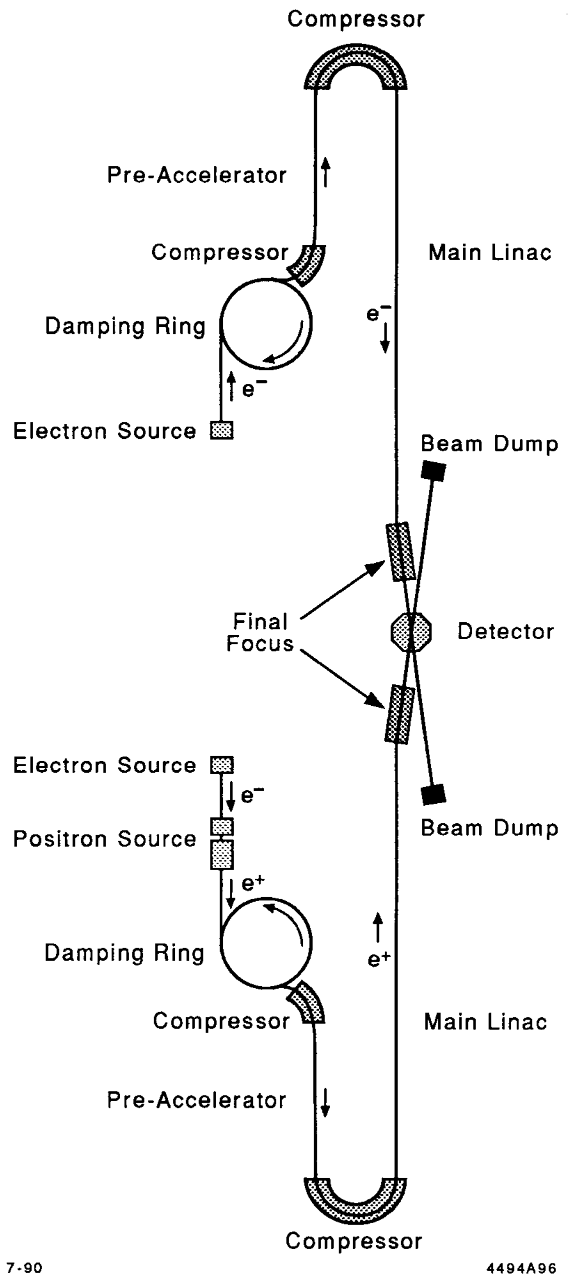

Fig. 1. NLC Layout

to shrink the beam size at the interaction point (IP). In addition, the beam cross section must be kept flat at the IP in order to minimize the amount of "beamstrahlung" radiation emitted as energetic electrons or positrons interact with the electromagnetic field of the opposing bunch.

The luminosity can be further enhanced by accelerating several bunches on each machine cycle. A single bunch of particles can in practice extract only a few percent of the energy available in the accelerating structure, without introducing an intolerable energy spread in the bunch. With additional bunches we get both greater luminosity and higher efficiency of energy transfer to the beam. The number of particles in each bunch, another factor that directly affects the luminosity, is limited by the RF energy that can be stored in the accelerating structure and by the amount of beamstrahlung radiation that can be tolerated. The obvious solution is to generate trains of successive bunches, each with fairly moderate numbers of particles.

Given these goals and constraints, we can now sketch a rough design of a linear collider able to achieve both the desired energy and luminosity. A possible layout is shown in Fig. 1. There are two complete linear accelerators, one for electrons and the other for positrons. Each linac is supplied with particle beams by a damping ring followed by a pre-acceleration section consisting of two bunch compressors and a pre-accelerator linac. After passing through the main linacs and final focus system, the beams collide at a small crossing angle inside a large particle detector like the present SLD.

To illustrate the basic features of the NLC operation, let's follow some electron bunches through the collider. A sequence of 10 bunches or so is created at the source and accelerated up to about 1.8 GeV in a pre-accelerator. This "batch" of bunches is then injected into a damping ring that serves to reduce the transverse and longitudinal phase space occupied by the electrons in each bunch. At the proper moment, these bunches are extracted from the ring and then compressed along their direction of motion by a bunch compressor, after which they are accelerated in a conventional linac and compressed in length a second time just prior to injection into the main, high-gradient linac. The entire batch is carefully steered and focussed as the electrons are accelerated up to full energy in the linac. Precision magnets in the final focus system squeeze the bunches down by about a factor of 300 just before they collide at the IP with similar bunches of positrons. Except for the fact that they were created differently (from the shower of particles that occurs when a beam of electrons from a special linac hits a metal target) these high-energy positron bunches have followed a similar evolution. After the beams collide, their debris is channeled out of the detector area and into shielded dumps.

In the remainder of this paper, we begin with a review of SLC performance. This is followed by a section detailing how the energy of the collider is obtained. The next section discusses those issues which contribute to the luminosity. Finally, we present a brief outlook. In all sections we have attempted to select a representative, but not exhaustive, set of references.

2. SLC Performance and Fundamental Limits

The Stanford Linear Collider[6] (SLC) was successfully brought into operation in 1989 and has completed the first round of accelerator physics and particle physics experiments. The SLC is the first scientific prototype of a new breed of colliders which will collide beams from two opposing linacs. The second round of experiments on the SLC is to begin in the fall of 1991. A continuing accelerator improvement program[7] is underway to produce a sizable luminosity with polarized electrons in 1991. The new particle physics detector, SLD, will be installed. Accelerator studies of small spot sizes, low emittances, and high currents will be done.

2.1 PROJECTED PERFORMANCE

The performance level of the SLC has increased steadily since construction was completed in 1987. The present performance uses bunch intensities up to 2×10^{10} e^-, spot sizes on the order of 3 μm at 120 Hz repetition rate. The peak luminosity is 3×10^{28} $\mathrm{cm^{-2}s^{-1}}$. The projected performance for the fall of 1991 will double the intensities, reduce the spot size to around 2 μm and achieve a luminosity on the order of 6×10^{29} $\mathrm{cm^{-2}s^{-1}}$.

The present four (major) practical limitations are detector backgrounds, unscheduled downtime (30%), positron yield (0.5 of desired), and difficulties of the injector chain to produce stable electron beams above 4×10^{10} e^-/bunch. Major improvement programs are underway to significantly reduce these limitations and also provide polarization. The projects include polarization equipment, nondisruptive emittance monitors, better accelerator protection systems, global fast feedback systems (correction every pulse), improved automatic tuning, and improved reliability of power supplies and kickers. These projects are well advanced.

2.2 FUNDAMENTAL LIMITS

During the planning, construction, and operation of the SLC, many practical and fundamental limits were addressed. In this section, the fundamental limits and issues which affect the next linear collider are discussed.

Longitudinal and transverse wakefield effects have been observed and controlled. The wakefield effects agree with the predictions indicating that calculations for the NLC should hold. New controls and cancellation methods (BNS damping) for transverse wakefields have been successful. Recently-discovered methods will likely provide improved control, e.g. autophasing[8, 9], dispersion-free trajectory correction[10] and harmonic linac movers[11]. Also, the acceleration of three bunches (e^+, e^-, e^-) on each RF pulse has been very successful. The energies of the three bunches are kept stable to about 0.2% by active feedback. The energy spectra are also stable at 0.35%. Reasonable control of a small number of bunches in the NLC is thus expected.

Pulse-by-pulse feedback systems (corrections made every RF pulse) for beam energy, position, and angle work very well. These systems regain the stability lost when colliders changed from circular to linear. Feed-forward systems to compensate beam properties on the *same* RF pulse are coming into use. Both feedback and feed-forward systems will be vital to the NLC and should work well if carefully planned. Application software, like feedback, is essential for linear collider commissioning and operation and must be a major initial activity.

Trajectory measurement and correction are now routine, allowing, for example, the SLC linac trajectories of both e^+ and e^- to be simultaneously reduced to below 100 μm rms. Lattice models describing many betatron wavelengths need special attention for phase mismatches. Beam-based alignment methods using magnet strengths and beam trajectories[12] have been shown to provide corrections of position monitor and quadrupole errors to below 100 μm. Component vibrations can cause real-time changes of beam trajectories. Present vibration[13] is controlled to about 0.1 μm. Feedback systems can reduce these effects up to frequencies of about 1/6 the pulse rate. Beyond that, passive or active vibration isolation is needed.

Finally, RF deflections of beams in linacs can be significant for trajectory and emittance control. Potential deflections in the NLC will place constraints on RF structure design.

Polarized beams are required for certain physics studies[14]. Work is required on reliable, high intensity, high polarization, long lifetime photo cathode sources.

Emittance preservation at the level of $\Delta\gamma\epsilon = 1 \times 10^{-5}$ m has been demonstrated during acceleration by a factor of 42. Damping ring designs seem adequate. Bunch lengthening and multibunch effects, however, need attention. Flat beam studies in the existing linac are needed to verify our understanding of the NLC.

The production of small spots in the final focus is now routine, fixing incoming dispersion, betatron, and chromatic effects. The NLC designs are reasonable extensions of the SLC design. Efforts to reduce the setup time through automatic procedures are needed.

Profile monitors are essential for good collider operation. Both two-dimensional intercepting monitors (screens) for halo studies and projection monitors (wires) for core studies are needed. Because of the small sizes in the NLC, new initiatives are needed for profile monitor designs.

Reducing detector backgrounds is a continual process at the SLC. The backgrounds come from far off-axis particles which either strike the vacuum chamber near the detector, make synchrotron radiation near the IP, or produce muons in the final focus collimators. Clean betatron and energy collimation is important to remove these off-axis particles. The sources of these particles include bunch generation in the damping ring, bunch compression, linac wakefields, edge scattering from primary collimators, and, to a small extent, beam-gas scattering. A multi-faceted collimation/masking system has been developed at SLC.

2.3 CONCLUSIONS

The parameters of the SLC are constantly being improved. Polarized collisions at sizable luminosity is expected in the SLC in early 1992 after a two-year period of major upgrades.

Experience at the SLC indicates that beam diagnostics, emittance control, tolerances, and automated tuning procedures need special attention for the NLC.

3. Obtaining the Energy in an NLC

3.1 LENGTH VS. GRADIENT

In the Introduction, the energy and luminosity goals of the NLC were defined and the philosophy of design was described. For this discussion, we will assume that the linear collider will begin operation with a center of mass energy of 0.5 TeV and will later expand its energy reach by a factor of three. This expansion in energy would necessarily need a luminosity which increases as E^2 and could be realized as a gradual continuous upgrade of the facility or as discrete steps with major changes to the accelerator complex.

The accelerating gradient which one chooses for an NLC is a parameter which impacts on almost every other parameter in the design. It is important in determining the necessary RF power sources, the overall size of the accelerator, its expansion capability and the cost. Over the past few years, there have been studies of designs with accelerating gradients ranging from 20 to 200 MeV per meter, using RF frequencies between 3 and 30 GHz. The lower gradient and frequency correspond to existing technology, e.g. the SLC, and the higher values are the subject of extensive worldwide research and development activities. The possible limits and problems with these parameters are more fully described below, but the present choice in the US, Europe, Japan and the USSR is for gradients less than or equal to 100 MeV per meter. The RF frequencies are now chosen to be at or below 14 GHz except for two-beam accelerator schemes which are discussed later in Section 3.4.

Let us now examine several approaches to the design of an accelerator complex which has an initial goal of 0.5 TeV in the center of mass, and which has an energy expansion capability of at least a factor of three. This forces a choice between three different strategies which make different assumptions in technology development, accelerator physics, time frame, and cost.

First let us consider an extension of today's technology. Two extended SLC-type linacs (3 GHz and 17 MeV/meter), each 20 km in length, could achieve the desired first-stage energy, but there are problems with beam stability and total power consumption. These are not impossible to design around and, using existing designs for damping rings and final focus systems, a satisfactory luminosity could be achieved. To expand the energy, one would either have to increase the length of the accelerators or increase the gradient, both of which would increase the power consumption and cost to extreme values. An alternative would be to upgrade the accelerator, using the as-built conventional facilities and many support systems, by replacing the RF systems and accelerator structures with higher frequency and higher gradient units. This long low-gradient approach has several difficulties, including cost, schedule and the delay in gaining experience with the new accelerator technologies. It has the advantage, however, that it could be built using existing power sources.

An alternate but still conservative strategy, which takes into account the considerable progress over the last few years with higher frequency RF, would be to build a long, low gradient—but high frequency—collider. For example, two linacs, each 5 km in length operating at 12 GHz, could achieve the initial energy of the NLC with low gradients and expand into the TeV range without exceeding a 100 MeV per meter gradient. The initial demand on the peak power required from the RF sources is modest and the energy extension would be adiabatic as these sources were further developed. Here, as in all designs, the challenge is the luminosity in the initial stage. A low-gradient long accelerator has problems with beam instabilities and emittance control. If these are controlled, then as the gradient is increased, the requisite luminosity increase comes relatively easily. This comes about because in addition to the decrease in absolute emittance with increasing beam energy, the current per bunch can increase while maintaining constant the fraction of the energy extracted from the structure; the higher the gradient, the higher the stored energy in the structure.

The third strategy is to build a shorter version of the TeV collider, then increase its length. Several studies have shown that this is a viable approach, where one increases the length of the high gradient accelerator without significant changes in the injection, damping ring, or final focus systems. This is a folded-back design with all of the above in a central complex. Moderate energy damped beams are transported to the beginning of the main accelerator where they are accelerated back towards the central complex of this facility (see Fig. 1). Energy extension then involves extending the length of the accelerator and conventional facilities, tunnels, utilities, etc. away from the center. With this approach, one attacks the technologies of high power RF generation and high gradient acceleration in stage one. One would learn much about the problems of going to higher energies. There is, of course, the risk of tackling too many new problems at once.

The correct strategy will probably be a judicious choice combining the above ideas. A moderately conservative accelerating gradient, built-in expansion capability in energy as a design criterion in all systems, and a high-luminosity parameter set for the initial lowest operating energy.

3.2 LIMITS ON ACCELERATING GRADIENT

Consider the copper surface of an accelerator structure prior to the application of an RF field. The surface will depart from an ideal smoothness due to topographic features such as machining marks, microprotrusions and discontinuities at crystal grain boundaries. The surface will depart from chemical purity due to exposed pockets of solid impurities (typically at grain boundaries), adsorbed gases, and layers of condensed dielectric contaminants. The RF field can be enhanced in localized areas by a factor of a hundred or more, due both to these geometric perturbations and to the layers and patches of surface contaminants[15]. The RF fields that can be reached at such a surface without breakdown are initially about 80 MV/m. In a procedure known as "RF processing," the field level can be raised gradually as the surface becomes smoother and cleaner as surface contaminants are "scrubbed" away by electron and ion bombardment. The rate at which the cavity or structure processes to a high field level can vary widely, depending on the quality of the original copper, its mechanical treatment (machining, polishing), its thermal history (annealing, brazing), the chemical cleaning procedures used, and its vacuum history.

The ultimate surface electric field that can eventually be reached does not seem to be too sensitive to the details of surface smoothness and purity, but there is evidence that the rate at which this level can be approached is definitely faster if one starts with a smooth and highly-polished surface. In a series of measurements using short standing-wave sections of disk-loaded structure in the 3–10 GHz frequency range, Loew and Wang [15, 16] have found that the maximum surface electric field as a function of frequency that can be attained after processing follows the relation

$$\mathrm{E_{pk}(MV/m)} \approx 195[\mathrm{f\ (GHz)}]^{\frac{1}{2}}.$$

Since this is the peak surface field (usually near the tip of a loading disk), the actual accelerating gradient is obtained by dividing this number by a factor in the range 2.0–2.5 for a typical travelling wave structure. Furthermore, since the field emission current increases exponentially as the surface field approaches $\mathrm{E_{pk}}$, there may be too much "dark current" to be tolerable, even at fields somewhat below the breakdown level. Here again, there is some evidence that the ultimate field-emitted currents below the breakdown level can be reduced if one starts with a reasonably smooth and pure surface.

At 11.4 GHz, assuming also a peak to average field ratio of 2.3, the above relation predicts an accelerating gradient at breakdown of 280 MV/m. Measurements on short resonant structures, however, may not give a true indication of the performance of an actual travelling wave structure on the order of a meter in length. In recent experiments at KEK, a gradient of 85 MV/m has been attained at 2.85 GHz. This is just over one half of the breakdown gradient expected on the basis of the above expression. Already at this field level, a dark current beam of close to 20 mA was measured. There is hope that changes in the design of the structure and couplers can reduce this value. In an experiment at 11.4 GHz, the output of the SLAC/LLNL relativistic klystron was used to drive a structure 25 cm in length to a gradient of 84 MV/m without breakdown. Some field emission current was evident ($\approx 30\ \mu$A), but the structure was not well processed. In summary, measurements to date indicate that a gradient of 100 MV/m at 11.4 GHz can probably be reached without breakdown. More experiments on realistic accelerating structures must be done to see whether the background current from field emission can be held to a tolerable level.

3.3 RF SOURCES, STRUCTURE, COMPRESSION

3.3.1 RF Sources for Linear Colliders

The next linear collider will likely operate at an X-band frequency (e.g. 11.4 GHz) since this gives a reduced accelerator length and reduced average power consumption compared with lower frequencies. Compared with higher frequencies, the development risk of the required microwave tubes is reduced, and the tubes could likely be developed within two to three years.

We choose to use a microwave pulse compression system after the microwave tubes (e.g. SLED-II, which compresses the pulse length 16× and multiplies power by 8–10×). Producing the required power level directly in each microwave tube would imply a tube voltage $\gtrsim$1 MV, and would require excessively expensive power supplies. Required performance parameters of the X-band microwave amplifier tubes, assuming the use of pulse compression, are as follows:

Peak power = 50–150 MW
(0.5–1.5 GW available after pulse compression)
Pulse duration $\tau \approx 1.6\ \mu$s
(assuming 16× pulse compression will be used)
Gain $\approx$ 50 dB
Repetition rate $\approx$ 120-180 Hz
Efficiency $\gtrsim$ 40%
(the higher the better)
Voltage $\lesssim$ 600 kV

The combined peak power and pulse duration requirement are beyond the capability of conventional X-band klystron designs because of RF breakdown in the output cavity gap. Modified klystron configurations or new tube concepts are called for. A summary of X-band microwave amplifier development projects with $\tau \approx 1\ \mu$s is presented in the following paragraphs.

A development program at SLAC is presently aimed at producing a 100 MW X-band klystron amplifier. In initial experiments with a single gap in the output cavity, output power has been limited to about 20 MW at $\tau \approx 1\ \mu$s. A subsequent klystron with a double output gap produced about 35 MW[17, 18]. Further studies aimed at achieving the full design power will include testing additional tubes with multiple-gap output cavities and perhaps tubes in which the output cavity is replaced with a short travelling-wave section. The successful operation of a travelling-wave output section at powers above 100 MW, albeit with short pulses (<100 ns), has been demonstrated in joint SLAC/LLNL/LBL experiments, in experiments at MIT on a Haimson Research klystron (see below), and also at other laboratories (e.g., Cornell University, and the Institute of Applied Physics in Gorky, USSR). In connection with construction of the VLEPP collider at Protvino in the USSR, a development program is underway which aims at realizing 130 MW klystron amplifiers operating at 14 GHz; initial experiments have produced power ~50 MW. Klystron development is also underway at KEK (Japan) and at Haimson Research (USA). Recently, a Haimson Research klystron[19] with a travelling wave output structure has achieved a 100-MW output power at a FWHM pulse width of 50 ns with an efficiency of about 45%. In general, klystron efficiencies of about 50% can be expected with further development.

An alternative approach to klystrons is embodied in tubes such as the gyrotron and magnicon which exploit coherent cyclotron emission and require no small cavity gaps with high peak surface fields. This approach is being explored at the University of Maryland[20, 21, 22] and at the

Institute of Nuclear Physics (Novosibirsk). The Maryland program is presently aimed at producing a 30–50 MW, X-band gyrotron amplifier (i.e. a gyroklystron). Recent operation of this device has produced output power of 20 MW at 30% efficiency with a gain of 26 db at a pulse length of 1 μs[23, 24]. Improvements are in progress which are expected to increase both the power and efficiency. The design efficiency of 42% is somewhat smaller than expected for a klystron. Scaling of the X-band gyroklystron design to the 100-MW level may be feasible, and use of a depressed collector would enhance efficiency. Work at Novosibirsk is centered on developing a 7-GHz, 60-MW magnicon (which injects a rotating pencil beam into a gyrotron-like output cavity); magnicon efficiency is predicted to be 60% and the device is in the fabrication stage.

Linear colliders in the 3–5 TeV range will require higher RF frequencies, perhaps $\gtrsim$ 20 GHz. The gyroklystron has potential for operating in this frequency range, for example, by operating the output cavity at the second harmonic of the cyclotron frequency. Other tube concepts which may be well matched to high frequency operation include the cluster klystron and some form of sheet beam amplifier (e.g. sheet beam klystron). The two-beam accelerator configuration, which is discussed below in Section 3.4, would eliminate separate microwave tubes and is also a candidate concept for a 3–5 TeV linear collider.

3.3.2 RF Pulse Compression

RF pulse compression is a means for increasing the peak power from an RF source at the expense of a reduced pulse width. Assuming that the output pulse width of the compression system is equal to the filling time, T_f, of the accelerating structure, then the peak power after compression is

$$\hat{P} = P_k \eta_c (T_k/T_f).$$

Here P_k and T_k are the peak power and pulse width of the klystron (or other RF source), and η_c is the compression efficiency. Since the typical output pulse width of a klystron driven by a pulse modulator is, in general, considerably larger than the structure filling time, the power gain from a pulse compression system can in principle be quite large.

The SLED scheme[25], currently in use at SLAC, is one method of RF pulse compression. In the SLED method, energy is stored in high Q resonant cavities over a relatively long period(T_k = 3.5 μs) and released during the relatively short structure filling time (T_f = 0.8 μs). The compression efficiency of SLED is 62%, giving a power gain of about 2.7.

A new method of RF pulse compression, the Binary Energy Compressor[26] (BEC), is currently being tested at SLAC. In this method of pulse compression, the low level RF input drive pulses to two klystrons are divided into 2^n time bins, where the width of each bin is equal to the desired output pulse width and n is the number of compression stages. A phase modulation of 0 or 180° is then applied to each bin, following a specified modulation pattern for each of the two klystron inputs. At each stage of the compression, a 3 db directional coupler acts as a switch. Triggered by the phase coding, the coupler shunts the first half of the combined input pulses into a delay line so that they become coincident in time with the last half of the combined pulses at the output of the line. At the final output of the system, all of the time bins are, in effect, stacked on top of each other to give a power gain of $(2^n)\eta_c$, where the compression efficiency η_c takes into account losses in the delay lines and other components.

The system undergoing testing at SLAC is an 11.4 GHz three-stage BEC, with all components designed for high vacuum and high power, delivering about a 50 ns flat-top output pulse. The gain at low power has been measured to be 5.8, giving a compression efficiency of 5.8/8 = 73%. Recent measurements at high power (25 MW input, 120 MW output) show a power gain of about five, although more work remains to be done to optimize the phase tuning and pulse flatness[27]. This compression system, driven by two such klystrons with an output power of 30 MW, will deliver sufficient power to drive a 1.5-m long accelerating structure to a gradient of 100 MV/m.

The original SLED method is limited by the fact that the output pulse shape is a sharply-decaying exponential which varies by a factor of two or more over the structure filling time. Not only is this pulse shape undesirable in itself for driving an accelerating structure, but it has the consequence that stages of SLED cannot be ganged effectively to produce still higher power gains. This difficulty is overcome in a refinement of the SLED concept in which the resonant cavities are replaced by lengths of low loss delay line. In this SLED-II scheme[28], a single stage can reach a power gain of about four with a compression efficiency of about 70% (100 ns output pulse at 11.4 GHz). As a typical example of a two-stage system at 11.4 GHz, a 1.6 μs pulse can be compressed to 100 ns with an efficiency of about 55% (power gain of 8.8).

One criticism of both the BEC and SLED-II compression systems is the awkward lengths of the delay lines. For example, the longest delay line in the two-stage SLED-II system mentioned above is about 40 m long. Using loading irises, however, it appears feasible to reduce the group velocity and hence the line lengths by a factor of ten or so, with only a modest increase in loss and some degradation in pulse rise time and flatness. A pulse compression system related in concept to SLED-II is under development for the VLEPP project at Protvino[29]. Instead of two parallel loaded delay lines attached to the output of a 3 db hybrid coupler, this scheme employs a single chain of coupled high-Q "open" resonators with an azimuthally rotating mode.

3.3.3 Accelerating Structures

The accelerating structure for a next-generation linear collider is expected to be similar to the well-known $2\pi/3$-mode disk-loaded structure as used in the present SLAC linac, except (as will be discussed later) for modifications to damp higher-order modes. The key parameter which determines the performance of such a structure is the ratio of the radius of the iris opening to the wavelength, a/λ. A low value for this ratio gives better structure performance (more accelerating gradient per unit of stored energy), but unfortunately also stronger longitudinal and transverse wake potentials. The longitudinal wake potential determines the energy spread within a single bunch and the bunch-to-bunch energy change in a bunch train. The longitudinal wake scales roughly as $(a/\lambda)^{-2}$. The transverse wake potential determines the head-to-tail emittance growth within a single bunch, and leads to beam breakup in a multibunch train. The transverse wake scales roughly as $(a/\lambda)^{-3}$.

The value of a/λ for the present SLAC structure is 0.11. A value this low, however, would lead to an intolerable single bunch energy spread and emittance growth in a next linear collider. Values for a/λ in the range 0.14 to 0.20 are presently being considered in various collider parameter lists. To calculate some specific parameters, we take $a/\lambda = 0.175$ as representative. This results in a significant performance loss compared to the SLAC structure (20% more RF energy per unit length is required to reach a given accelerating gradient), but wakefield effects are greatly reduced.

For $a/\lambda = 0.175$, the group velocity is $0.06c$. The length of the structure must then be adjusted to make efficient use of the RF power. At 11.4 GHz this results in a structure length per RF feed of 1.5 m, a filling time of 83 ns and a peak power requirement of 220 MW/m for an accelerating gradient of 100 MV/m. These parameters will be used in the RF system example in the following section. Scaled to another frequency at the same gradient while keeping a/λ constant, the filling time and structure length would vary as $\omega^{-\frac{3}{2}}$, the peak power requirement as $\omega^{-\frac{1}{2}}$, and the stored energy per unit length as ω^{-2}.

A simple SLAC-type disk-loaded structure, but with the larger a/λ, would work fine for accelerating single bunches in a collider. For a train of 10 or 20 bunches, however, the long-range wakefields would produce unacceptable energy and transverse emittance differences between consecutive bunches. These long-range wakefields are mainly due to a dozen or so higher longitudinal modes, and the lowest few dipole modes. Various structures having appropriate slots and apertures for damping these unwanted modes, and some radically different "open" accelerating structures, are under consideration at both KEK and SLAC[30]. The damping must be strong enough so that the wakefields decay significantly between successive bunches for both the longitudinal and transverse modes. Another way to say this is that the Q's for the modes must be reduced to the order of 20 (for the lowest dipole mode) to 50 (for lower-order longitudinal modes), while leaving the accelerating mode substantially unaffected. Calculations and measurements on models have shown that this can be achieved. A practical structure design, however, must take into account many other considerations such as ease of fabrication, tolerances (will tuning of each individual cell be required?), mechanical support, and adequate vacuum pumping. As implied in Section 3.2, the need to control field emission and dark currents may also influence the structure geometry and the required surface finish, and may necessitate a strict protocol for subsequent surface treatment and cleaning procedures.

An alternate method for reducing the effect of long-range wakefields is to introduce a spread in the frequencies of the unwanted modes, while maintaining a constant accelerating mode frequency. Initial calculations[31] indicate, for example, that a 15% spread in the frequency of the lowest dipole mode is almost sufficient to reduce multibunch emittance growth to a tolerable level. Probably a mix of the two techniques, some mode damping and some mode detuning, will be used in the final structure design.

3.3.4 *Example of an 11.4-GHz RF System*

In this section we combine results from several preceding sections to give rough RF system parameters for an 11.4 GHz linear collider. We assume that klystrons having a microperveance of 1.2 can produce 150 MW of RF output power at an efficiency of 50%. Two 75-MW klystrons could, of course, be used if a single klystron could not produce this much power. We assume that the klystrons are powered by a modulator with a 25:1 turns ratio pulse transformer, producing 575 kV at the klystron cathode for 1.6 μs. The klystron in turn drives a two-stage SLED-II pulse compression system with a power gain of 8.8, an output pulse length of 100 ns and a compression efficiency of 55%. The peak output power after compression of 1.3 GW is sufficient to drive four accelerating sections, each 1.5 m in length having $a/\lambda = 0.175$, at a gradient of 100 MV/m.

The modulator, klystron and RF pulse compression system, together with the four accelerating structures, comprise one RF module. A 1-TeV collider would require 10 km of active structure length and 1667 such modules.

An important question is the efficiency of the RF system in converting wall plug power to power delivered to the accelerating sections. Although the structure filling time is 83 ns, we have assumed a 100-ns pulse length to allow for switching time in the pulse compression system and for timing errors. The energy in the RF pulse is thus 220 MW/m × 6 m × 100 ns = 132 J. The modulator efficiency is estimated to be about 55% for a 1.6-μs pulse, using the 25:1 pulse transformer mentioned previously. The overall RF efficiency is then

$$\eta_{rf} = 0.50\ (\text{klystron}) \times 0.55\ (\text{modulator}) \\ \times 0.55\ (\text{pulse compression}) = 0.15\ .$$

The energy per pulse per RF modulator at the wall plug is therefore 132 J/0.15 = 880 J. For a repetition rate of 120 pps, the average power per module is 105 kW. This times 1667 modules gives a total wall-plug power of 175 MW. Adding on a 20% overhead to allow for multibunch operation, correction of the BNS damping energy spread, etc., we arrive at a total linac length of about 12 km and a wall-plug power of 210 MW. For a center-of-mass energy of 0.5 TeV, the accelerating gradient is reduced to 50 MV/m, the klystron power to 40 MW and the wall-plug power to about 50 MW.

The collider energy could be increased to 1.5 TeV by increasing the linac length (corrected for overhead) to 18 km, and the number of RF modules to 3000. However, the wall-plug power would also increase to 315 MW. For the long-range future, it is possible to foresee multiple-beam RF sources (e.g. cluster klystron) or extended beam sources (e.g. sheet beam klystron) which operate at a lower voltage and higher efficiency. The lower voltage allows the use of a lower turns ratio pulse transformer in the modulator, which in turn leads to a faster pulse rise time and hence better modulator efficiency. Such sources would also be capable of higher peak RF output power and hence would allow a pulse compression system with lower power gain and better efficiency. The RF system efficiency is then estimated to be

$$\eta_{rf} = 0.70\ (\text{modulator}) \times 0.60\ (\text{klystron}) \\ \times 0.60\ (\text{pulse compression}) = 0.25\ .$$

The wall-plug power now becomes a more reasonable 190 MW.

In summary, the initial 0.5-TeV accelerator could be powered by a standard SLAC-type modulator and an X-band klystron under development at several laboratories. The SLAC version has already delivered close to the required 40 MW, although many engineering refinements still remain to be made. A new lower-voltage, extended-beam (or multiple-beam) RF source must be developed to power the 1.5-TeV extension. However, this is further down the road and sufficient time remains to do the necessary R&D. A 1.5-TeV, 18-km collider is, however, near the end of the line for 11-GHz technology. Extending the length could provide only a slightly higher energy at great expense, assuming a 200-MW limit on wall-plug power is maintained. To go to still higher energies, a higher RF frequency is required.

3.4 TWO-BEAM ACCELERATORS

The design of linear colliders at TeV energies has concentrated on scaling the structure to high frequencies (10–30 GHz). At X-band, one can probably employ multiple power tubes plus RF pulse compressors to power the collider at an acceptable capital and operating cost. Two-beam accelerators (TBA) seek to offer superior performance or similar performance at a substantially lower cost.

In TBAs, the kinetic energy of a high peak current drive beam accelerated with an induction linac or a superconducting RF accelerator is transformed to high peak power microwaves at 10–30 GHz with a free electron laser (FEL) amplifier[32] or with travelling wave transfer cavities[33], respectively. The TBA concept stems from several observations:

1. RF-power flow in the collider must be in the range of ≈150 MW/m (@ 30 GHz) to ≈250 MW/m (@ ≈10 GHz); therefore, the drive electron beam must have a power flow of several GW.
2. FELs operating at 35–140 GHz have demonstrated conversion efficiencies >35% and peak powers >1000 MW.
3. Unlike transfer cavity converters, the power flow in the FEL is independent of RF-wavelength.

In another variant[34] most suitable for X-band, the relativistic electron beam from the induction linac excites klystron-like transfer cavities. An experimental study[35] of such relativistic klystrons by a SLAC/LLNL/LBL collaboration has demonstrated high gradients of approximately 100 MeV/m, but also suggests that either travelling or standing wave transfer cavities tend to destabilize the low voltage drive beam rapidly as the drive beam current is increased beyond several hundred amperes. To surmount these difficulties one may revert to the FEL converter[36] or raise the drive beam energy to a few GeV[33] as in the superconducting variant discussed below. For FEL converters the beam-to-microwave conversion efficiency may be raised from demonstrated levels of ≈35% to ≈70% or more, if induction cells can re-accelerate the beam that has passed through the wiggler or klystron converter cavity. Re-acceleration of a bunched beam in induction modules raises questions[37] requiring both theoretical and experimental study. These issues, under experimental study at Pulse Sciences, Inc., include 1) parasitic loss of RF-power into the induction gaps and 2) destablization or degradation of the drive beam by longitudinal and transverse wakefields increasing the emittance during the deceleration-acceleration cycle (and degrading the FEL performance).

Because the induction structure is independent of the drive beam current, induction linac driven TBAs will have unacceptably high capital and operating costs[37, 38] unless a) the beam current can be raised to a few kiloamperes, b) the size of the induction modules are reduced, and c) the induction system is operated heavily beam loaded. Heavy beam loading makes voltage and current regulation in the induction linac a challenging and expensive task[39] and is under extensive study at LLNL. Without such improvements, it will be difficult for TBAs using induction linacs to compare favorably with more conventional power sources for the next linear collider that are projected to operate at moderate gradients 50–100 MeV/m at frequencies in the range of 10–15 GHz.

The third and rather different variant of the TBA[33] is being studied at CERN under the name of CLIC, CERN linear collider. For CLIC the drive beam is given an energy of several GeV and is periodically re-accelerated with CW superconducting cavities operating at a few hundred MHz after power is extracted from the beam with travelling wave transfer cavities. The high energy of the drive beams eliminates the difficulties of phase stability and phase shifts of the RF-power and permits a single drive beam to run along the entire high gradient linac. The superconducting re-acceleration cavities can be grouped into re-acceleration stations spaced by several hundred meters. Therefore, a large fraction of this TBA variant will consist of beam transport and passive structures only. This scheme has three basic problems.

1. The generation of a suitably intense and tightly bunched drive beam is somewhat beyond the present state of the art. A demonstration test facility is under construction at CERN.
2. A low-impedance transfer structure with acceptable wakefield properties must be developed. This work is ongoing and appears promising.
3. The somewhat extreme choice of 30 GHz for the high gradient structure is now mandatory to limit the stored energy in the system. The attendant problems of wakefields, fabrication and alignment tolerance for the high gradient structure are also under study. A first prototype of a 30 GHz structure has been produced and an alignment system with sub-micron reproducibility have been developed in the ongoing studies at CERN.

We conclude that two-beam accelerators will require a level of research and development beyond that which is needed for high gradient accelerators with multiple power sources. The TBA in any of the three variants may, therefore, not be ready in time for the next linear collider. The prospects of using TBAs to drive linear colliders are brighter for colliders with energies >2 TeV, which may require gradients >200 MeV/m and operate at frequencies well above X-band in order to keep size and power consumption within reasonable limits.

3.5 SUPERCONDUCTING LINACS[40]

Making the main linacs superconducting would seem to be an ideal solution in several ways. Since the superconducting cavities can store electromagnetic energy at very low loss, the peak power problem is completely eliminated. The high conversion efficiency permits very high values of beam power with corresponding relief of the final focus. Moreover, since relatively low frequencies (3 GHz or less) can be employed without loss of efficiency, wakefield effects will be small.

Unfortunately, reliably achievable average accelerating gradients (including dead space between cavities) are limited to values clearly below 10 MV/m at present, and the cost is high, reflecting the extreme care required in the manufacturing process. Much higher gradients have been obtained in laboratory tests but were, so far, limited to very short samples of structure and were not always well reproducible. A record surface field of up to 130 MV/m has been obtained[41]. This would, potentially, correspond to roughly half that value for peak acceleration but was, in fact, limited to less than a square centimeter surface in a test cavity not directly useful for acceleration. Also, a six-cell 1.5 GHz structure has reached 17.5 MV/m actual accelerating gradient (M. Tigner, private communication).

In addition, quality factors will have to be increased to operational values of 10^{10} or more if gradients substantially above 10 MV/m should become possible at tolerable cryogenic power. Even so, the linacs have to be slowly pulsed in order to reduce the power loss and the implications of this remain to be studied.

The fact that superconducting linacs present potentially the ideal solution is more than sufficient justification for a continued and vigorous effort of research and development in this area. This work should aim at bridging the large gap which now exists between record performance in small cavities and fully-engineered storage ring cavities operating at 5 MV/m at best. For the time being, it appears impossible to predict whether and when a fully-superconducting linear collider may become a realistic possibility.

4. Obtaining the Luminosity in an NLC

The luminosity can be calculated from the beam characteristics at the collider IP,

$$\mathcal{L} = \frac{N_{e+} N_{e-} f_{rep} n_b H_D}{4\pi\sigma_x\sigma_y}.$$

In this formula, $N_{e+}(N_{e-})$ denotes the number of positrons (electrons) in a single bunch; f_{rep} is the rate at which the accelerator is cycled, the repetition rate; n_b is the number of bunches accelerated on one cycle (i.e., in a single "batch"); H_D is an enhancement factor due to the mutual focusing of the beams at the interaction point; and σ_x and σ_y are the horizontal and vertical gaussian rms beam sizes at the interaction point.

The bunch size at the interaction point can be further separated,

$$\sigma = \sqrt{\beta\epsilon}$$

where ϵ is the emittance of the beam and β is the depth of focus of the focusing system.

In the next few sections, we discuss constraints and limitations on each of the various terms in the luminosity formula. The emphasis is to increase those terms in the numerator while decreasing those in the denominator in a consistent fashion.

4.1 SINGLE-BUNCH INTENSITY AND REPETITION RATE

The single bunch intensity in a linear collider is intimately related to the allowed wall-plug power:

$$eNEf_{rep} = \eta_{rf}\eta_b P_{wall}$$

where N is the number of particles per bunch, E is the center-of-mass energy, f_{rep} is the repetition rate, η_{rf} is the efficiency of wall-plug to RF power, η_b is the fraction of energy extracted by a single bunch. η_{rf} is about 20% for conventional RF sources. η_b is limited by the allowed spread in bunch energies in the final focus. It is typically a few per cent (we take 2% for our example).

The repetition rate has a lower bound due to the vibration of the ground. Since the amplitude of ground motion falls rapidly above 10 Hz, and since we need another factor of six to sample that motion for feedback, the lower bound on

f_{rep} is $\sim$ 60 Hz. Designs tend to be a factor of two or three higher than this number, with the exception of CLIC which is an order of magnitude higher. For our example, we take 180 Hz. Higher values have a large impact on the source and damping ring. The energy of the beam we take to be 0.5 TeV. Having specified these quantities, the number of particles per bunch is determined. In this example, it is 2×10^{10} for a total wall-plug power of 150 MW.

This number is entirely independent of the method of acceleration; it is an upper bound on the charge per bunch given the above limitations.

If the repetition rate is increased, then the intensity of the bunch must decrease given the power limit. Viewing this trade-off in isolation from the other limitations, the largest luminosity comes from a low repetition rate and high single bunch current.

Positron production can be accomplished via several methods. A scaled conventional "SLC"-type source is workable if adjustments are made to accommodate the higher pulse intensities (10 bunches $\times 2 \times 10^{10}$/bunch) desired. A higher-acceptance positron collection system combined with an additional pre-damping ring is an attractive improvement over existing sources. This would likely mean larger aperture, lower frequency (1 GHz) accelerator sections combined with strong focusing magnets. The higher beam energies in the NLC open other interesting possibilities. Positrons are usually made by pair production (nuclear beta decay has also been considered as an alternative), with a variety of potential ways to produce the photons. Synchrotron radiation from a beam of energy 150 GeV or higher provides photons adequate for positron-electron pair production. If generated from a helical undulator magnet, circularly polarized photons are obtained and give longitudinally polarized positrons. This capability may improve the analyzing power in an experiment. Other possibilities such as Compton back-scattered lasers or photons produced by channeling are also under consideration as an alternative to the usual electromagnetic cascade shower. While a challenge, adequate positron production is likely obtainable by several different techniques.

4.2 EMITTANCE PRODUCTION/PRESERVATION[33, 42 – 46]

The most important factor in increasing the luminosity in the NLC is the reduction of the beam sizes. Part of the size reduction comes from reduced emittances generated by damping rings with natural emittances about an order of magnitude below the SLC. In addition, many designs make use of the naturally flat beams produced by damping rings. It is generally agreed by experts in the field that the damping rings to produce these emittances ($\gamma\epsilon_x \sim 3 \times 10^{-6}$, $\gamma\epsilon_y \sim 3 \times 10^{-8}$) are straightforward and are rather similar to synchrotron light sources now under construction.

To prepare the bunches for acceleration in a high gradient linac, it is also necessary to compress the bunch length and to pre-accelerate the bunches at some longer RF wavelength to about 5–15 GeV. Both the bunch compression and pre-acceleration are straightforward extensions of present techniques.

After the bunches enter the main linac, there are many effects which conspire to dilute the emittance. To begin with, the optical functions, dispersions, beta functions, etc., must be matched to those of the linac transport system. If this is not done, filamentation due to the finite energy spread can dilute the emittance.

One of the dominant sources of emittance dilution and beam sensitivity is the transverse wakefield. This couples the head and tail of the bunch and can cause the tail to grow. This effect can be compensated by a technique called BNS damping (or autophasing for very strong wakefields). BNS damping uses a correlation between energy spread and bunch length to make use of the energy-dependent focusing of the linac to cancel the wakefield effects. This technique has been tested at the SLC where it dramatically decreases the sensitivity to wakefields.

However, even with BNS damping, due to the lack of complete cancellation at the local level, and due to the uncorrelated energy spread, there can be emittance dilution due to the sequence of dipole kicks by misaligned magnets and steering magnets. If left uncorrected, this can result in alignment tolerances which range from one to 10 μm depending upon the design. Fortunately, new trajectory correction techniques which have recently been studied can increase these tolerances well above the 100-μm range[10, 47]. If wakefields are weak by design, then the accelerator alignment tolerances are no worse than the quadrupole/BPM alignment tolerances; however, some designs have very large wakefields–comparable, in fact, to the external focusing force. In these designs, accelerator alignment is critical. A new technique of harmonic linac structure movement may significantly relax these alignment tolerances.

It is also necessary that upon exiting the linac, the bunch position be stable from pulse to pulse. It is possible to feedback on motion which occurs at frequencies less than about $f_{rep}/6$. For higher frequencies, the bunch must be passively stable to values less than the beam size. Fortunately, it is the local beam size that sets the scale for magnet vibration tolerances because the beam centroid motion is demagnified along with the size, provided that motion is induced upstream of the final focus system.

The tolerances for random vibrations of quadrupole magnets from pulse to pulse are a small fraction of the beam size in the linac. For flat beam designs, they are $\sim$ 40 nm. Fortunately, at frequencies of 30 Hz and above, natural ground motion is far below this value. Active and passive control will be needed for low frequencies.

4.3 MULTIPLE BUNCHES[48 – 50]

Extracting the energy from a linear accelerator is a rather inefficient process with a single bunch. This is due to the coupling between the energy extracted and the resulting energy spread induced in the bunch. It is possible, however, to extract a significant portion of the RF energy while maintaining good beam quality with the use of a train of bunches. This has the additional advantage of increasing the luminosity. Presently, the SLC makes use of this concept by accelerating both the electron and positron bunch on the same cycle of the SLC linac, as well as the electron bunch used to make positrons.

For the NLC, however, it is desirable to have many more bunches ($\sim$ 10) to extract a significant fraction of the RF energy. To obtain a useful train of bunches in this case, it is necessary to control both the bunch-to-bunch energy and the transverse stability of the bunch train.

The transverse stability can be destroyed by multibunch beam breakup. This must be controlled by providing RF structures in which the transverse wakefield decreases rapidly behind a bunch. This can be accomplished either by damping the higher order modes using external waveguides, or by providing a large cell-to-cell frequency change so that the higher order modes interfere destructively (see Section 3.3.3). Both of these techniques have been tested: Qs as low as 10 have been measured at SLAC in model structures[51].

Multiple bunches also impact the design upstream and downstream of the linac. In particular, one must assure the stability of the train in each system of the linear collider. The train of bunches places a special emphasis on the time stability of magnetic components in the system, especially kickers. New controls and monitors are needed to handle closely spaced bunches.

4.4 FINAL FOCUS AND CROSSING ANGLE[52 – 56]

4.4.1 Final Focus Issues

All final focus systems under consideration for NLCs have the same structure as the Final Focus Test Beam facility now under construction at SLAC. The FFTB system has been carefully analyzed over the last two years, and together with experience with the SLC final focus system, these systems are becoming well understood.

The NLC designs proposed by K. Oide and D. Helm are roughly twice as long (300 m) as the FFTB; they have system quadrupoles of about twice the strength, and a final quadrupole of about six times the strength. A final quadrupole with these specifications (0.5 mm half aperture, 1.4 Tesla pole tip field) has been built and measured at Protvino, USSR.

The main aberrations which control the design are the chromaticity, strong-long sextupole aberrations, and the chromo-geometric aberration coming from the chromatic failure of the minus I transform between sextupoles. Having chosen a design of length sufficient to control these aberrations, one must insure that feeddown aberrations, such as dispersion, normal and skew quad, and chromatic skew quad, are either compensated or minimized by controlling beam position at crucial points within the system. It is these feeddown aberrations which determine the alignment and stability tolerances of the system.

Recent analysis indicates that mechanical alignment tolerances can be quite generous (100 micron) as well as "BPM-center-to-element-center" tolerances (50 micron). What is required for the main diagnostic are BPMs located at maximum beta points throughout the system which are sensitive to changes in orbit position. When used in pairs located π apart, the sum signal of these BPMs can be used to adjust the linear system and position the beam at critical elements. When the linear system is tuned, BPM sum readings will not depend on upstream changes of the beam. For the 250 GeV NLC, if the BPMs have a scaling accuracy of about 1%, are free of cross-talk between the horizontal and vertical channels to about 1%, and can detect orbit changes of three microns, then the spot size will be within a factor of 10 of the design spot size and the final tuning can be readily accomplished with correctors.

To carry out final tuning, which is based on observation of final spot size, previous corrections must remain stable. Stability could be monitored by the BPM pairs, if a stable sensitivity to changes of 0.3 microns is obtained. Presently, BPMs in the SLC have the required scaling and cross-talk tolerances, however their relative motion sensitivity is limited to about five microns. FFTB sensitivity is specified to be one micron. A smaller NLC beam pipe diameter will be helpful, but further BPM improvements are needed.

At the SLC, the beam-beam interaction provides a distinct signal to determine beam offsets and spot sizes. This signal is expected to be present throughout the NLC parameter domain. The stability of the beam centroid at the IP is maintained by feedback from this beam-beam signal to steering coils adjacent to the final quads. The most sensitive tolerance required to maintain beam collision comes from the final quads. They must not move relative to one another by more than the final spot size (2 nm). Since the feedback signal can be expected to suppress motion at frequencies below 20 Hz, the support structure of the final quad must be rigid at frequencies above 20 Hz. At these frequencies earth motion is typically a few nanometers or less.

For linear colliders of higher energy and smaller emittance, one can say roughly that the final focus system length will scale linearly with energy, and that BPM sensitivity will scale with beam size in the final focus system. The minimum spot size achievable with a final focus system is limited by synchrotron radiation in the final quad. This limit, the Oide limit, is a simple function of the beam emittance. It appears that systems can always be designed to achieve the Oide limit.

In summary, it appears that final focus systems can be built to perform as required in the NLC regime; however, we still need some BPM precision improvements and a method for determining the beam spot size of a single beam at the IP to aid the tuning process.

4.4.2 Crossing Angle

Because of beam blowup at the IP, an exit path of larger diameter than the final quad aperture is required. Hence the beams must cross at an angle, which lies somewhere between 10 and 50 mr. Above 10 mr, a crab cavity must be employed to twist the bunch as it travels to the IP. Simulations of masking and tests of crab cavity performance in the FFTB facility will determine the optimal choice of this angle for the contemplated range of NLC energies.

4.5 BACKGROUNDS & MASKING[57 – 61]

Experience with the SLC has shown that the problems associated with detector operation at e^+e^- linear colliders are quite different from those encountered at storage rings. The pulse rate at linear colliders ($\sim 10^2$ Hz) is many orders of magnitude less than the repetition rate of a circular machine, but the beam is created anew on each pulse, so the extremities of the beam momentum and betatron phase space are continuously repopulated. These beam "tails" create backgrounds that disrupt data acquisition and later analysis.

The detector is vulnerable to backgrounds created in several ways. Synchrotron photons are emitted by beam particles as they are bent onto the final straight trajectory leading to the IP, and as they are focused to a waist by the final quadrupole lenses. The energies and numbers of photons that are emitted depend critically on the demagnification of the final focus optics, and on how well the beam tails are removed by the upstream collimation system. Even with good collimation and beam control, it is necessary to shield the detector from synchrotron photons with carefully designed masking.

Particles that are lost from the beam at points upstream of the detector will shower in the materials that make up the apertures of the beam line, and create debris that contributes to the background in the detector. A fraction of these secondary particles eventually strike apertures nearer to the detector and can create tertiary backgrounds. Even a single high-energy particle striking the aperture of the final quadrupole lens or synchrotron masking per pulse can render the machine useless.

Bethe-Heitler production of energetic muons also occurs in the showering process started when particles are lost from the beam. Once produced, these penetrating secondaries are difficult to stop.

It is necessary to properly collimate the beams at points well upstream of the final focus in order to bring these backgrounds under control. Collimation of small-emittance high-power beams is not easily achieved. Beam particles will scatter from the edges of the collimating surfaces, and will repopulate the tails of the beam phase space so that repetitive collimation must be done. The use of conventional collimating devices is problematical because wakefields that can disrupt the bunch profile are generated by the collimating surfaces as the beam passes. Furthermore, the bunch is sufficiently intense that it is capable of destroying the collimator if it is targeted entirely onto the material of the device.

A new kind of background will appear at the collision point of future e^+e^- colliders. These machines will focus the beams to densities that are high enough that the particles in one beam will interact sufficiently strongly with the collective field of the opposing beam to produce "beamstrahlung" photons and low-energy electron-positron pairs. It is necessary to include these particles in considering the masking of the detector.

4.5.1 Interaction Region Masking

A successful Interaction Region (IR) masking scheme necessarily results from a delicately balanced compromise between the following (partially contradictory) requirements:

1. Prevent synchrotron radiation photons, produced either in the last bending magnet, or in the final quadrupole lenses, from entering active regions of the particle detector. The most sensitive detector element is probably the tracking system. Here, relic synchrotron photons carry energies in the range 0.01-10 MeV, and a few percent of these will be absorbed in the gases in the wire chambers and produce free electrons. The masking must be designed to reduce the occupancy in the detector to acceptable levels, thereby limiting the tolerable penetrating photon flux to several thousand per crossing (depending on their energy spectrum).
2. Intercept beamstrahlung-induced e^+e^- pairs (or their shower debris) before they can penetrate active regions of the detector. This point is discussed in some detail in Section 4.5.2.
3. Maximize the beam stay-clear apertures close to the detector, so that the number of intercepted synchrotron radiation photons or high energy stray beam particles is kept to a bare minimum, as such hits become powerful re-radiating sources on their own.

Calculations combining charged particle tracking through the final quadrupole fields, and EGS simulations of photon rescattering and electromagnetic showering off masks are used to evaluate the amount of debris traversing various parts of the detector. Preliminary results indicate that as in the SLC case, the primary synchrotron flux, for a given magnetic configuration, is extremely sensitive to beam tail populations: upstream collimation at the 4-5 sigma level reduces the flux by about 4 orders of magnitude, and the mean critical energy by about a factor of 30. Beam centering on the quadrupole axis is similarly critical, at the level of 1 sigma (50-100 microns). Radiation from the last (soft) bending magnet hits primarily the final quadrupole bore and face; rescattering effects and tolerance levels remain to be evaluated.

A large fraction of the primary photons hit either the bore of the quadrupoles themselves, or (in the case of radiation produced in the penultimate quadrupole) the upstream face of the last lens. The necessity to prevent transmission of this intense flux, transversely to the beam direction, through the quadrupole bodies and into the detector, implies the magnets may have to be enclosed inside an absorbing photon shield, which in turns complicates the IR layout and may limit the minimum achievable IP-to-quadrupole distance.

However, and in contrast to present experience, the large outgoing beam aperture already imposed by the safe disposal of the disrupted beam reduces the contribution of rescattered synchrotron energy (dominated by Compton scattering and atomic fluorescence) from the face and bore of the downstream quadrupoles to a negligible level. Synchrotron radiation which hits onto, and is rescattered from, the beam pipe inside the detector can also be neglected provided the pipe can be kept larger than about 1 cm. The problem would of course become much more severe if physics requirements dictated the installation of a very small radius (a few mm) vertex detector.

A potentially devastating source of backgrounds lies in high energy beam particles hitting the vacuum pipe or the quadrupole bore within a few meters of the interaction point (IP), thereby creating an intense high multiplicity electromagnetic shower that saturates the readout and pattern recognition capability of the detector. Such stray electrons can arise, for instance, from imperfect upstream primary collimation, from secondary particles generated on the edges of collimators, or from beam-gas beamstrahlung in the straight section between the last bending dipole and the IP. EGS simulations as well as SLC experience indicate a rate as high as one hit per crossing to be unacceptable. The best remedies known to date consist in the combination of a relatively large (10-15 sigma) beam stay clear in the final quadrupoles, supplemented by a carefully-sized absorbing mask that precisely shadows the quadrupole bore.

In summary, successful synchrotron radiation masking schemes appear achievable for beam crossing angles ~10 mr that are necessary for the extraction of the disrupted beams. The sensitivity to tail population imposes stringent requirements on upstream collimation, which must efficiently suppress all phase space beyond about five sigma. Stay-clear apertures of 10 sigma are required to reduce hits of high energy stray beam particles to the required level of much less than one per crossing. Finally, as discussed below, synchrotron radiation masking will also shield the detector from most particles produced in the beam-beam interaction.

4.5.2 Beam-Beam Backgrounds

The beam-beam interaction (between bunches) will produce beamstrahlung photons in addition to radiative Bhabha events (arising from individual e^+e^- interactions). These photons can turn into e^+e^- and $\mu^+\mu^-$ pairs during collision. The generally low-energy particles will be strongly deflected by the same collective field of the opposing beam and could be a potential source for backgrounds.

When the collective field of the opposing beam is sufficiently intense, it will interact "coherently" with a photon and turns it into a pair. This process, although the dominant channel of pair production if the beamstrahlung parameter is larger than unity, can nevertheless be exponentially suppressed by properly adjusting the beam parameters such that the field intensity is below a certain level. This is indeed attainable for NLC at 0.5 to 1.5 TeV center-of-mass energy. When this is achieved, the incoherent Breit-Wheeler and Bethe-Heitler processes become important by comparison. In addition, pairs can also be produced through virtual photons both coherently and incoherently.

Theoretical calculations on beamstrahlung are now fairly mature[57, 58, 62]. There has also been general agreement on the yield of coherent pairs. Further investigation of the energy and angular distributions of the pair particles is in progress. Initial results show that nearly all secondary particles are contained in the very forward cone and carry very little energy. However, this qualitative assessment will have to be evaluated in much greater detail to yield a realistic design of the IP beam pipe, and, in particular, of the minimum tolerable vertex detector radius.

Complete modelling with EGS on the effects of these particles on detectors is presently in progress. From initial calculations it appears that masking designed to shield the detector from synchrotron radiation will also shield the detector from the vast bulk of the particles produced from the beam-beam interaction. Backscattering of particles into the detector is negligible, and the energy that is deposited in forward luminosity monitors is well below the signal level expected from Bhabha-scattered beam particles. These calculations will continue, and will be refined as more accurate models of the beam-beam interaction become available.

4.6 BEAM COLLIMATION AND MUON CONTROL

4.6.1 Beam Collimation[63]

Recent calculations on scraper design have shown that wakefields (fields induced by the bunch acting back on the bunch) can be made small by having very small taper angles (1-10 mr) from the beampipe radius to the scraper gap. A single bunch train of 10^{11} 250 GeV electrons, however, contains 4000 Joules of energy. Since the beam radius will be 50 microns or less, even one missteered train striking a standard collimator will melt a small core region and create a damaging thermal shock.

A collimation system, similar in character to the final focus system, with sextupoles or octupoles placed at π phase apart, with a scraper located midway between, may provide a solution to the collimation problem. The non-linear elements blow up the beam size, which is then collimated, and then returned to its original size at the second non-linear element. This has the advantage that if the beam is indeed off axis in the first sextupole, it is sufficiently blown up that it will not damage the collimator. The disadvantages of this system are that it must be very long (all beam phases must be collimated) and has alignment tolerances similar to final focus systems.

Laser (Compton) collimation has also been proposed. An intense photon beam is focused at the edge of the beam. The particles there interact with the photons and are sufficiently degraded in energy that they can be scraped from the beam downstream.

More research work is required on collimation systems.

4.6.2 Muon Control[61]

Muons are created profusely at collimator locations, and a series of 10 m long magnetized iron toroids must be installed around the beampipe to deflect these muons and prevent them from reaching the detector. Detailed design of these systems has just begun. Initial simulations show that only three muons with energies greater than 50 GeV reach the detector from 1,000 produced at the collimator. Toroidal arrangements aimed at controlling and diverting higher momentum muons are expected to improve this attenuation to the required value of 1 in 10,000. The small bend angles within the final focus designs are sufficient to insure that even the highest energy muons enter a toroid at some point.

4.7 RELIABILITY, TOLERANCES & FEEDBACK

4.7.1 Reliability

The linear collider availability for physics and accelerator studies is required to be at least 90%. Every system with its constraints must meet this combined goal. In the case of the accelerator RF systems, a small fraction of the RF stations will be held in standby mode, ready to be switched in to replace any other nearby system that fails. This mode of operation causes minimum perturbation to the beam trajectory and energy, and results in only a short downtime. The number of magnetic and RF components is of the order of 5000, to be compared to SLC (2100), the Tevatron (2200), and the SSC (18000). Thus, the components in the NLC must be two to five times more reliable than those in existing machines. Unlike pp and p̄p machines, there is no accumulation or refilling time in the case of a lost "fill." However, the circular colliders can coast through an injector problem. There does not appear to be any single system in the NLC that is not expected to meet the reliability requirements.

4.7.2 Alignment and Tolerances

Alignment tolerances are of the order of 30 to 100 μm for most components. This appears to be achievable using presently available survey and alignment techniques[64]. The beam must be aligned within a few tens of microns of the center of the RF structures to prevent head-tail effects. Magnetic elements must be aligned to a few tens of microns to prevent emittance dilution due to the chromatic phase advance (betatron oscillations combined with energy spread)[44]. Because the beam itself can be used as an alignment reference, the mechanical alignment tolerances can be relaxed.

Vibration and ground motion are potentially very problematic, but because their frequencies are normally below about 10 Hz, operating with repetition rates over 100 Hz allows making pulse-to-pulse closed-loop feedback correction with averaging times short compared to the period of the motion. Earth tides, water table variations, diurnal effects, power supply drifts, all can be corrected using beam feedback.

In general, the pulse-to-pulse beam centroid position must be stable to the order of the rms bunch size in order not to adversely affect the luminosity. Because beam sizes are the order of $\sigma_x \approx 10-20$ μm and $\sigma_y \approx 1-2$ μm in the linac, pulse-to-pulse jitter exceeding these values can lead to loss of luminosity. It is therefore desirable to have beam position monitors (BPMs) with resolution and accuracy in the range of 1 μm and 10 μm respectively. For a BPM with a half-aperture of 5 mm, this corresponds to ±0.02% and ±0.2% of the half-aperture, about a factor of five better than has been obtained with 11 mm half-aperture stripline BPMs at SLC[65].

Mechanical tolerances do not scale with aperture, however, so new pickup geometries with more easily controlled mechanical tolerances are desirable. One possibility is slot-coupling (four slots) through the beam pipe wall to four separate transmission lines outside the beam pipe[66]. Because the coupling strength is determined by small holes in the beam pipe wall, the mechanical tolerances are considerably improved.

In order to achieve the 0.2% accuracy, it is necessary to measure the relative BPM signal amplitudes to about ±0.6%. Because of cable attenuation disparities and electronics offsets, it will be necessary to excite the BPM itself with a calibration signal, probably by sequentially exciting each slot and recording the output response from the others. This may attain the needed accuracy. The fundamental limit to resolution is set only by the signal-to-noise ratio. A 1000:1 (60 dB) signal-to-noise ratio should yield resolution of about ±0.02%.

A fundamentally different approach for a BPM design is for the beam to excite higher order TM modes in small rectangular resonant cavities[67]. Two disadvantages of this are the possibility of excessive wakefields, and the fact that the signal amplitude is proportional to both the beam displacement and the beam current. In the latter case, the normalization would have to be done in the electronics using a known beam intensity signal.

4.7.3 Feedback and Control

Whereas storage rings involve magnets and

RF cavities with some inertia due to their inductance and Q, and feedback-stabilized power sources, linear colliders generate their accelerating fields each pulse and are subject to pulsed kicker magnet fluctuations. Because of this, the type of beam diagnostic and control equipment needs to be somewhat different. For example, a frequency spectrum analyzer which is very useful in storage rings is of limited use at a linear collider. On the other hand, using feedback on the beam is an important tool to achieve stable collisions.

A linear collider is (or should be viewed as) one big feedback loop, with many sensor inputs and control variables. The sensor inputs come from the beam pickups (intensity, position, angle, size, tails, etc.). Control variables are steering magnets, RF drive and phase, source intensity, and possibly quadrupole and other magnet settings. A linear system would have a matrix which gives the change in each control setting for a set of error signals from the sensors. The dimensionality of this matrix would be large since the beam parameters would need to be stabilized at several locations. While it would be desirable to have several small loops, one at each location which run independently, this is difficult since up-beam controls will affect most down-beam pickup sensors. Even with separate beams from separate linacs, the feedback loops for each would get coupled together at the final focus interaction point in order to keep the beams in collision. One traditional method for decoupling loops which otherwise are coupled is to let each have its own frequency space. For example a slow, low frequency loop would work outside a fast, AC coupled loop. This might be useful if the machine repetition rate is high enough, but for most designs, there is little frequency space to split. To achieve the best performance, sampling, averaging, and correcting at the machine repetition rate would be desirable. The feedback control system should be built to allow an efficient (easy to set up and modify) global fast beam feedback system. The SLC machine is an excellent laboratory to develop such a feedback system.

RF system amplitude and phase must be controlled to a fraction of a percent and a few psec respectively. Damping ring systems, including the injection and extraction kicker amplitude and timing, also require tight tolerances. Based on the successes at SLC, there do not appear to be any technological barriers.

Feedback loops using the beam position monitors will keep the beam on axis via x and y correctors at regular intervals. Beam energy (pulse-to-pulse) stability of about 0.01% will be accomplished using BPMs placed in high dispersion areas between the linac and the final focus. For example, the dispersion after a 2 mrad bend is probably of the order of 25 cm, thus requiring several BPMs with resolution of about 10 μm to measure the rigidity.

5. Example Designs and Future Potential

Exploring the ultimate potential of linear colliders is more difficult than for the SSC, FNAL, or even a B-factory, because there is not an initial starting point. In addition, one can explore both energy and luminosity upgrades. To simplify the situation we will consider only designs that have a luminosity such as to give the same event rates, independent of energy (i.e. we will assume that the luminosity must rise as the square of the energy, and we will start from the values suggested for "full exploitation" in the Luminosity Goals section). For instance, we can consider:

	C of M Energy (TeV)	**Luminosity cgs**
NLC1	0.5	2×10^{33}
NLC2	1.0	1×10^{34}
NLC3	1.5	2×10^{34}
NGC1	3.0	1×10^{35}
NGC2	6.0	4×10^{35}

It should not be surprising to find that the technical difficulty of achieving the very high luminosities is more severe than that of simply increasing the energy. What is surprising is that self-consistent designs for such machines are, even now, possible. Many innovations are needed and their practicality is not yet known, but it is encouraging.

We will briefly discuss how one might achieve each of the above requirements.

Next Linear Collider, Phase 1 (NLC1)

As an example of a possible approach to a 0.5 TeV collider, we give a parameter set of the type being considered now at SLAC and KEK. The idea is to design an upgradable machine, which would initially use a relatively low gradient (50 MeV/m), a modest repetition rate (120 Hz), X-band (11.4 GHz), and low initial wall power consumption (40 MW, assuming a 20% RF power source efficiency).

Table I
Example Designs for Linear Colliders

		NLC1	NLC3	NGC1	ZFAC
Energy		0.5 TeV	1.5 TeV	3.0 TeV	0.1 TeV
Luminosity ($cm^{-2}s^{-1}$)		2×10^{33}	2×10^{34}	1×10^{35}	1×10^{34}
Linac Length		13 km	13 km	32 km	2.2 km
Accel. Gradient		50 MeV/m	50 MeV/m	100 MeV/m	50 MeV/m
RF Frequency		11.4 GHz	11.4 GHz	30 GHz	11.4 GHz
# Particles/Bunch:					
	Linac	1.2×10^{10}	1.3×10^{10}	0.14×10^{10}	0.6×10^{10}
# Bunches, n_b		10	28	150	120
Repetition Freq.		120 Hz	120 Hz	330 Hz	360 Hz
Wall-Plug Power		40 MW	170 MW	180 MW	50 MW
RF Source Eff.		20%	40%	40%	20%
IP Beam Size:	σ_y	3.6 nm	1.2 nm	0.5 nm	7.0 nm
	σ_x	160 nm	130 nm	16 nm	180 nm
	σ_z	100 μm	50 μm	40 μm	140 μm

To achieve the required luminosity of 2×10^{33} $cm^{-2}s^{-1}$, this example employs multiple (10), relatively small bunches (1.2×10^{10} particles).

Next Linear Collider, Phase 2 (NLC2)

The NLC1 design could be upgraded to 1 TeV by simply increasing the RF power (by a factor of four) and thus the gradient by a factor of two. The luminosity rises by the required factor of four almost automatically, but the wall power rises by the same factor to 140 MW.

Next Linear Collider, Phase 3 (NLC3)

A further upgrade along the same lines would increase the wall power requirement to 300 MW. It is not unreasonable to assume, however, that a more efficient power supply will become available (e.g. 40%). One could then achieve the required luminosity at 1.5 TeV and 150 MeV/m accelerating gradient. This assumes, of course, that such a gradient does not generate excessive dark current or breakdown, which is yet to be determined. If such high gradients are not attainable, it would be possible to increase the length by 50% to obtain the increase in energy.

It may also become necessary, in order to control the beamstrahl- ung, to increase the number of bunches, and decrease the fraction of energy each removes from the structure by each bunch ($\eta = 0.6\%$). This requires a damping ring with rather high frequency RF and very short bunches. Such a ring, although its performance has been predicted by extrapolation, remains to be designed.

Next Generation Collider, Phase 1 (NGC1)

In order to move to a higher energy and luminosity, we need a more radical approach. In order to get more luminosity without increasing the wall power, we need more efficient coupling of power to the beam. In the above designs, using short trains of bunches, the RF to beam efficiency would be only about 15%. Far higher efficiencies are possible in continuous wave (CW) operation of a linac. One could thus consider such operation with the structure operated for a few fill times at each pulse.

If such long pulse operation were used at X-band at an energy of 3 TeV, then the stored energy would be very high, and the repetition rate would become very low, causing problems in the detection of vibration for feedback. The conclusion is that one must go to a higher RF frequency, for instance, 30 GHz (as studied at CERN). A reasonable initial accelerating gradient would be 100 MeV/m (again as proposed by CERN).

The example given in the table achieves the required luminosity with reasonable beamstrahlung and other parameters. As in the later phase NLC, it would employ a very large number of small bunches. The damping ring must have even higher frequency RF, and will clearly require some use of combined function magnets to increase the horizontal partition function and lower the longitudinal.

Next Generation Collider, Phase 2 (NGC2)

Upgrading the energy of the above NGC to 6 TeV might be possible by increasing the gradient to 200 MeV/m, or it could be done by increasing the length. In either case, achieving the required luminosity of 4×10^{35} cm^{-2}s^{-1} would require another significant innovation. There seem to be two possibilities.

One approach would be to significantly improve the damping ring performance. The use of Robinson Wigglers could significantly increase the transverse damping, and radically decrease the longitudinal. This, in combination with a further decrease in the horizontal-vertical coupling, would seem to achieve the required performance. Clearly, no such ring has been designed.

The other idea is to use "super-disruption". In this case, a larger weak bunch is accelerated a few hundred microns ahead of each main bunch. These precursor bunches contribute little to the luminosity, but after passing through one another they act to strongly focus the main bunches. As the main bunch is focused by the precursor, it will synchrotron radiate, and thus focus less effectively (this effect gives rise to the Oide limit). But at these energies, and with a short precursor, the radiation is in the quantum regime, and thus suppressed. The details of such a scheme are far from understood, but it would seem to provide the needed luminosity improvement.

To conclude this section, the energy range from 0.5 to 1.5 TeV seems attainable with consistent luminosity and an RF frequency of about 11 GHz. The higher-energy accelerators rely on a higher frequency (30 GHz) and also require more bunches to achieve the necessary luminosity.

A Z-Factory Postscript (ZFAC)

It must also be noted that the application of some of these ideas to lower energy colliders has not yet been studied in much detail. A high luminosity linear Z-factory should be looked at again. For example, the use of the long pulse in an X-band collider, combined with some damping ring innovation, would seem to give a Z-factory with a luminosity of 10^{34}, which would have considerable physics interest.

6. Outlook

6.1 TECHNOLOGY LIMITATIONS AND RESEARCH PROGRAMS

To review the technology limitations in linear colliders, it is convenient to follow the beam through the subsystems which make up the accelerator complex. The generation of suitable beams of electrons is relatively straightforward, but the production of intense positron beams encounters technological problems. Positrons are produced by pair production on a target and here we have problems in both the peak and average power that the target can handle. There are several solutions being studied including multiple targets, liquid metal targets, accumulator storage rings and systems which use very high energy incident beams of photons (> 100 GeV) impinging on thin targets. This is not seen as a hard technology limit, but rather one of cost and reliability optimization.

The beams are then damped to small emittances in damping rings which are well understood. In fact, the damping rings are not unlike the storage rings being built today as low emittance synchrotron light sources. The limits on performance of these types of rings is in the physics rather than technology. Again, the issue is cost or complexity, as multiple stacked rings can overcome almost any difficulties including assisting the above positron production problem.

After initial acceleration and further bunch length compression, the particles enter the main linear accelerator. Here we must confront several technological problems. We want a high accelerating gradient, not only to get high energy, but to optimize luminosity. The maximum gradient that one can achieve is limited by dark current, i.e. field emission inside the accelerating structure which loads down the accelerator and could produce intolerable backgrounds. There are several programs around the world which are studying this problem and trying to define a practical limit.

It is based on this continuing research that present designs do not exceed gradients of 100 MeV per meter.

To achieve high gradients and luminosity we need high peak power RF generators at high frequency. There are many R&D programs worldwide exploring various technological approaches to this issue. Much of this work is collaborative as there are many ideas to be explored. They range from conventional klystrons pushed to higher power, through new technology generators, to two beam accelerators. This has been discussed in Chapter 3. These research programs are making rapid progress and again the selection of the best approach will depend on cost, efficiency and reliability. Because the cost and electrical efficiency of RF generators is such a fundamental design parameter for linear colliders, we see these research programs continuing beyond the next linear collider.

The control of the emittance of high current multi-bunch trains of particles during acceleration introduces several technological problems. They include accelerating structures which damp wakefields, high precision alignment techniques, high precision beam instrumentation, and beam control systems. Much of the research in the accelerator physics of linear colliders is directed towards defining and minimizing tolerances on the parameters which impact the above. We now feel that these problems are difficult, but not impossible with today's technology. Furthermore, the continuing experience being gained from the SLC development program is invaluable in pushing technology forward on a variety of fronts.

Experience with the SLC has shown that during acceleration, the beams can develop significant tails in both energy and transverse phase space. To control backgrounds, these tails must be removed by some collimation system in the linac and the early part of the final focus. With the high power and low emittance beams of TeV linear colliders, this presents technological problems. In a fault condition, any conventional collimator can be damaged by a single errant beam pulse. This is also true of many other accelerator components. Machine protection system design will be complex and new approaches to collimation will be required. Many ideas, including nonlinear magnetic fields or laser beams, are actively under study.

The final focus systems which reduce the beam size by typically a factor of several hundred are well understood, but a few of the final components are limited by alignment and vibration control technology. The international collaboration building the Final Focus Test Beam at SLAC is a research program which addresses these technology questions and will be used to test solutions to these problems.

6.2 ACCELERATOR PHYSICS QUESTIONS/RESEARCH

A key feature in many designs for Next Linear Colliders is the acceleration of many bunches on each fill of the accelerator. This bunch train impacts the design of every subsystem from the source to the final focus. Studies have shown that the train can be kept stable in all these subsystems; however, this requires special RF structures which damp the wakefield strongly in the main linac.

Although the emittances seem to be straightforward to generate, the small beams induce special requirements on the measurement systems to achieve good optical matching. Much progress has been made on the study of emittance preservation in the linac. There are now designs emerging in which the emittance can be preserved with conventional alignment tolerances for all components in the linac, provided that BPM resolution is in the micron range. Future studies must ensure that this can be accomplished in the face of all errors present in a real linac and in the face of the expected motion of all components (both slow drifts and fast jitter).

The RF power source for the Next Linear Collider is itself a complicated dynamical system. For two-beam accelerators, the stability of the drive beam needs careful study. For more conventional power sources, such as klystrons, theoretical studies show that designs should produce about 100 MW with about 50% efficiency. Other designs which use gridded cathodes and/or permanent magnet focusing, however, may have lower costs or higher efficiency and must be investigated.

The final focus for the Next Linear Collider is similar in spirit, but different in detail from that of the SLC. Optical designs are now being studied with regard to tunability and stability. This effort is coupled closely to the Final Focus Test Beam, where these ideas can be experimentally tested.

The SLC has shown that in order for a linear collider to operate with high reliability, it must have an appropriate feedback network. This is a key feature of linear colliders; they must be stabilized during tuning and colliding beam operation. The feedback necessary for each system in the collider needs to be specified together with the interaction or coupling of feedback between systems.

Many of the effects just discussed have been under extensive theoretical study. Indeed, during the past few years, a general theoretical foundation has been laid for the Next Linear Collider. This has been supported by actual experience with the SLC and extensive technical progress on RF power sources and structures.

We believe it is now possible using this foundation to specify in fair detail the design of a Next Linear Collider. This, however, must be supported by ongoing experimental tests at the SLC to verify key features of the design. Provided that technical efforts are successful for an RF power source, and provided that there is sufficient support, a conceptual design of a Next Linear Collider could be completed in the mid-1990s.

References

[1] T. Barklow, these proceedings and SLAC–PUB–5364.

[2] *Proceedings of the Workshop on Physics of Linear Colliders*, Eds. L. Palumbo, S. Tazzari and V.G. Vaccaro, Capri, Italy (1988), available from INFN Frascati, Italy.

[3] *Proceedings of the Summer Study on High Energy Physics in the 1990's*, Snowmass CO, July 1988 (World Scientific, Singapore, 1989).

[4] *Proceedings of the International Workshop on Next-Generation Linear Colliders*, SLAC, Stanford CA, Dec. 1988, SLAC-Report-335.

[5] *Proceedings of the 2nd International Workshop on Next-Generation Linear Colliders*, Eds. S. Kurokawa, H. Nakayama and M. Yoshioka, KEK, Tsukuba, Japan, March 1990, KEK Internal 90–22.

[6] R. Erickson Ed., "SLC Design Handbook," SLAC Document, Dec. 1984.

[7] M. Breidenbach et al., "SLC Performance in 1991," SLAC Document, Dec. 1989.

[8] V. Balakin, "Autophasing," in Ref. [1].

[9] J. Seeman and L. Merminga, "Mutual Compensation of Wakefields and Chromatic Effects of Intense Bunches," 1990 Linac Conference, Albuquerque NM, and in SLAC–PUB–5220.

[10] T. Raubenheimer and R.D. Ruth, "A Dispersion Free Trajectory Correction Technique for Linear Colliders," SLAC–PUB–5222 (1990), to be published.

[11] J. Seeman, "New Compensation of Transverse Wakefield Effects in a Linac by Displacing Accelerating Structures," 1990 Linac Conference, Albuquerque NM, and in SLAC-PUB-5337.

[12] C. Adolphsen et al., "Beam-Based Alighment Technique for the SLC Linac," IEEE Particle Accelerator Conference, Chicago, 1989, and in SLAC-PUB-4902.

[13] G. Fischer, "Alignment and Vibration Issues in TeV Linear Colliders," Int. Conf. on High Energy Accelerators, Tsukuba, 1989, and in SLAC-PUB-5024.

[14] J. Clendenin, "The SLC Polarized Source," Workshop on Polarized Electron Sources and Spin Polarimeters, Bonn, West Germany, September 1990.

[15] G.A. Loew and J.W. Wang, "RF Breakdown Studies in Room Temperature Electron Linac Structures," *13th Int. Symp. on Discharges and Electrical Insulation in Vacuum*, Paris, France, June 1988, and in SLAC–PUB–4647 (May 1988).

[16] G.A. Loew and J.W. Wang, "Field Emission and RF Breakdown in Copper Linac Structures," in *Proceedings of the 14th Int. Conf. on High Energy Accelerators, Part IV, Part. Accel.* 30, 225 (1990), and in SLAC–PUB–5059 (August 1989).

[17] G. Caryotakis, "Multimegawatt RF Power Sources for Linear Colliders," Presented at the 1991 Particle Accelerator Conference, San Francisco CA, May 1991, and in SLAC–PUB–5508.

[18] A. Vlieks, "Numerical Simulation of the SLAC X–100 Klystron Using RKTW2D," Presented at the 1991 Particle Accelerator Conference, San Francisco CA, May 1991, and in SLAC–PUB–5497.

[19] J. Haimson and B. Mecklenburg, "Use of TW Output Structures for High Peak Power RF Generation," *Proc. of the 1990 Linear Accelerator Conference*, Los Alamos National Laboratory Report LA–12004C (1991), pp. 244–246.

[20] V. Granatstein, P. Vitello, K.R. Chu, K. Ko, P.E. Latham, W. Lawson, C.D. Striffler and A. Drobot, "Design of Gyrotron Amplifiers for Driving 1 TeV e^+e^- Linear Colliders," *IEEE Trans. Nucl. Sci.* **NS-32**, 2957 (1985).

[21] A. Mondelli, D. Chernin, V. Granatstein, P. Latham, W. Lawson and M. Reiser, "RF Frequency Scaling and Gyroklystron Sources for Linear Supercolliders," *Lecture Notes in Physics*, M. Month and J. Turner eds. (Springer-Verlag, 1988), pp. 533–606.

[22] W. Lawson, J. Calame, B. Hogan, W. Main, P.E. Latham, C.D. Striffler and V.L. Granatstein, "High Power Microwave Generation from a Two-Cavity Gyroklystron Experiment," Conference Digest of the 15th Intl. Conf. on IR and mm-Waves, ed. R.J. Temkin, *Proc. of the SPIE* **1514** (1990), p. 669.

[23] W. Lawson, et al., "Efficient Operation of a High Power X-Band Gyroklystron," to be published in *Phys. Rev. Lett.*

[24] W. Lawson, et al., "Amplification Studies of a Multi-Megawatt, Two-Cavity X-Band Gyroklystron," *Proc. of the IEEE Particle Accelerator Conference*, San Francisco CA, May 1991, to be published.

[25] Z.D. Farkas, H.A. Hogg, G.A. Loew and P.B. Wilson, "SLED: A Method of Doubling SLAC's Energy," *Proceedings of the 9th Int. Conf. on High Energy Accelerators*, SLAC, Stanford CA, May 1974, pp. 576–583, and in SLAC–PUB–1453.

[26] Z.D. Farkas, "Binary Peak Power Multiplier and its Application to Linear Accelerator Design," *IEEE Trans.*, **MTT 34**, 1036 (1986), and in SLAC–PUB–3694.

[27] T.L. Lavine, et al., *1991 Particle Accelerator Conference*, San Francisco CA, May 1991, and in SLAC–PUB–5451.

[28] P.B. Wilson, Z.D. Farkas and R.D. Ruth, "SLED–II: A New Method of RF Pulse Compression," *Proceedings of the 1990 Linear Accelerator Conference*, Albuquerque NM, September 1990, Los Alamos National Laboratory Report LA–12004–C, pp. 204–206, and in SLAC–PUB–5330.

[29] V.E. Balakin and I.V. Syrachev, "A New Approach in RF Power Multiplication," VLEPP–Note–06/1990, Branch of Novosibirsk Inst. Nucl. Phys., Protvino, USSR (March 1990).

[30] H. Deruyter, et al., "Damped and Detuned Accelerator Structures," *Proceedings of the 1990 Linear Accelerator Conference*, Los Alamos NM, September 1990 (Los Alamos National Laboratory Report LA–12004–C) pp. 132–134, and in SLAC–PUB–5322 (1990).

[31] K.A. Thompson and J.W. Wang, "Simulation of Accelerating Structures with Large Stagger-Tuning," Presented at the 1991 Particle Accelerator Conference, San Francisco CA, May 1991, and in SLAC–PUB–5465.

[32] D.B. Hopkins, A.M. Sessler, and J.S. Wurtele, "The Two-Beam Accelerator," *Nucl. Inst. & Meth. for Phys. Res.*, **228**, pp. 15–19, (1984).

[33] G. Guignard, "Status of CERN Linear Collider Studies," *Proc. of the Linear Accelerator Conference*, Albuquerque NM, Sept. 1990.

[34] S.S. Yu and A.M. Sessler, "The Relativistic Klystron Two-Beam Accelerator," *Phys. Rev. Lett.*, (1987).

[35] M. Allen et al., *Proc. of the European Particle Accelerator Conference*, Rome (1988), and Nice (1990).

[36] D.B. Hopkins, et al., "FEL Power Sources for TeV Linear Colliders," *Proc. of Int. Linac Conference*, Williamsburg, VA, (1988).

[37] W.A. Barletta, "Electron Accelerator with Pulsed Power Drives," *Proc. of the 2nd European Particle Accelerator Conference*, Nice, France (1990).

[38] W.A. Barletta, "Cost Optimization of Induction Linac Driven Power Sources for Linear Colliders," *Proc. of Int. Workshop on New Methods of Acceleration*, Orsay, France (1986).

[39] W.E. Nexen, et al., "The ETA-II Induction Linac as a High Power FEL Driver," *Proc. of the 11th Int. Conference on Free Electron Lasers* (1989).

[40] D.L. Rubin, et al., "Superconducting RF Linear Colliders," *Proc. IEEE Particle Accelerator Conference*, Chicago IL (1989), p. 721.

[41] D.L. Moffet et al., CLNS 89/923, SRF 890531/3.

[42] T.O Raubenheimer, W.E. Gabella, P.L. Morton, M.J. Lee, L.Z. Rivkin and R.D. Ruth, "A Damping Ring Design for Future Linear Colliders," *Proc. of the IEEE Particle Accelerator Conference*, Chicago IL, March 1989, and in SLAC–PUB–4912.

[43] M. Bassetti, S. Guiducci and L. Palumbo, "General Considerations on a Damping Ring for a 5–10 GeV Linear Collider," *Proc. 1988 European Particle Accelerator Conference*, Rome, Italy, p. 508.

[44] R.D. Ruth, "Beam Dynamics in Linear Colliders," *Proc. of the Linear Accelerator Conference*, Albuquerque NM, Sept. 1990, and in SLAC–PUB–5360.

[45] V. Balakin, A. Novokhatsky and V. Smirnov, *Proc. of the 12th Int. Conf. on High Energy Accelerators*, Fermilab, 1983, p. 119.

[46] J. Tuckmantel, "Beam Tracking With RF-Focussing in CLIC," CLIC-Note 87 (1989).

[47] T.O. Raubenheimer, "A New Technique of Correcting Emittance Dilutions in Linear Colliders," To be published in *Nucl. Instr. Meth.*, and in SLAC–PUB–5355 (1990).

[48] R.D. Ruth, "Multibunch Energy Compensation," SLAC–PUB–4541 (1989), and in Ref. [1].

[49] K.A. Thompson and R.D. Ruth, "Controlling Transverse Multibunch Instabilities in Linacs of High-Energy Linear Colliders," *Phys. Rev. D* **41**, 964 (1990), and in SLAC–PUB–4801.

[50] K.A. Thompson and R.D. Ruth, SLAC–PUB–4962, to be published.

[51] H. Deruyter, et al., "Damped Accelerator Structures," *Proc. of the 2nd European Particle Accelerator Conference*, Nice, France, June 1990, and in SLAC–PUB–5263.

[52] K. Oide, "Synchrotron Radiation Limit on the Focusing of Electron Beams," *Phys. Ref. Lett.* **61**, 1713 (1988).

[53] J. Irwin, "The Applications of Lie Algebra Techniques to Beam Transport Design," *Nucl. Inst. & Meth.* **A298**, 468 (1990), and in SLAC-PUB-5315.

[54] J. Buon, "Final Focus Test Beam for the Next Linear Collider," *Proc. of the 2nd European Particle Accelerator Conference*, Nice, France, June 1990.

[55] O. Napoly and B. Zotter, "Progress on the Final Focus for CLIC," CERN/SL/90–63 (AP) (1990).

[56] R. Brinkmann, "Optimization of a Final Focus System for Large Momentum Bandwidth," DESY–M–90–14 (1990).

[57] P. Chen, "Disruption, Beamstrahlung, and Beamstrahlung Pair Creation," SLAC–PUB–4822 (1988), and in Ref. 2.

[58] P. Chen and V.I. Telnov, "Coherent Pair Creation in Linear Colliders," *Phys. Rev. Lett. 63*, 1796 (1989), and in SLAC–PUB–4923.

[59] R. Jacobsen, et al., "Detector Background Conditions at Linear Colliders," Invited Talk at the 5th International Conference on Instrumentation for Colliding Beam Physics, Novosibirsk, USSR, March 1990, and in SLAC–PUB–5205.

[60] T. Tauchi, et al., "Background Problem at Interaction Point for an e^+e^- Linear Collider," these proceedings.

[61] L.P. Keller, "Calculation of Muon Background in a 0.5 TeV Linear Collider," these proceedings.

[62] R. Blankenbecler, S.D. Drell and N. Kroll, "Pair Production Photon-Pulse Collisions," *Phys. Rev. D* **40**, 2462 (1989).

[63] L. Merminga, et al., "Collimation Systems in the Next Linear Collider," these proceedings, and in SLAC–PUB–5436.

[64] G.E. Fischer, *Particle Accelerators* **31**, p. 47 (1990).

[65] J.L. Pellegrin, M. Ross, et. al., *Proc. IEEE Particle Accelerator Conference*, p. 673 (1987).

[66] L. Faltin, *Nucl. Inst. and Meth.* **148**, p. 449 (1978).

[67] Kroes, Maaskant, Sluyk and Spelt, *IEEE Trans. Nucl.. Sci.* **NS-28**, p. 2362 (1981). See also Chan, Bird, Coulas, McKeown, and Shin, *IEEE Trans. Nucl. Sci.* **NS-28**, p. 2328 (1981).