SLAC-PUB-3357 June 1984 (M)_

THE FUTURE OF HIGH ENERGY PHYSICS*

W. K. H. PANOFSKY Stanford Linear Accelerator Center Stanford University, Stanford, California 94305

I am taking the liberty of interpreting the title "Future of High Energy Physics" as "The Future of High Energy Colliders." This is, of course, grossly incorrect for a number of reasons. The first is that there has been recently an increase in the number of important non-collider experiments relevant to elementary particle physics such as proton decay and neutrino experiments carried out underground, and the DUMAND underwater experiment. Further, a discussion of the future of high energy physics should rightfully be preceded by a recital of the large variety of theoretical concepts bearing on the expected and speculative phenomena in the next region of high energy. Then detector and data handling technology continues to be challenged by the advances in collider energy and luminosity. My only excuse for restricting this discussion to accelerators and colliders is that all subjects bearing on the future of high energy physics cannot possibly be covered within one hour.

A further fact is that, notwithstanding the necessary guidance which the theoretical predictions give to the definition of future accelerator and collider parameters, the predicted thresholds for new phenomena are rarely sharply defined. Also historically, the detector designers have ultimately produced devices able to match collider performance, albeit at high and often unanticipated cost. Moreover, I would like to remind you that in the past accelerators and colliders

^{*} Work supported by the Department of Energy, contract DE-AC03-76SF00515. (Lecture presented at the University of Hawaii, Honolulu, March 14, 1984.)

have rarely been built for the right reason! By this I mean that while new accelerator have generally been able to support those experiments which had been the basis of the theoretical justification for those accelerators, the principal impact on science these accelerators have made has been in unexpected fields. Thus one can say with justification that in the past the progress of high energy physics has been largely paced by accelerator, collider and detector technology.

The conventional way in which the growth of accelerator performance has been characterized in the past has been through the chart (Figure 1) originated by Livingston. This chart demonstrates that the spectacular exponential growth of laboratory or equivalent laboratory energy in the past has been made possible through a succession of technologies; as any one technology reached its natural limit it was superceded by new developments. Figure 2 shows a similar pattern for electron-positron colliders. Here thus far only a single technology has supported the exponential growth in this field—colliding beam storage rings.

The interest in high energy physics is worldwide, and Figure 3 shows the performance of the world's high energy accelerators and colliders, albeit on a somewhat simplified basis. We all know that collision energy is not the only relevant parameter; therefore we also plot here the luminosity which is the number by which the cross-section of the process of interest is to be multiplied to yield the data rate.

Figure 3 contains both electron and proton machines, both in the form of colliders and stationary target machines. Again, this gives to some extent a misleading impression. Historically electron and proton machines have had roles in high energy physics that overlap only partially. Proton machines have generally led the energy frontier; this is a direct consequence of the fact that in the past attainment of a given collision energy had been less expensive for protons than for electrons. On the other hand, the results of electron scattering experiments, and in particular electron-positron annihilations, tend to be easier to interpret. In the former we are exploring unknown structures with known interactions, since the electron does not carry the strong force. In the latter we are creating final states of particular simplicity, since one-photon annihilation of electrons and positrons produces a state of well-defined quantum numbers.

Since we now understand that the proton is composite, while down to the smallest dimensions investigated thus far (less that 10^{-16} cm) electrons are still point-like, there is another energy which is highly relevant: this is the collision energy measured not in the center-of-mass frame of the particles colliding in the laboratory but in the center-of-mass frame of those constituents thus far considered to be elementary, in other words, the center-of-mass frame of leptons, quarks and gluons. Figure 4 shows the energy equivalence between electron and proton colliders in these terms. Naturally, there is not an exact conversion factor between the experimental reach in terms of collision energy from protons to electrons; such a correspondence depends somewhat on the process under investigation, on theoretical assumptions about the momentum distribution of quarks and gluons in the proton, and on the available luminosity. The lines in Figure 4 derive from calculations by John Ellis on these subjects for a number of assumed "yard-stick" reactions in the TeV region. If we take this correspondence literally, then Figure 3 can be replotted in terms of collision energy in the elementary constituent frame. This plot (Figure 5) shows that proton machines still lead in terms of reach into new energies, but this lead is no longer anywhere near as large.

The worldwide distribution of high energy accelerators and the growth in energy would, of course, have been impossible had it not been for a drastic decrease in the unit cost, that is cost per unit of energy in the construction of accelerators and colliders. Figure 6 gives a rough representation of the unit cost decrease in time. It is extremely difficult to generate systematic data on this subject because construction of each accelerator or collider facility has to meet special conditions. Moreover, many of the recent advances have been made by adding a new installation to an older machine which is used as the injector; in that case Figure 6 gives only the incremental cost. Also, since we are talking about accelerators in many countries, there are questions of the true equivalence of currencies, accounting practices, etc. However, ignoring all this, we see that the unit cost of accelerator and collider construction has decreased by about 5 orders of magnitude during which time the center-of-mass energy has increased by 6 orders of magnitude. The result is, of course, as is painfully clear to anyone, that the actual cost of each new accelerator and collider construction has grown as the frontier has advanced, the decrease in cost per MeV notwithstanding. It is for this reason that there is a difficult struggle in each continent to maintain the progress of high energy physics within the available budget.

The price which has been paid for this dramatic advance in energy has been high. There has been a decrease in the number of high energy accelerator and collider installations operating at the frontiers of the field, and this has caused the social pattern in utilizing these machines to undergo drastic changes.

Figure 7 shows the life and death cycle of American high energy physics accelerators since World War II. The number of active machines has indeed shrunk

4

drastically and the future of those machines and laboratories operating them is the central focus of current planning and debate.

As you know, in the United States the focus of future planning for high energy facilities has been the High Energy Physics Advisory Panel which advises the Department of Energy. This Panel, in turn, has at frequent intervals appointed subpanels to deal specifically with this issue. The most recent such panel was that chaired by Stan Wojcicki which reported on July 1983 after an extensive series of internal and external hearings and debates.

The reasons the decisions as to future facilities faced by this subpanel were so difficult are the following:

- (a) Two construction projects breaking new boundaries and offering substantial promise are now in progress. These are: (1) Tevatron I at Fermilab, hopefully leading to proton-antiproton collisions of 1 TeV per beam, and (2) the SLC at SLAC which serves the dual purpose of giving the United States an inexpensive entry into the 100 GeV center-of-mass region for electron-positron collisions where neutral intermediate vector bosons are expected to be produced copiously, and furthermore to demonstrate the linear collider principle.
- (b) The superconducting conversion of the Fermilab machine was progressing well but was not yet fully completed.
- (c) The ISABELLE 400 GeV proton-proton collider had suffered earlier delays and was being reproposed as the "CBA." At the same time Fermilab proposed a 2 TeV vs. 2 TeV proton-antiproton collider, the so-called Dedicated Collider (DC).

5

These available proposals employ the same basic technology: the use of superconducting ring magnets as storage rings to produce proton-proton or protonantiproton collisions. Looking at the Livingston chart, one might just judge that the next construction project using this technology might well be the last. This in turn creates a strong presumption that this last step should be as large as might be economically sustainable. At the same time theoretical guidance indicated that at energies near 1 TeV in the quark-quark collision frame new and important entities might become evident.

In parallel with these fundamental issues remains the fact that the exploration of truly new technologies for high energy has been progressing relatively slowly. In this respect the SLC is a key undertaking which examines the feasibility of collisions between micron diameter beams. However until this has been demonstrated it is difficult to incorporate machines depending on the practicality of such collisions in specific future planning. Various proposals for basic new methods of acceleration are in their infancy. For instance, acceleration of beams of quality sufficient to lead to collisions useable for high energy physics based on lasers, wakefield acceleration, 2-beam acceleration, etc. is well into the future. I will discuss some of the reasons for this shortly.

Against this general background the Wojcicki committee decided, after extensive debate, to recommend termination of the CBA, and non-acceptance of the Dedicated Collider, and instead to recommend that a machine generating collisions of 20 TeV vs 20 TeV protons and aiming for a luminosity of 10^{33} cm⁻² sec⁻¹ be given the highest priority in the U.S. program. This recommendation was indeed difficult to reach, not only because in involved the termination of a project which had, after a rocky start, overcome most of its problems, but also because the total amount of analysis underlying the recommendation for the "highest priority machine" was very limited indeed. In fact only two significant organized efforts formed the basis of this recommendation: one was the Snowmass study organized by the Division of Particles and Fields of the American Physical Society, held from June 28-July 16, 1983, and the other was a one week long study of an ultra-high energy proton-proton collider which was held at Cornell University. The Snowmass study established the expected richness in terms of new physics of the multi-TeV collision energy range, and the Cornell study documented the fact that no orbit dynamics obstacles were known to exist which would make a 20 TeV per beam collider infeasible. Cost estimates were extremely uncertain at the time the HEPAP Subcommittee made its recommendations. The report of the HEPAP Subcommittee was accepted unanimously by the full HEPAP on July 11-12, 1983 and transmitted to DOE.

This history created a situation unprecedented in U.S. high energy physics, and in fact unprecedented internationally. A recommendation was made to discontinue a project to which a large amount of effort had been dedicated while neither a specific proposal nor a design study supporting the newly recommended project existed. However, at the same time a great deal of research and development is going on in U.S. laboratories, in particular with respect to superconducting magnets, which advance the technology on which the new machine called the SSC for Superconducting Super Collider would have to be based.

In view of all this complexity the U.S. government could not simply say yea or nay to the HEPAP recommendation. Rather, some of the funds supporting the BNL CBA program were reprogrammed to support further SSC research and development. A subcommittee of HEPAP was appointed to advise DOE on

é

7

distribution of effort in support of R&D for the SSC both at BNL and other laboratories. This committee, in turn, recommended strongly that a more formal R&D management structure be set up so that work done at the various laboratories in support of the future SSC would truly proceed towards a common goal. A difficulty is, of course, that in the absence of a formal decision to proceed or even an agreed conceptual design, it is difficult to establish central direction for a total R&D effort which is quite substantial but dispersed among several laboratories. Yet without additional work to put the project on a more solid basis it is justifiably difficult to get a formal decision for construction.

The present resolution of this problem was the establishment of a "Reference Designs Study" which is now proceeding under an April 30 deadline. This study, chaired by Maury Tigner of Cornell, is to provide a much better cost estimate than is currently available but is to focus on a single design for all elements of a new SSC laboratory with the exception that excursions using three different values of magnetic fields and corresponding machine radii are to be included. During this effort many very arbitrary decisions will have to be made which should in no way be binding on the design for the actual installation when it finally proceeds. Thus the term "Reference" designs. Hopefully, this design, in turn, after extensive review, will inspire enough confidence on the part of the government to generate a green light by early August to proceed with a formal Phase I, that is research and development work for this project. This Phase I will involve further refinement of magnet technology and much other R&D, but primarily will focus on the production of a formal detailed conceptual design and a proposal for construction. At this time it is not certain whether the SSC will in fact be built. The cost estimate which will result from the Reference Designs Study is, of course, critical. If it is too high, then specifications for this machine would have to be reduced or it would have to be built under international sponsorship or not at all. Quite apart from this are all the other questions relating to management, contractor, site selection, etc. which will involve many political as well as technical questions.

So much for historical and administrative factors. Let me turn to the technical inputs, to the extent we understand them, which affect the future of accelerators and colliders for high energy physics.

I said above that the SSC will probably be the last of the proton colliders built using what one now dares to call "conventional" superconducting magnet technology applied to proton storage rings and synchrotrons. The principal basic improvement which might go into such machines has to do with superconducting materials. The possibility of using niobium tin rather than niobium titanium is already being incorporated in the planning for the SSC. Such machines obey a linear scaling law between cost and energy and therefore would not be instrumental in following the decreasing cost per MeV trend indicated in Figure 6. There can, of course, be economies of scale as the size increases further, but there are also increasing cost dealing with requirements for more extended communication and greater reliability which are concommitant to increases of scale. For all these reasons I doubt that Fermi's projected machine which encircles the earth in a Saturn-like ring will ever become a reality!

The cost of electron-positron storage rings used as colliders scales roughly as the square of the collision energy. The reason is that the required RF power to compensate for synchrotron radiation varies as the fourth power of the energy divided by the machine radius, while most other costs go up roughly linearly with the radius. As a consequence an overall cost optimization leads to the quadratic scaling law. For this reason it is unlikely that an electron-positron storage ring collider larger than LEP II at CERN will ever be built; this machine has a circumference of 28 km and a cost near 1 billion dollars. Thus the established technologies both for electron machines and proton machines are approaching their limits, and new technologies must take over lest the entire high energy collider enterprise come to a halt.

All other potential candidates for high energy colliders using new concepts are in essence linear colliders whether they are for protons or electrons. At this primitive stage most discussions of the subject are restricted to obtaining higher energies at reasonable cost and within reasonable physical space, with luminosity being largely ignored. Yet it may well be luminosity which is just as serious an obstacle towards further advances as simple collision energy. If you accept the conventional wisdom that cross-sections decrease inversely as the square of the mass range of interest, then luminosities should increase correspondingly above the 10^{32} cm⁻² sec⁻¹ range which has remained an approximate goal for machines either for electrons or protons in the 100 GeV to 1 TeV range. If one translates this requirement into beam intensities and beam powers, one obtains the following equation:

$$P \sim 10^{-2} LE/(n/A)$$

where P is the average combined beam power in megawatts, E is the energy per beam in GeV, n/A is the number of particles per unit beam cross sectional area in units of 10¹⁰ particles per (micron)² and L is the luminosity in units of 10³² $cm^{-2} sec^{-1}$. Note also that \mathcal{L} should increase as E^2 to maintain a constant rate of interesting events!

Numerically this means that for linear colliders, either for electrons or protons, beam powers in the many megawatt range are required, while beam diameters must slip well below 1 micron. The SLC has, among other purposes, the goal of demonstrating the feasibility of collisions between micron diameter beams; the verdict is not as yet in, but no fundamental principles appear to be violated to make the SLC a success although there are many practical problems related to stability, alignment, and freedom from vibrations, and there are also very challenging requirements in generating the required high beam "brightness," that is concentrating the very big number if particles into a very small phase volume required for this application.

Much attention has been given to the production of high gradients, since the building of linear machines to reach many TeV using gradients as low as 20 million volts per meter (this is roughly the gradient at which the SLAC Linac will operate for the SLC) is thoroughly unattractive. While increasing the gradient will of course decrease land requirements, it is not at all clear that the highest possible gradient is the right choice when one looks at other requirements such as power economy, etc.

This is not the place to give more than an outline of the various ideas for new methods of acceleration. A whole class of devices hopes to exploit the very large electric fields contained in high intensity laser beams. Obviously a plane electromagnetic wave cannot accelerate particles. Thus laser-powered accelerators are based on the interaction of laser beams with external boundaries such that a wave with a longitudinal field component results. Such boundaries can be gratings or one can utilize propagation in a gas which yields an electromagnetic field configuration corresponding to the inverse of the Cherenkov effect. Another approach is to produce laser originated standing wave patterns in a gas. Such patterns, if intense enough, will induce a plasma in the gas of periodic density variation. These density variations will in turn generate electric fields which can produce acceleration.

None of these schemes has as yet been carried out far enough to evaluate their promise. However, extremely serious obstacles must be overcome both of a practical and theoretical nature. On the practical side we have the problem that the power tolerance of gratings, windows, or other devices might be a serious obstacle. In addition the power efficiency of the laser itself thus far has not exceeded several percent; if that is combined with the efficiency of the accelerating mechanism, then the expectation to accelerate multi-megawatt beams appears remote. There is some hope for the free electron laser which theoretically can have good efficiency but again we are talking about a very large extrapolation. On the theoretical side I might add that no careful work has as yet been done to examine the orbit dynamics of intense beams which might be accelerated by any of these devices.

Another class of novel accelerators might roughly be characterized as linear "step-up transformers." Let me explain. There has been a great deal of progress in generating multi-MeV beams of intensity in the kiloampere range. Stemming from this fact a class of inventions has been made which are, simplistically speaking, step-up transformers from such devices to low current, ultra-high energy accelerators. One approach (Voss and Weiland) employs a disc-loaded waveguide as a step-up transformer. The high current low voltage beam passes a region in such a waveguide where the electric field is low but the magnetic field is high, while the beam to be accelerated passes through an aperture where the inverse is the case. Many unsolved problems for this approach remain, in particular those related to higher mode excitations and various stability problems.

Another approach is to pass the high intensity low voltage beam through a set of decelerating structures and couple the electromagnetic energy induced in such structures through a series of waveguides into a more conventional RF linear accelerator. In other words, this approach uses the low energy high current beam "decelerator" as a continuous RF source to power the high energy, low current accelerator.

Then there is the large class of accelerators known as "collective devices." Generically this term describes a class of devices in which many degrees of freedom of charged particles combine to give their energy to a single particle. Specific applications of this principle have mainly focused on capturing a relatively small number of protons in intense electron beams and letting the space charge of the electron beams trap and accelerate the protons. The well-known "smokatron" accelerator pioneered by the Soviets is an example of this class of device. Enough work on these approaches has now been carried out to be fairly persuasive that these devices might be useful for heavy ion accelerators to moderate energies but do not have much promise as sources of high energy beams for elementary particle physics.

The one approach which has been analyzed somewhat further is to build a linear collider based on conventional radiofrequency structures, that is to have two radiofrequency linear accelerators aim beams at one another. Here the situation is somewhat different for protons and electrons. Since both proton linear accelerators and proton synchrotrons using superconducting magnets obey linear scaling laws, the economic issue is simply whether the cost per GeV of one or the other of these devices is lower. It is clear that a proton linear accelerator can only be the winner in this competition if new technology, in particular in respect to RF sources, is developed. As far as electrons are concerned, the case for linear colliders is somewhat more compelling—not by choice, but by necessity. This necessity arises, of course, from the quadratic scaling law of cost with energy which we discussed before.

A special problem for electron-positron linear colliders stems from the character of the beam-beam interaction. In circular colliders the beam-beam interaction sets the fundamental limit to the attainable luminosity because the non-linear focusing effects of one beam on the other limit performance. Conventionally this limit is specified by an empirical "tune shift" which is the shift in betatron frequency produced by the linear focusing action of one beam on the other. As a practical matter a figure of 0.06 for this tune shift is the largest number which has been attained for stable operation, and this in turn can be put into the standard expression for luminosity, leading to a practical limit for that quantity for a given available length between the magnetic elements surrounding the interaction region. This limitation does not apply to linear colliders since after beam-beam interaction the particles are discarded into a "dump." In fact the beam-beam interaction for electron-positron conditions is benign in terms of luminosity because the two beams focus one another, resulting in an enhancement factor H(D) which can become as large as six. Here D is the so-called disruption parameter which is the ratio of the longitudinal bunch size measured by its standard deviation σ_z to the focal length F produced by the beam-beam interaction. However, not all the beam-beam interaction effects in respect to linear electron colliders are benign. There is also the production of "beamstrahlung," that is the electromagnetic radiation produced by one electron or positron when traveling in the collective electromagnetic field of the other beam. This broadens the energy spectrum of the beams during interaction by an amount ΔE given by

$$\Delta E \propto n^2 E^2 / \sigma_r^2$$
; $\Delta E / E \propto L E$

where σ_r is the radial beam size. What energy spread is permissible here depends, of course, on the physics problem under investigation. Should there be narrow resonances at very high energies, then the principal effect of energy spread is to decrease the peak counting rate and interference with the measurement of the natural width of the peak. However, considering the previous proportionality relation, the peak counting rate can be adjusted to the same value independent of the beam-beam interaction by proper choice of the number of electrons per bunch, i.e. by trading luminosity and energy width. If one observes a continuum in energy then the beam broadening does not control the counting rate but simply affects the energy resolution with which one wishes to observe the process in question.

Thus as a practical matter the number of particles per bunch which is permissible is set by the required ratio of luminosity to percentage energy spread set by the experimental requirements. If we assume that this ratio is given by:

$$10^{33} \text{ cm}^{-2} \text{ sec}^{-1} = \mathcal{L}/(\Delta E/E) = .0024 f \sigma_z/E$$

where f is the repetition frequency in Hertz, E is in TeV, and σ_z is the bunch length in millimeters, then one can easily calculate that the beam power measured in megawatts is given by the expression:

$$P = (\mathcal{L}/3.5 \times 10^{31}) \sigma_z/DH(D)$$

where \mathcal{L} is measured in cm⁻² sec⁻¹. This, in turn, implies that there is relatively little latitude in design for a given luminosity to minimize the average beam power, since the quantities other than beam power in this equation are limited by relatively narrow bounds. In other words, the magnitude of the luminosity to energy width ratio can always be worse than that specified if we do not learn how to focus beams to submicron diameters, but it cannot continue to improve if we learn how to focus to better than the required amount.

Let me emphasize again that these basic considerations are totally independent of the mechanism which produces high intensity beams. At SLAC we have looked at the possibility of producing such high intensity beams in the TeV range with "conventional" rf acceleration and find that for a TeV per beam machine the resulting parameters are not really more extreme than for some of the SSC designs. As an illustration, Figure 8 gives a table of such parameters. As emphasized before, none of this should be taken too seriously pending the demonstration of usable micron-size beam-beam collisions in the SLC. Moreover, the rf systems envisaged in this illustrative example are certainly inelegant, and the conventional rf approach may not be the best way to produce such beams for linear colliders, depending on progress with the more esoteric methods.

I hope I have given you a rough overview of the expectations for the extension of high energy colliders and accelerators into the extremely high energy range. It appears likely that the SSC or something like it will be the "last gasp" of the "conventional" method of producing high energy proton-proton collisions using synchrotron rings with superconducting magnets. It is likely that LEP will be the highest energy e^+e^- colliding beam storage ring built. The future beyond that depends on the successful demonstrations of new technologies. The linear collider offers hope in this respect for some extension in energy for electrons, and maybe even for protons, but it is too early to judge whether, by how much, or when such an extension will indeed take place.

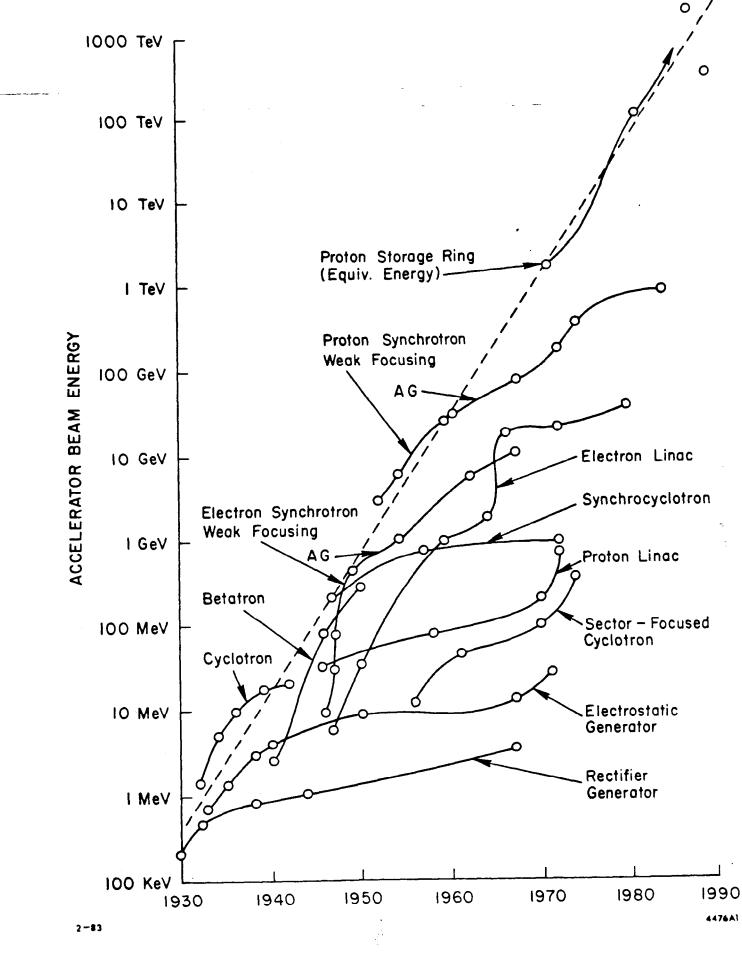


Figure 1

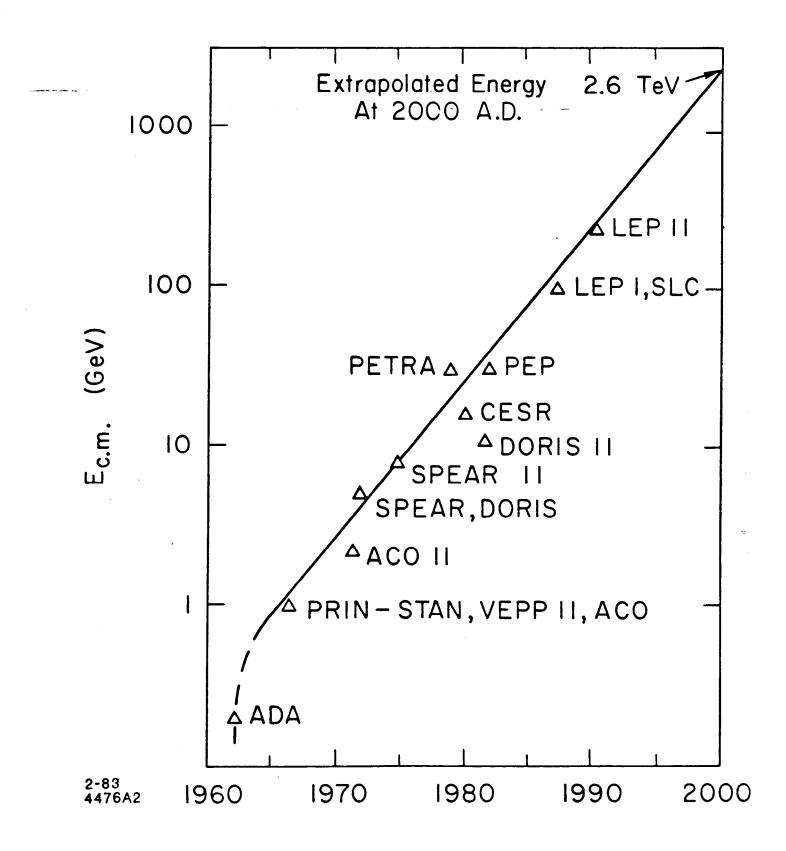
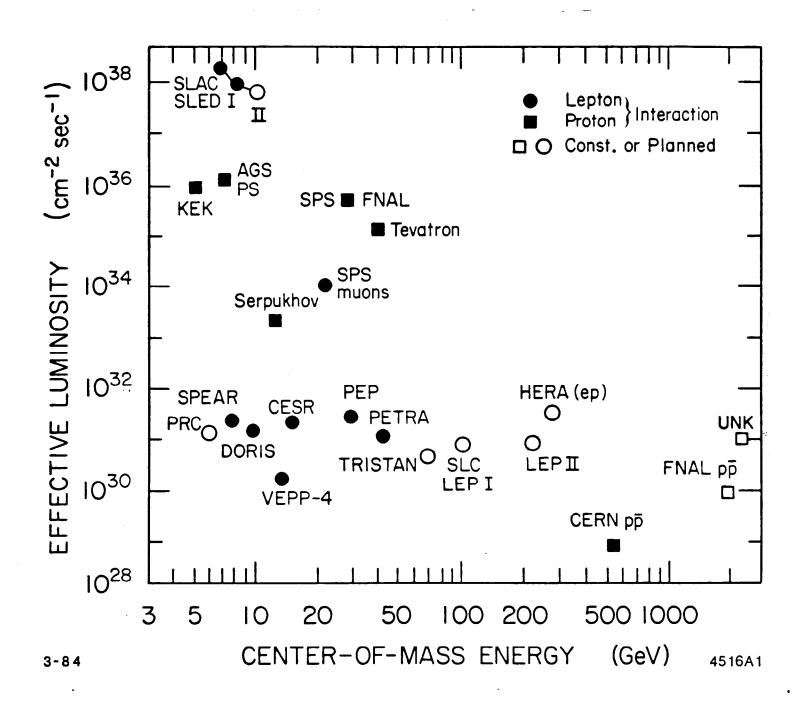


Figure 2



Effective luminosity vs center-of-mass energy for the largest accelerators and storage rings now operating, in construction or planned. For accelerators, the target is assumed to be one meter of liquid hydrogen, except in the case of "SPS muons," where 50 meters of LH_2 is assumed.

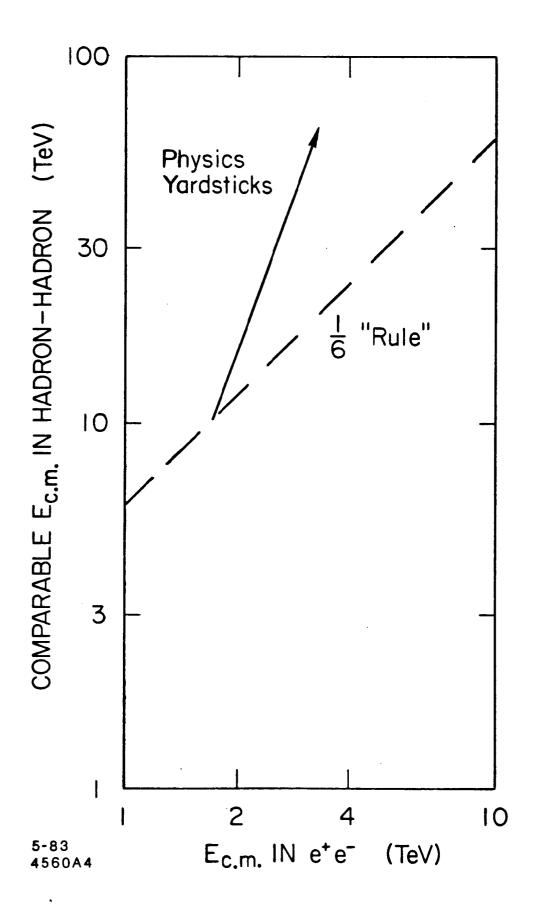
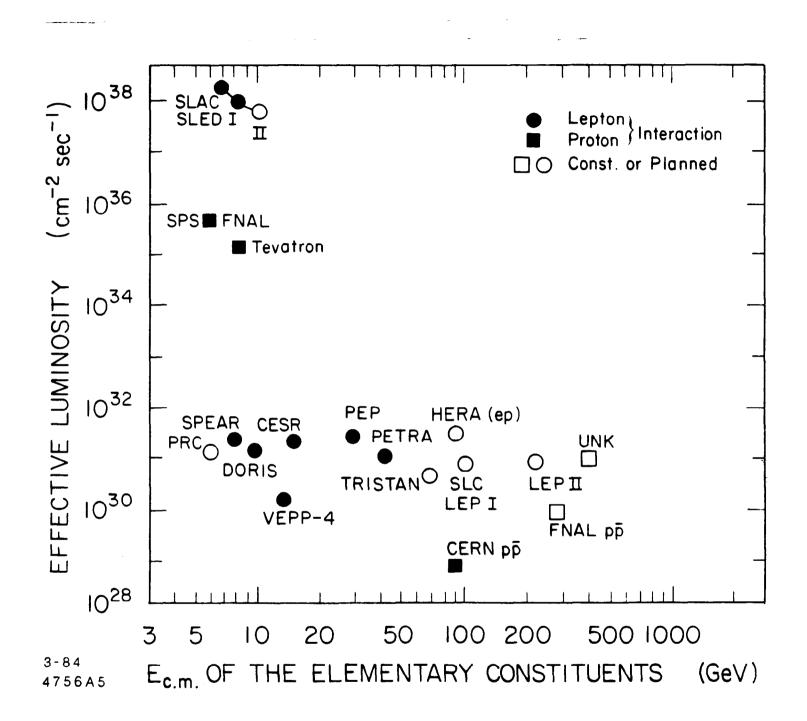


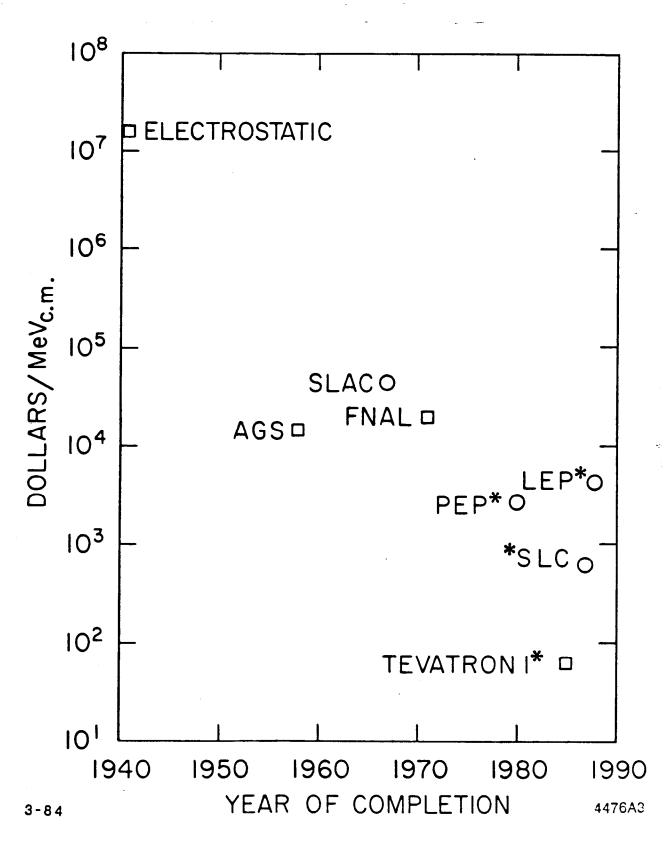
Figure 4

- .

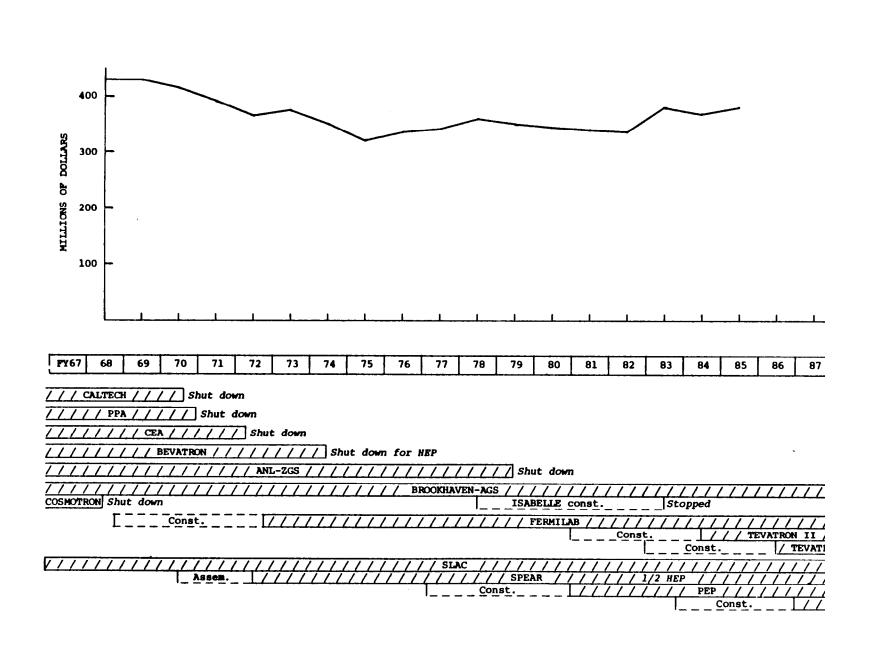


19.8%

Effective luminosity vs. center-of-mass energy of the elementary constituents for the largest accelerators and storage rings now operating, in construction, or planned. For accelerators, the target is assumed to be one meter of liquid hydrogen. Figure 6



* Based on incremental costs in existing laboratories and using existing injection systems



1. K. P. P.

U.S. DOE operations funding in equivalent FY1985 dollars, and "life and death" of U.S. facilities.

- 24 - 14 - 1

Figure 7

.

BASIC ACCELERATOR PARAMETERS FOR EACH ACCELERATOR

. ...

Beam Energy	1 TeV
Operating Frequency	2856 MHz
Length	50 km
Gradient	20 mV/m
Pulse Repetition Rate	180 pps
No. of Bunches / Pulse	12
Bunch Length (σ_{g})	2 20
No. of Electrons / Bunch (N)	1.4×10^{10}

44 Jan 1

ACCELERATOR STRUCTURE DESIGN PARAMETERS

Frequency	2856 MHz
Accelerator Structure	Disk-Loaded / Const.Gradient
Shunt Impedance (aver.)	55 MΩ/m
Q (aver.)	13,900
Attenuation Parameter	0.23 nepers
Filling Time	0.36 µsec
Length of Accel. Section	1.75 m
No. of R.F. Feeds	28571

FIGURE 8a

ACCELERATOR ENERGY-RELATED FACTORS

.

 $1201 \pm 10^9 W$ Total Peak RF Power $81.9 \times 10^6 W$ Total Aver. RF Power No. of Feeds / Klystron 8 3571 No. of Klystrons 336 MW Peak RF Power / Klystron 22.9 kW Aver. RF Power / Klystron 6λ (2.0 nsec) Bunch Separation 21.8 nsec Beam Pulse Length 1.23 A Peak Beam Current $1.36 \times 10^{12} W$ Peak Beam Power Aver. Beam Power (P) 4.83 MW Energy Carried by Beam/Pulse (W_b) 26.8 x 10³ J $3.09 \times 10^5 J$ Energy Stored in Accel./Pulse (W_{st}) 0.087 $W_b / W_{st} = \eta$ 48.0 GeV (4.8%) Beam Loading Modulator Effective Pulse Length 0.45 µsec P.S./Mod./Klystron Efficiency 50X RF Transmission Efficiency 90X Total AC Power Per Accel. 196 MW Total AC Power for 2 Accel. 392 MW

/1 All figures are for 1 accelerator except last line.

FIGURE 8b