Proceedings of the Linear Accelerator Conference 1990, Albuquerque, New Mexico, USA

BEAM DYNAMICS IN LINEAR COLLIDERS*

Ronald D. Ruth
Stanford Linear Accelerator Center (SLAC)
Stanford University, Stanford, California 94309

Introduction

In this paper, we discuss some basic beam dynamics issues related to obtaining and preserving the luminosity of a next generation linear collider. In Figure 1 you see a diagram illustrating the main subsystems of one-half of the collider. The beams are extracted from a damping ring and compressed in length by the first bunch compressor. They are then accelerated in a preaccelerator linac up to an energy appropriate for injection into a high gradient linac. In many designs this pre-acceleration is followed by another bunch compression to reach a short bunch. After acceleration in the linac, the bunches are finally focused transversely to a small spot.

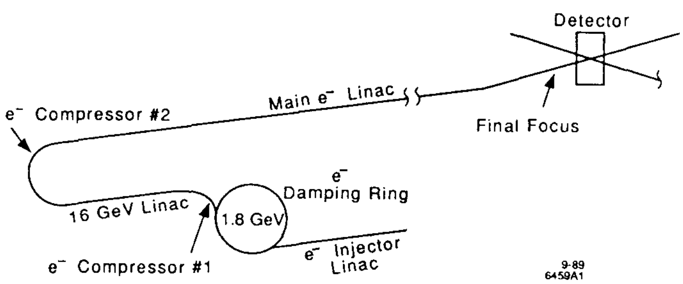

Fig. 1. Schematic of a Next Linear Collider

Before discussing each subsystem, it is useful to discuss the overall philosophy and parameters of this paper.[1,2] The energy range presently considered in various designs throughout the world varies from 1/2 TeV to 2 TeV in the center of mass while the desired luminosity varies from 10^{33} – $10^{34}\mathrm{cm}^{-2}\mathrm{sec}^{-1}$. The energy will be achieved by RF acceleration at acceleration gradient $\mathcal{E}_z$ for a certain length L. The acceleration gradients currently under consideration are in the 100 MV/m range while the RF frequencies range from 10–30 GHz. In this paper we only discuss the RF in so far as it affects the luminosity. Although obtaining the energy of a linear collider may be very expensive and require technical development, it is, in a sense, the easy part of the problem. The hard part is to obtain the luminosity.

The increase in luminosity over the SLC is obtained primarily in two ways. First, the spot cross-sectional area is decreased. Second, the energy extraction is improved by the use of multiple bunches per RF fill which effectively increases the repetition rate of the collider. Both of these techniques lead to many beam dynamics questions.

The proposed vertical beam sizes at the interaction point are the order of a few nanometers while the horizontal sizes are about a factor of 100 larger. This cross-sectional area is about a factor of 10^4 smaller than the SLC. However, the main question is: what are the tolerances to achieve such a small size, and how do they compare to present techniques for alignment and stability?

These tolerances are very design dependent. Alignment tolerances in the linac can vary from 1 μm to 100μm depending upon the basic approach. It is the premise of this paper that in order to achieve a next linear collider in this century, we must make design and correction choices which move most alignment tolerances into the 100 μm range. We begin the discussion with the damping rings.

Damping Rings

The SLC damping ring has achieved normalized emittances of $\gamma\epsilon_x = 3 \times 10^{-5}$ and $\gamma\epsilon_y = 5 \times 10^{-7}$. A next generation linear collider will need a horizontal emittance at least an order of magnitude smaller. In addition, most designs use $\epsilon_x/\epsilon_y \simeq 100$. This type of emittance ratio is naturally produced in an electron storage ring provided that the vertical dispersion and coupling are controlled. This sets tolerances for vertical alignment in the 50–100μm range which might be loosened by the addition of skew quadrupoles for compensation.

The ring designs typically include wigglers to decrease the radiation damping time. As mentioned earlier, most plans include the use of multiple bunches per RF fill. In order to efficiently use the circumference it is possible to damp several "batches" of bunches at once, each batch having the order of 10 bunches each. The batches must be separated by a distance which allows a kicker rise and fall time so that one batch can be extracted while allowing the remaining batches to continue damping.

Due to the small dispersion of the ring, the broad band impedance must be quite low ($Z/n \lesssim 0.5\Omega$) in order to avoid bunch lengthening. The long-range wakefield must also be controlled to avoid coupled-bunch instabilities. Because of the very close spacing of the bunches within a batch ($\sim$ 30 cm), inter-batch feedback would be quite difficult.

Example designs for a damping ring are given in Ref. 3. Aside from higher energy ($\sim$ 1.8 GeV) and larger circumference (155 m), this design uses combined function bends to enhance the horizontal damping at the expense of the longitudinal. Similar designs have been developed also at KEK, CERN and INP; therefore, it seems that damping rings which produce flat beams of the desired emittance are relatively straightforward.

* Work supported by the Department of Energy, contract DE–AC03–76SF00515.

Bunch Compression & Preacceleration

In order to prepare the bunches for injection into a high-gradient structure, it is necessary to reduce their length by bunch compression. Actually, there are two primary reasons for bunch compression. First, the bunch length must be less than the β^* at the interaction point. Since in many designs $\beta^* \sim 100\mu$m, we must have $\sigma_z \lesssim 100\mu$m. In addition, we should reduce the bunch length to reduce transverse wakefields.

If the bunch length and the relative energy spread in the damping ring are 5 mm and 10^{-3} respectively, then two bunch compressions are needed to reach 50 μm bunch length. Each compression reduces the bunch length by a factor of 10, and they are separated by a preacceleration section which reduces the initial relative energy spread at the second compression back to about 10^{-3}.

The energy spread is kept to $\sim 1\%$ during each compression in order to avoid emittance dilution due to chromatic and dispersive effects in the compressors. After compression the bunch must be matched into the linac lattice. Studies of bunch compressors have been completed and tolerances are presently under study.[4]

Linac[5]

Injection Errors

As the beam enters the linac, it is necessary to match the lattice functions to those of the linac. In particular the dispersion must vanish. For typical flat beam parameters, the beam size is about $2 \times 20\mu$m which yields a tolerance on dispersion D given by

$$\begin{aligned} D_y &< 0.2 \text{ mm} \\ D_x &< 2 \text{ mm} \quad . \end{aligned} \tag{1}$$

This is an additive effect. There are also multiplicative effects due to the mismatch of the lattice functions. If the beam were monoenergetic, these mismatches would not filament; however, since this is not the case, there will be some filamentation. Allowing for complete filamentation, the emittance dilution is given by[6]

$$\frac{\epsilon}{\epsilon_o} = \frac{1}{2}\left[\frac{\beta_o}{\beta}(1+\alpha^2) + \frac{\beta}{\beta_o}(1+\alpha_o^2) - 2\alpha\alpha_o\right] \quad , \tag{2}$$

where α_o and β_o are the matched values, and α and β are the mismatched values. For $\alpha = \alpha_o$ and $\beta = \beta_o + \Delta\beta$, this reduces to

$$\frac{\Delta\epsilon}{\epsilon_o} \simeq \frac{1}{2}\left(\frac{\Delta\beta}{\beta}\right)^2 \quad . \tag{3}$$

For incomplete filamentation, the emittance dilution will be somewhat less.

Wakefields and BNS Damping

Wakefields are a key problem not only for linear colliders, but for all accelerators and storage rings. The standard solution to this problem is to first reduce the wakefield forces until they are small compared to the applied external fields. Then compensation can be used in the form of feedback, or we can simply live within the limits by keeping the number of particles in the bunch sufficiently small.

For linear colliders the transverse wakefield within the bunch can be reduced first by keeping the RF frequency sufficiently small or by increasing the iris size. Secondly, the β-function in the linac must be kept sufficiently small. Then compensation can be applied by using BNS damping—the use of a correlated energy spread to cancel wakefield effects. The BNS correlated energy spread is given by[7]

$$\left(\frac{\Delta E}{E}\right)_{\text{BNS}} \equiv \delta_{\text{BNS}} = \frac{e^2 N W_\perp(\sigma_z)\beta_o^2}{4E_o} \quad , \tag{4}$$

where N is the number of particles, $W_\perp(\sigma_z)$ is the transverse wakefield evaluated at σ_z, and β_o is the β-function at energy E_o. For this paper I define a small wakefield by the condition

$$\left(\frac{\Delta E}{E}\right)_{\text{BNS}} \lesssim \frac{1}{\psi_{\text{tot}}} \lesssim 1\% \quad , \tag{5}$$

where ψ_{tot} is the total phase advance in the linac.

If the wakefield is large, then one can still satisfy Eq. (4) with a variation of focusing strength along the bunch rather than energy variation. In this case, however, coherent oscillations filament rapidly. To avoid emittance dilution with strong wakes, the alignment and trajectory tolerances are less than the beam size. This leads to 1 μm alignment tolerances.[8,9] As we shall see in the next sections, these tiny tolerances can be avoided by keeping the wakefields weak.

In the weak wakefield regime, BNS damping has been tested at the SLC linac.[10] In this case the tail growth due to a coherent oscillation was reduced by an order of magnitude. BNS damping has since been adopted as the normal running configuration for SLC.

Chromatic Effects

Upon injection into the linac, the compressed bunch has about a 1% uncorrelated energy spread. As the beam is accelerated, this relative spread decreases inversely with energy. At the same time a correlation between energy and bunch position is introduced due to the longitudinal wake and the curvature of the RF. Thus, the distribution in phase space becomes a wavy line which, when projected on the energy axis, yields an effective energy spread. At any location along the accelerator, the overall energy spread is a combination of the damping injected energy spread and the variation of energy along the bunch. After the bunch emittance is sufficiently damped, the relative energy spread remains constant unless deliberately increased by phase changes. For this reason it is useful to consider two models; one with constant energy spread and one with damping energy spread.

The first chromatic effect to consider is that of a coherent betatron oscillation. If the variation of the phase advance with momentum (chromatic phase advance) is much greater than unity, the oscillation filaments. In this case the oscillation amplitude must be less than the beam size to avoid emittance dilution. If the chromatic phase advance is small ($\delta\psi_{\rm tot} < 1$), then the tolerance on a coherent oscillation of size $\hat{x}_o$ is

$$\begin{aligned} \hat{x}_o &< \frac{\sigma_\beta}{\delta_o \psi_{\rm tot}} = \frac{\sigma_\beta}{\delta_o \psi_{\rm cell}} \frac{2}{N_q} \quad , \\ &< 2\sigma_\beta \quad ({\rm ILC}) \end{aligned} \tag{6}$$

where $\delta_o = 2 \times 10^{-3}$ is the constant relative momentum, $\psi_{\rm cell}$ and $\psi_{\rm tot}$ are the phase advance per cell and total phase advance respectively, and N_q is the number of quadrupoles. In all cases we give not only the formula but also the value for an example design of an Intermediate Linear Collider (ILC) of energy 0.5 GeV in the center of mass.[1,2] For the case of a damping energy spread with initial value $\delta_i = 0.01$, the tolerance is

$$\begin{aligned} \hat{x}_o &< \frac{\sigma_\beta}{\delta_i \psi_{\rm cell}} \frac{2}{N_q} \left(\frac{\gamma_f}{\gamma_i}\right) \\ &< 5\sigma_\beta \quad . \end{aligned} \tag{7}$$

For the case of a corrected trajectory let us consider the model of a sequence of random bumps. In this case the tolerance on the trajectory or alignment is

$$\begin{aligned} (\Delta x)_{\rm rms} &< \frac{\sigma_\beta}{\delta_o \psi_{\rm cell}} \left(\frac{3}{N_q}\right)^{1/2} \\ &< 30 \mu{\rm m} \quad , \end{aligned} \tag{8}$$

for a constant energy spread δ_o. For an initial damped energy spread δ_i, we have

$$\begin{aligned} (\Delta x)_{\rm rms} &< \frac{\sigma_\beta}{\delta_i \psi_{\rm cell}} \left(\frac{1}{N_q}\right)^{1/2} \left(\frac{\gamma_f}{\gamma_i}\right)^{3/4} \\ &< 30 \mu{\rm m} \quad . \end{aligned} \tag{9}$$

Misaligned Accelerator Sections

BNS damping only cures the growth and filamentation of coherent oscillations in the linac; it is an average compensation rather than a local one. In an actual linac, the wakefield kicks are not cancelled locally by adjacent quadrupoles. This leads to an incoherent growth of wakefield tails due to a random sequence of misalignments between the trajectory and the accelerator structure. If we parameterize the strength of the wakefield kick by $\delta_{\rm BNS}$ as defined in Eq. (4), the tolerance on random accelerator misalignments is given by

$$\begin{aligned} (\Delta x_{\rm structure})_{\rm rms} &< \frac{\sigma_\beta}{\delta_{\rm BNS} \psi_{\rm cell}} \left(\frac{3}{N_q}\right)^{1/2} \\ &< 25 \ \mu{\rm m} \quad . \end{aligned} \tag{10}$$

for $\delta_{\rm BNS} = 2.5 \times 10^{-3}$. From Eqs.(8) and (10) above, we see that the structure tolerances and quadrupole alignment tolerances are comparable provided that $\delta_{\rm BNS} \sim \delta_0$, that is, provided that the energy correlation needed for BNS damping is equal to the minimum energy spread in the linac.

Compensation of Chromatic/Wakefield Effects

The alignment tolerances shown above assume that the trajectory is a random sequence of bumps. There is no particular reason that it has to be random. Let us for the moment neglect wakefields. Then it is possible to measure the trajectories for particles of different energy and choose a trajectory which yields a small difference. Such a difference trajectory can be generated by scaling all the magnetic fields in the linac by a small amount so that the entire beam has an effective energy which is changed. By choosing the corrector sequence to minimize this difference trajectory (as well as the actual trajectory), the dispersion generated by misalignments is cancelled locally.

This technique is called dispersion-free correction. Provided that the beam position monitors have precision the order of 1 μm, it is possible to essentially decouple the quadrupole misalignments from the dispersive effects.[11] This increases the tolerances given in Eqs. (8) and (9) by an order of magnitude.

When we include wakefields, the coherent motion is BNS-damped and the incoherent motion gives rise to a random tail growth which can be controlled by tight tolerances. All that really matters for this effect is the value of the offset of the bunch within the structure. The offsets can be caused by two effects: misalignments of structures and trajectory offsets in structures. The trajectory is under our control; therefore, it is possible to use a trajectory which cancels the wakefield effects locally. Recently, T. Raubenheimer at SLAC has shown that by modifying the dispersion-free trajection technique, he can obtain a trajectory which cancels both the wakefield effects and the energy variation of the trajectory.[12]

Finally, we are left with the misalignments of accelerating structures. The most straightforward technique is to simply align the structure to the beam by using a BPM which is geometrically linked to the structure center. Such a BPM could consist of simply measuring the transverse wakefields induced by the beam.[13] One can use this information to either move the structure or move the trajectory to minimize the wakefield effects. Alternatively, for weak wakes, it is possible to deliberately move the beam or the structure to add a wakefield which cancels the effect of the rest of the accelerater.[14]

Beam Tilt

If there are RF kicks due to construction errors in the accelerator sections, the tail of the beam receives a different kick than the head. This can give a tilt to the beam. If we assume a random uncorrelated sequence of RF kicks, and compensate the center of the bunch with dipole correctors, the tilt tolerance is given by

$$(\Theta_{\rm rms})\left\{\beta_o < \sin\phi_o >_{\rm rms} \frac{2\pi}{\lambda_{\rm rf}} \sqrt{N} \left(\frac{\gamma_o}{\gamma_f}\right)^{1/2}\right\} < \frac{\sigma_y}{\sigma_z} \tag{11}$$

where $\Theta_{\rm rms}$ is the rms RF kick angle for a beam with energy γ_o, N is the number of accelerator sections and σ_z is the bunch length. For the ILC we have

$$\Theta_{\rm rms} < 2\mu\text{rad} \quad . \tag{12}$$

If such a kick is caused entirely by the systematic tilting of irises in a section (the bookshelf effect), then the tilt angle of the iris must be restricted by

$$\Theta_{\rm iris} < 0.3 \text{ mrad} \quad . \tag{13}$$

Jitter and Vibration: Motion Pulse to Pulse

Feedback is essential to handle the "slow" drift of x, x', y, y', E. In practical cases it is possible to feedback at $f \lesssim \frac{f_{rep}}{5}$. This sets the scale for what we consider slow. Time variation has many sources in linear colliders, for example: damping ring kicker jitter, power supply variations and ground motion. The jitter of the kicker in the damping ring must be kept small compared to the natural divergence of the beam at the kicker. Tolerances in power supply variations are also set in many cases by the beam divergence. The effects of ground motion depend upon the design and assumptions for the motion. If the wakes are weak and chromatic effects are kept small, there is no filamentation, and the beam moves coherently from pulse to pulse. If wakes are strong, and there is a large spread of betatron wave number, there is filamentation so that the beam size varies from pulse to pulse with a smaller centroid motion.

If we assume coherent motion, then for random magnet-to-magnet jitter the tolerance is

$$\begin{aligned}(\Delta x)_{\rm rms} &< \frac{\sigma_\beta F}{\beta}\left(\frac{3}{N_q}\right)^{1/2} \\ &< (0.04)\sigma_\beta \quad (\text{ILC}) \quad ,\end{aligned} \tag{14}$$

where F is the focal length of a lens. If, on the other hand, there is magnet-to-magnet correlated motion, then the dominant effect occurs when the wavelength is equal to the betatron wavelength. However, since the betatron wavelength changes $\propto \gamma^{1/2}$, the resonance is only temporary. If $2\pi\beta_i < \lambda < 2\pi\beta_f$, then the tolerance is given by

$$\begin{aligned}\Delta x_\lambda &< \sigma_\beta \frac{2}{(\pi\psi_{\rm cell})^{1/2}}\left(\frac{\gamma_f}{\gamma}\right)^{1/2}\left(\frac{2}{N_q}\right)^{1/2} \\ &< (.1 \text{ to } .4)\sigma_\beta \quad (\text{ILC}) \quad ,\end{aligned} \tag{15}$$

where γ is the energy at which $2\pi\beta = \lambda$.

Multibunch Effects

In order to efficiently extract energy from the RF, it is possible to accelerate many bunches per RF fill. This can increase the luminosity by an order of magnitude. To achieve the largest luminosity, we should put the maximum charge in a single bunch subject to restrictions on single bunch effects and beam-beam effects, then we should increase the number of bunches to extract as much energy from the RF as possible. This is not trivial in that the use of multiple bunches impacts every system.[15]

The most difficult problem, however, is the main linac, where the primary problems are bunch-to-bunch energy spread and transverse beam breakup. The basic tolerance for bunch-to-bunch energy spread is that it be less than the single bunch energy spread. This assures that the bunch-to-bunch chromatic effects will be no worse than single bunch ones.

Transverse beam breakup in the linac is a very difficult effect to control. For a normal traveling wave structure at 11.4 GHz, the 10th bunch blows up by many orders of magnitude by the end of the linac. Fortunately, there are solutions to this problem. The most useful approach seems to be the damping of the transverse modes in the structure to Q's $\sim$ 10–40 using external waveguides.[16] For the larger Q's, damping alone is not completely sufficient; however, if the frequency of the first higher mode is also adjusted with a tolerance $\sim 0.5\%$, the 2nd bunch can be placed near the zero crossing of the wake and the blowup vanishes.[17] This technique of damping high modes has also been shown to be useful for controlling coupled-bunch instabilities in the damping ring.[18] Recently, Q's as low as 8 have been measured in damped structures at SLAC.[19,20]

It is also possible to reduce the net wakefield by detuning the higher order modes in each cell of the structure. Provided that there is a large enough frequency spread, the net wakefield averages to zero by the time the second bunch arrives.[20] This technique might be much easier to implement if it proves to be successful.

Final Focus

Much progress has been made on the design of final focus systems.[21,22,23] As mentioned earlier, the final spot size desired is in the range 2–5 nm × 100–

300 nm. The limiting effect seems to be the radiation of the particles in the final quadrupoles which yields a minimum vertical spot size in the nanometer range.[24]

Once the design is specified, one is led to the question of the sensitivity of the design to different types of errors. The most serious vibration tolerance is in the final doublet, but there seem to be solutions to provide the required isolation.[25] Alignment tolerances in the absence of any correction are quite tight; however, it has recently been shown that one can recover from misalignments in the range 10–30 μm.[26] There is much more work to be done here, but the initial results indicate that tuning will be possible in the presence of errors.

A Final Focus Test Beam is presently being constructed at SLAC by a collaboration from SLAC, INP, KEK, Orsay, and DESY.[27] The purpose of this test is to study a flat beam final focus system which can demagnify the spot by a factor of about 300 in the vertical direction. This is precisely the demagnification necessary for the Next Linear Collider. For this experiment, due to the larger emittance of the SLC beam, the goal is to produce a spot with dimensions $\sigma_y \times \sigma_x = 0.06\ \mu\text{m} \times 1.0\mu\text{m}$.

Outlook

Before completing a realistic design of a next-generation linear collider, we must first learn the lessons taught by the first generation, the SLC. Given that, we must make designs fault tolerant by including correction and compensation in the basic design. We must also try to eliminate these faults by improved alignment and stability of components. When these two efforts cross, we have a realistic design. I believe this will not occur in the 1 μm alignment range. However, from the results presented here, I do believe that, with compensation, designs exist which move us into the 100 μm range and closer to a realistic design.

Acknowledgements

I would like to thank K. Bane, J. Irwin, S. Kheifets, N. Kroll, T. Raubenheimer, J. Seeman, and K. Thompson for their useful discussions, and T. Boysen for help in preparation of the manuscript.

References

1. *Linear Collider Working Group Reports From Snowmass '88*, Ed. R. D. Ruth, SLAC-Report-334 (1989).
2. *Proc. of the Int. Workshop on Next-Generation Linear Colliders*, SLAC-Report-335 (1988).
3. T. O. Raubenheimer, L. Z. Rivkin and R. D. Ruth, "Damping Ring Designs for a TeV Linear Collider," in Ref. 1 and SLAC-PUB-4808.
4. S. Kheifets, R. D. Ruth and T. H. Fieguth, "Bunch Compression for TLC," SLAC-PUB-5034.
5. R. D. Ruth, "Emittance Preservation in Linear Colliders," *Lecture Notes in Physics*, **296**, Eds. M. Month, S. Turner, Springer-Verlag, Berlin, and SLAC-PUB-4436 (1988).
6. W. Spence, private communication.
7. V. Balakin, A. Novokhatsky and V. Smirnov, *Proc. of the 12th Int. Conf. on High Energy Acc.*, 119 (1983).
8. J. Tuckmantel, "Beam Tracking With RF-Focussing in CLIC," CLIC-Note 87 (1989).
9. G. Guignard, these proceedings.
10. J. Seeman et al., SLAC-PUB-4968, to be published.
11. T. Raubenheimer and R. Ruth, "A Dispersion-Free Trajectory Correction Technique for Linear Colliders," SLAC-PUB-5355 (1990), submitted for publication.
12. T. Raubenheimer, "A New Technique of Correcting Emittance Dilutions in Linear Accelerators," SLAC-PUB-5355 (1990).
13. This technique was suggested by V. Balakin.
14. J. Seeman, "New Control of Transverse Wakefield in a Linac by Displacing Accelerating Structures," these proceedings and SLAC-PUB-5337 (1990).
15. K. A. Thompson and R. D. Ruth, "Multibunch Instabilities in Subsystems of 0.5 and 1.0 TeV Linear Colliders," in Ref. 1 and SLAC-PUB-4800.
16. R. B. Palmer, "Damped Accelerator Cavities," in Ref. 1.
17. K. A. Thompson and R. D. Ruth, "Controlling Transverse Multibunch Instabilities in Linacs of High Energy Linear Colliders," *Phys. Rev. D* **41**, No. 3 (1990) 964, and in SLAC-PUB-4801.
18. K. A. Thompson and R. D. Ruth, "Transverse and Longitudinal Coupled Bunch Instabilities in Trains of Closely Spaced Bunches," *Proc. of the IEEE Part. Acc. Conf.* and SLAC-PUB-4872 (1989).
19. H. Deruyter et al., "Damped Accelerator Structures," *Proc. 2nd European Part. Acc. Conf.*, SLAC-PUB-5263 (1990).
20. H. Deruyter et al., "Damped and Detuned Accelerator Structures," these proceedings and SLAC-PUB-5322, (1990).
21. K. Oide, "Final Focus System for TLC," in Ref. 1 and SLAC-PUB-4806.
22. J. Irwin, "The Applications of Lie Algebra Techniques to Beam Transport Design," to be published in *Nucl. Inst. & Meth.*, and in SLAC-PUB-5315 (1990).
23. J.J. Murray, K.L. Brown, and T.H. Fieguth, "The Completed Design of the SLC Final Focus System," *Washington PAC 1987:1331*, and SLAC-PUB-4219 (1987).
24. K. Oide, *Phys. Rev. Lett.*, Vol. 61, No. 15 (1988) 1713.
25. W. W. Ash, "Final Focus Supports for a TeV Linear Collider," in Ref. 1.
26. J. Irwin, K. Oide, private communication.
27. J. Buon, et al., "Final Focus Test Beam for the Next Linear Collider," *Proc. 2nd European Part. Acc. Conf.*, and LAL-RT-90-10 (1990).