THE DEVELOPMENT OF THE NEXT LINEAR COLLIDER AT SLAC

B. L. Bahr
Stanford Linear Accelerator Center
Stanford University, Stanford CA 94309

ABSTRACT

At SLAC we are pursuing the design of a Next Linear Collider (NLC) which will be ten times larger than the linear collider at SLAC. The luminosity is designed to be 10^33 cm^-2 s^-1 at the lower energy and 10^34 cm^-2 s^-1 at the higher energy. Our basic approach is to extend the techniques and accelerator physics which are being used for the SLAC to the next generation. To obtain the higher energy without excessive length we choose an RF frequency of 114 GHz which is four times the frequency of the SLAC linear. We would like to increase the accelerating gradient to 1000 MV/m in a few of these to go beyond the SLAC.

In order to obtain the higher energy without excessive wall plug power we focus the beams at the interaction point to a small ribbon. This controls the beamstrahlung radiation while allowing a small cross-sectional area. The luminosity is further enhanced by the acceleration of beam bunches on each cycle of the accelerator.

A possible layout for the NLC is shown in Fig. 1. There are two complete linear accelerators, one for electrons and the other for positrons. Each linac is supplied with particles from a damping ring followed by a preacceleration section consisting of two bunch compressors and a 10 GeV linac. After passing through...

---

This document is a presentation on the development of the Next Linear Collider (NLC) at SLAC. It discusses the design of a linear collider that is ten times larger than the existing SLAC linear collider. The luminosity is targeted to be 10^33 cm^-2 s^-1 at the lower energy and 10^34 cm^-2 s^-1 at the higher energy. The approach is to extend the techniques and accelerator physics used for the SLAC collider to the next generation. The goal is to increase the accelerating gradient to 1000 MV/m in a few cycles to go beyond the SLAC's capabilities.

The layout of the NLC is described, showing that it consists of two complete linear accelerators, one for electrons and the other for positrons. Each linac is supplied with particles from a damping ring followed by a preacceleration section with two bunch compressors and a 10 GeV linac. After passing through...

---

**DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED**
the main knots and hind legs system. The beams collide at a small crossing angle inside a particle detector.

In order to demonstrate the basic features of the NLC operation, consider the transport of electrons through the collider as a single particle with a mass of 10⁻¹⁵ kg. A beam of 10 bunches is separated at the source and accelerated to about 1.5 GeV in a pre-accelerator. The bunches are then injected into a damping ring that serves to reduce the transverse and longitudinal coherence among bunches. The purpose of the damping rings is to remove energy from the transverse and longitudinal modes of the bunches. The bunches are then accelerated to about 1.5 GeV and compressed to a small transverse radius into the main high gradient linac.

The final bunches are carefully steered and focused as the electrons are accelerated in small bunches in the main linac. Precisely matched in the final focus system, the bunches are focused by about a factor of 300 just before they collide at the IP with another bunch of positrons. A loop for the fast washout of the positrons leaves the interaction area when the beam of positrons hits a metal target in the main pion production. After the beams collide, their debris is directed out of the detector area at finite momentum.

The patent lists a few NLCs, one of which is a beam of 10⁻¹⁵ kg at 1.5 GeV. The LHC can be used to accelerate the transverse and longitudinal modes of the bunches as well as a pairing third bunch which is used to produce the next round of beams. This technique was possible because of the high energy of about 10⁻¹⁵ GeV, it is possible to bend 10th beams without loss of energy at relativistic dilution. This led to the characteristic beams of the SLC.

### Table 1: NLC Parameter Options

<table>
<thead>
<tr>
<th>Option</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>1 + 1/2 TeV</td>
<td>1 + 1/2 TeV</td>
<td>1 + 1/2 TeV</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2 x 10^{34}</td>
<td>2 x 10^{34}</td>
<td>2 x 10^{34}</td>
</tr>
<tr>
<td>Linac Length</td>
<td>14 km</td>
<td>14 km</td>
<td>14 km</td>
</tr>
<tr>
<td>Acceleration Gradient</td>
<td>100 MV/m</td>
<td>50 MV/m</td>
<td>100 MV/m</td>
</tr>
<tr>
<td>RF Frequency</td>
<td>11.4 GHz</td>
<td>11.4 GHz</td>
<td>11.4 GHz</td>
</tr>
<tr>
<td># Particles/Beam</td>
<td>6 x 10^{14}</td>
<td>3 x 10^{14}</td>
<td>6 x 10^{14}</td>
</tr>
<tr>
<td>Linac</td>
<td>1 x 10^{14}</td>
<td>3 x 10^{14}</td>
<td>6 x 10^{14}</td>
</tr>
<tr>
<td>FF</td>
<td>1 x 10^{14}</td>
<td>3 x 10^{14}</td>
<td>6 x 10^{14}</td>
</tr>
<tr>
<td># Bunches, mg</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Beamstrahlung, %</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Reptition Rate</td>
<td>100 Hz</td>
<td>100 Hz</td>
<td>100 Hz</td>
</tr>
<tr>
<td>Wall Phg Power</td>
<td>30 MW</td>
<td>30 MW</td>
<td>30 MW</td>
</tr>
<tr>
<td>IP Beam Size</td>
<td>4 mm</td>
<td>4 mm</td>
<td>4 mm</td>
</tr>
<tr>
<td>#</td>
<td>200 mm</td>
<td>200 mm</td>
<td>200 mm</td>
</tr>
<tr>
<td>σ_{x}</td>
<td>100 μm</td>
<td>100 μm</td>
<td>100 μm</td>
</tr>
<tr>
<td>σ_{x}</td>
<td>0.15%</td>
<td>0.15%</td>
<td>0.15%</td>
</tr>
</tbody>
</table>

#### 2 The First Linear Collider: The SLC

During the past 20 years, the first linear collider the SLC was constructed and developed. This design uses the existing linear in a novel way to

 accelerate the electron and positron colliding bunches as well as a second round of beams. This technique was possible because of the high energy of about 10⁻¹⁵ GeV, it is possible to bend 10th bunches without loss of energy at relativistic dilution. This led to the characteristic beams of the SLC.

In spite of the distinction of a folded design, however, the SLC is the technology and accelerator physics base for the Next Linear Collider. During the detailed design and subsequent commissioning and development of the SLC, we have had to face the real issues which limit performance of linear colliders. Many lessons have been learned. In fact, hundreds of articles have been written on the details of design, and operation, a working linear collider. In this section we will discuss a few of the highlights from SLC which form the foundation of SLAC's approach to the NLC.
2.1 Low Emittance Production

The SLC damping rings were designed to provide very low emittance electron beams at a very high repetition rate (160 Hz). Although the technology and accelerator physics of electron storage rings is well understood, the SLC damping rings were the first designs which combined very low emittance and high repetition rate. The basic design of these rings has been confirmed in practice; they achieve their design asymptotic emittance.

One of the most critical components of the entire collider has turned out to be the damping ring extraction kicker. During SLC development it has been necessary to develop very stable kicker power supplies and also techniques for shaping the pulse of the kicker to vary the kick separately on each of two extracted bunches. This will be especially important in an NLC where trains of 10 or more bunches will be extracted at once.

Damping rings are especially sensitive to the longitudinal impedance seen by the beam as it passes through the vacuum chamber. An excessive impedance causes bunch lengthening and an increase in the relative energy spread of the bunches. The bunch lengthening observed in the SLC damping rings agrees very well with theoretical bunch lengths using detailed calculations of the damping ring impedance.

Finally, upon extraction from the SLC damping rings, it is necessary to compress the bunch length from about 7 mm to 5 mm. This requires a transport line with careful dynamic and dispersive compensation. Although the SLC bunch compressor initially did not perform as expected, they also were responsible for a stable portion of the emittance dilution. To eliminate this dilution, it has been necessary to develop both linear and nonlinear compensation techniques to match the beam from this complex transport line into the linac to be accelerated.

2.2 Application

This brings us to the heart of any linear collider, the main linac. Although the SLC retains the basic accelerator structure and components from the SLC beam line, it was necessary to modify many of the remaining parts of the RF power system and low power phase control system. New klystrons were developed which produce 45 MW of peak RF power at 2.36 GHz in pulses 3.5 ns long. Although at one time these klystrons had development problems, they are now very reliable and have lifetimes far longer than initially expected. This experience is being directly applied to the development of klystrons for the NLC at 11.4 GHz.

The power level of 60 MW is not sufficient to reach the required acceleration gradient of 17 MV/m. A new device called SLED was invented at SLAC which compresses RF pulses while increasing the peak power up to about 100 MW. This has proved a valuable tool and a new modification, dubbed SLED II, is planned for the NLC.

- The phase control requirements and the management of the linear energy profile are much more critical in a linear collider than in a conventional linac. It is necessary to shift the phase of each bunch on the RF wave by a precise amount in order to compensate single-bunch energy spread. The multibunch energy control is provided by matching the rate of extraction of energy to the rate of input. All of the technology, control system and accelerator physics of this acceleration process has been tested in detail in the SLC main linac. The problems in an NLC are very similar to those which have been solved for the SLC; the primary difference is the factor of four higher frequency.

2.3 Emittance Preservation

Acceleration and transport in the main linac can lead to the dilution of the beam emittance. Several developments at the SLC have made it possible to preserve the low emittance produced by the damping rings. The routine beam diagnostic measurement at SLC is done with a precision of 20 μm. With these precise measurements it is possible to control the trajectory of the electron beam at the 20 μm level. With this level of precision, beam-based alignment has been used to align the entire SLC linac.

Single bunch transverse beam breakup can be cured by using a correlated energy variation (BNS damping). This technique was experimentally tested at the SLC and is part of the normal running configuration. Residual tail growth of the bunches is compensated by deliberate bunch offsets to cancel the growth.

The BNS correlated energy spread is a measure of the relative size of the wakefield forces and the external focusing forces. In the SLC, at design current, the value is several percent. In this case alignment tolerances tend to be of the order of the beam size. For the NLC we would like tolerances much smaller than the beam size. To do this, it is necessary to have a much smaller BNS energy spread. For the NLC design at SLAC the BNS energy spread is a few parts of a percent, an order of magnitude smaller than that for the SLC.

While these compensations are necessary to achieve robust stable running, they are not sufficient to prevent drift which causes the beam to deteriorate. This drift is held in check with a sequence of feedback systems which steer the beam to a reference trajectory. For the NLC this puts special emphasis on keeping the repetition rate high enough to provide effective sampling for feedback.

2.4 Matching Diagnostics and Phase Space Certification

In order to match the beam condition to its ideal value, it is necessary to measure beam sizes and lengths rather precisely. In the SLC the beam size measurement is done with wire scanners for beam sizes in the range 1.5 μm to 100 μm. This information is used to adjust compensating magnets in order to
match the beam to its desired size and divergence. In the longitudinal direction it is very difficult to change the bunch length directly, however by using a longitudinal energy modulation together with known dispersion the bunch lengths are measured down to the smallest minimum bunch length provided by the bunch compressors. This bunch length is adjusted as the intensity is varied in order to achieve the optimum energy spread at the end of the linac. All of these techniques can be directly applied to the NLC.

**Foil Beam Transport**

After the beams leave the hall, they are separated and travel through two arcs which lead to the NLC final focus. These arcs are constructed in a plane and direct transport to the beam without distorting the structure. Because of the very strong focusing, the quadrupole magnets are used as correctors in adjusting the position with respect to the main magnets. The NLC arcs have provided the first large scale experience with computer controlled alignment. In order to compensate errors due to misalignments in the SLAC arcs, the foil transport magnets have been completely restructured using beam data. This information has been used to absorb uncompensated system to bring the beam to the desired performance. This type of compensation may also be necessary to control collisions in the NLC bunch compressors systems.

**SLC Final Focussing**

Perhaps the most well known about the final focus system is actually up stream of the beam transport to the storage rings necessary to collimate the beams well before they enter the final focus. This process is essential for successful physics runs and will also be important in the NLC. The design of the SLC final focus has played an important role in the beam's behavior and has provided a challenge to optimize the system. From time to time the beams are scanned across each other. This is the characteristic of a large collection process which can be used to determine the positions of the spots as well as the position for low beam collisions. This spot information is used to adjust the magnetic configuration in order to tune the spots with minimum configuration to achieve minimum spot size. For the NLC design we have found the importance of separating the functions of phase space matching from the final focus. We have also learned the importance of detailed tuning strategies which are already being applied to the design of the Final Focus at the SLC (see Section 6.1).

**Apparatus**

Some of the most important lessons learned thus far from SLC experience are also now difficult to quantify. There are few technically serious problems. Qualitatively as a system, an electron storage ring is rather like a damped pendulum. It has a stable configuration which is only perturbed athetically by parameters. The system is perturbed every time and then for injection, but in all other aspects it is a steady state system.

**In a linear collider, the beam is accelerated on each cycle of acceleration.** The electrons remember their initial conditions and all downstream parameters are affected by changes upstream. As a system, the linear collider behaves rather like a pendulum with your best hope in order to keep the balanced pendulum stable it is necessary to provide feedback at the base. For the linear collider feedback is the key to stable operation. It must deal with changes every 10 pulses or so as well as day night drift due to small temperature changes.

The stable operation of the system is also coupled tightly to the overall complexity and the reliability of individual components. In order to provide sufficient time for stable running, it is necessary to eliminate tuning of the linear chain of components. High reliability and stability help reduce this to a minimum. In a complex system such as a linear collider, one cannot expect peak performance from all systems simultaneously. For reliable operation experience at SLC indicates that a significant margin is required in each system.

### 3 Obtaining the Energy

#### 3.1 The Basic Approach

The energy for the NLC is obtained by a combination of length and acceleration gradients. At SLAC we have chosen to push acceleration gradients a factor of three to up to 300 MeV. In order to accomplish this while still controlling the valid power, we increase the frequency by a similar factor up to 3.4 GHz. Other than this change, the NLC RF system is modeled after the successful SLC RF system. A schematic of the system is shown in Fig. 2. The RF power flows from the source through waveguides to a travelling wave structure. After the structure is full, the beam is accelerated and the remaining energy flows out of the structure into a cooled load. Perhaps the best place to begin the discussion of the NLC RF system is at the RF structure.

#### 3.2 The RF Structure

The RF structures planned for the NLC are rather similar to those presently being used at the SLC except for the factor of four change in frequency. The beam parameters are shown in Table II. The key difference is a modification in design to reduce the transverse wakefields induced by bunches which are offset in the structure. First in order to control the wakefield within a bunch the emittance has been increased by about a factor of two relative to the wavelength. Although this slightly increases the required power, the short range transverse wakefield is reduced by about an order of magnitude by this simple modification. Second,
long range transverse wakefield

This can be accomplished by two techniques. In the first method, the cavity design is altered so that the deflecting fields are strongly coupled to external waveguides. After a bunch passage, the fields in the structure die out quickly as they propagate out the waveguide into a matched load. The second technique relies on the cancellation of the deflections from cell to cell. If the cells in a single structure are designed so that the deflecting modes oscillate at different frequencies, then the average deflection over the structure effectively damps due to the decoherence of the various cell wakefields.

Both techniques have been theoretically studied and experimentally tested at SLAC. Present designs at SLAC are focusing on the "detuned" option due to its simplicity. An experimental test of this technique is shown in Fig. 3, together with theoretical estimates of the wakefield for a model detuned structure. This model used a Gaussian distribution of frequencies for the higher-order modes, but the effect was exaggerated in a structure with a small number of cells. The experiment was a success and showed excellent agreement with theory.

![Image of transverse wake potential](image-url)

Fig. 3. Transverse wake potential for the HEM13 detuned cavity disk-loaded structure. Top: calculation by LINACBIBO ($\eta = 0$). Bottom: Measurement result at AART ($\eta = \pm 0.5$ mm).

The actual structure will have a Gaussian distribution in the higher-order modes which is tailored by adjusting the cell dimensions while keeping the funda-
In order to verify that it is possible to reliably accelerate beams with gradients in the range 50-150 MV/m, we are conducting high power experiments with model structures. The latest high field test was conducted on a seven cell 11.5 GHz standing wave structure. Although the peak field produced was limited by the available input power, the structure was powered to peak surface fields in excess of 500 MV/m. This would correspond to a travelling wave accelerating field of about 200 MV/m.

High power tests in 1992 will include a 30 cm long structure and a 75 cm long structure. The focus of these experiments will be on the generation of dark current due to field emission. Early indications are that the dark current will not be a problem at 50 MV/m, but may cause problems for acceleration gradients in excess of 100 MV/m.

3.3 The RF Power System

The elements of the NLC RF power system are shown in Fig. 6. The RF power is supplied by a klystron amplifier. In order to obtain the correct pulse length and peak power, the RF pulse is compressed by about a factor of six by an RF pulse compression system. The overall system is supplied by a modulator (unfed power supply). The primary goal for the system is to provide RF power with the correct peak power and pulse length as efficiently as possible.
The objective of RF pulse compression is to convert a long RF pulse of moderate power into a short RF pulse with high power. Ideally, a factor of five decrease in pulse length could yield a factor of five increase in peak power. Due to inefficiencies, the factor is always somewhat less. The RF pulse compression system SLED (SLAC Energy Development) is presently used at SLAC to boost the klystron power by about a factor of three before powering the SLAC linac. This system uses storage cavities to allow the RF to build up. A phase switch from the klystron causes the klystron power and the power emitted from the cavity to add coherently, yielding a narrow pulse of high peak power. Unfortunately, this system gives a pulse shape which is sharply spiked due to the exponential decay of the fields in the storage cavities. For an NLC it is useful to have a flat-top pulse to control multibunch energy spread.

This flat-top pulse can be obtained by two different methods. The first method, called binary pulse compression (BPC), uses delay lines to delay the leading portion of an RF pulse so that it is coincident in time with the trailing portion. This yields an RF pulse which is one half as long, but with nearly twice the power. This process can be repeated in a sequence to achieve more and more multiplication. Due to losses in components and waveguides, the method is limited to about three compressions.

Figure 7 shows a schematic diagram of a two-stage BPC system which was constructed at SLAC. The 3-dB hybrid shown in Fig. 7 is a four-port device which combines two power inputs into one or another output port depending upon phase. In this way phase shifts can be used as high power RF switches. A three-stage system of analogous design has been constructed at SLAC and has achieved a multiplication factor of 4.5 while reducing the pulse length by a factor of eight. This system, together with new high power klystrons, is used to test the RF structures described above.

![2 STAGE BPC SCHEMATIC](image)

One disadvantage of the BPC method of pulse compression is that it uses rather long delay lines. The waveguides which are used have a group velocity very close to the speed of light, and they are only used once as transmissive delay lines. This problem has led to the development of a new pulse compression scheme called SLED II. The system as shown in Fig. 8 is a similar to the SLED system at SLAC except that the cavities for storing the RF are replaced by resonant delay lines. Each of these delay lines has a round-trip delay time equal to the output pulse length. A resonant buildup of energy stored in the lines takes place during an input pulse length which is an integral number of delay periods, typically in the range of four to eight. A phase reversal of the input pulse effectively triggers the compression to produce a flat-top output pulse during the final delay period. An example of a SLED II pulse compression by a factor of four is shown in Fig. 9. Measurements from a low-power SLED II system with a power gain of four have shown excellent agreement with theory.

![Fig 8. A block diagram of SLED II](image)

Comparing SLED II with BPC, the amount of waveguide delay line to achieve a similar compression is reduced by about a factor of five. This is due to the...
although negative in nature, the delay lines are used repeatedly as the RF waveform builds up. In addition, there is the use of the SLED II system in series to provide even higher compression of the beam. A high power SLED II pulse compression system will be constructed at SLAC in 1992 to investigate this promising technique further.

5.1 The Klystron

Although a 50 MW klystron would probably suffice for a 500 GeV linear collider, the goal for klystron development at SLAC is to produce a reliable and efficient 100 MW 11.4 GeV klystron with a pulse length of about 1 μs. The design of these klystrons is based on the successful experience with 65 MW SLC klystrons during the past eight years. There are few differences due to the higher power and frequency. The solenoidal field which confines the klystron beam is increased to decrease the beam size transversely. The beam energy is increased and the electric fields at the output are increased due to the smaller wavelength. A summary of klystron performance thus far is shown in Fig. 10. The third klystron AC3, reached power levels similar to those reached by AC2.

![Measured Power Limits](image)

**Fig. 10** Summary of klystron performance

The problems which have thus far limited performance are related to the tube power and clock frequency. These are RF breakdown at the output window, inhomogenity, and vacuum breakdown. Each of these problems is being addressed in new designs which will be tested in 1992. We hope by the end of 1992 to have a reliable 11.4 GeV klystron which produces more than 50 MW in a pulse duration 1 μs long.

5.6 The Modulator

The modulator is a critical component in an NLC due to its impact on cost and efficiency. The challenge is to produce a beam voltage pulse in the range of 400-600 kV for about 1 μs with a rapid rise time to maintain efficiency. Conventional SLAC type modulators have slow rise times due to the 200 Hz ratio of the transformer. A new design developed at SLAC uses a multiplying Blumlein pulse forming network which allows an reduction in the rise time of the pulse transformer to about 6. It is output pulse shape for this new design is compared with the present modulator pulse in Fig. 11. The improved rise time and a substantial gain in efficiency are evident from the figure. The new design will be tested in 1992 by converting one of our present klystron test stands into the new design.

5.7 The NLC Test Accelerator

The previous sections have discussed the status and plans for the RF system and acceleration studies for an NLC. In order to integrate these separate development efforts into an actual N-band accelerator capable of accelerating beams necessary for an NLC, we propose to build an NLC Test Accelerator (NCTA). The goal of the NCTA is to bring together all the elements of the entire acceleration system to construct and test the components and modules of a section of a high gradient linac suitable for the next linear collider. The NCTA will serve as a test bed for the design of the NLC studies and will provide a model on which a reliable cost estimate of an NLC could be based. In addition to testing the RF acceleration system, the NCTA will be able to address many questions related to the dynamics of the beam during acceleration.

5.7.1 Basic Parameters

The NCTA consists of a high gradient N-band linac injected by a simple
4. Obtaining the Tolerance

The table below shows the tolerance values for different parts of the system. The tolerances are specified in millimeters (mm) and are critical for ensuring precise alignment and operation of the system.

<table>
<thead>
<tr>
<th>Component</th>
<th>Tolerance (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF Power Coupler</td>
<td>0.4</td>
</tr>
<tr>
<td>Power Supply</td>
<td>0.1</td>
</tr>
<tr>
<td>RF Module</td>
<td>0.2</td>
</tr>
<tr>
<td>RF Amplifier</td>
<td>0.1</td>
</tr>
<tr>
<td>RF Filter</td>
<td>0.05</td>
</tr>
<tr>
<td>RF Transformer</td>
<td>0.3</td>
</tr>
<tr>
<td>RF Cavity</td>
<td>0.2</td>
</tr>
<tr>
<td>RF Mixer</td>
<td>0.1</td>
</tr>
<tr>
<td>RF Mixer Mount</td>
<td>0.2</td>
</tr>
<tr>
<td>RF Mixer Mount Base</td>
<td>0.3</td>
</tr>
<tr>
<td>RF Mixer Mount Base Holder</td>
<td>0.4</td>
</tr>
<tr>
<td>RF Mixer Mount Base Mount</td>
<td>0.1</td>
</tr>
</tbody>
</table>

The tolerances are strictly adhered to ensure the system's performance and reliability.
The luminosity is the same as for a circular collider except that there is an additional term, an enhancement factor due to the mutual bunching of the beams. The luminosity is given by

\[ L = \frac{N_s N_b f_{\text{rep}}}{\alpha_s \alpha_b} \]  

(1)

where \( N_s \) is the number of positrons per bunch, \( f_{\text{rep}} \) is the repetition frequency, \( N_b \) is the number of bunches accelerated on each cycle of the accelerator, \( H_p \) is the bunch enhancement factor, and finally, \( \alpha_s \) and \( \alpha_b \) are the rms beam size of the insertion slit at the interaction point. Each bunch is assumed to collide with only one other bunch in the opposing bunch train.

The objective is to increase the luminosity to \( 10^{34} \) cm\(^{-2}\)s\(^{-1}\) for the energy range 2 to 10 TeV. To do that, we must minimize the denominator of Eq. 1 and decrease the enhancement factor as much as possible. For the luminosity, we have at our disposal the number of particles per bunch, the repetition rate, and the number of bunches on each cycle that we must satisfy the constraint that the wall plug power is less than 100 MW. To decrease the cross section, we have to decrease \( \alpha_s \) and \( \alpha_b \). To do this, we must keep the beam flat to control beamstrahlung.

In the next few sections, we discuss each term in the luminosity formula. The discussion of beam size is subdivided into a few sections. In the next section we begin with the denominator of Eq. 1.

### 2 Inter-particle and Repetition Rates

First, let's discuss the single bunch intensity \( N_s \) and the repetition rate \( f_{\text{rep}} \) from conservation of energy, we must have:

\[ M_{\text{wall}} = \eta_{\text{rep}} P_{\text{wall}} \]  

(2)

where \( \eta_{\text{rep}} \) is the efficiency for converting wall plug power to RF power, \( M_{\text{wall}} \) is the fraction of the energy extracted by a single bunch, and \( P_{\text{wall}} \) is the total wall plug power supplied to the bunches. The wall plug to RF efficiency, \( \eta_{\text{rep}} \), is about 30% for the proposed RF system. This is a fairly realistic estimate including all of the losses in the RF system between the RF power and the RF power, which were discussed in the last section. There are new ideas which could raise this to perhaps 50-60%, however, with the system shown in Section 3. \( \eta_{\text{rep}} \) is about 30%.

For some what different reasons, the single bunch extraction efficiency is limited to about 2%. The single bunch beam loading causes an energy spread within the bunch. Although the beam can be compensated by shifting the RF phase to obtain a linear shape, the higher order effects are difficult to compensate. However, very small energy spread is required to keep the beam/current from increasing, to keep the beam in a small spot, and finally, to provide a narrow spectrum for experiments. This limits single bunch energy extraction to a few percent.

For the purpose of this discussion, let's select a wall plug power of 100 MW for an \( f_{\text{rep}} = 1 \) Hz.

Because the required bunches have a very small transverse dimension, it is necessary to control the offset pulse to pulse with a feedback system. In order for this feedback system to work efficiently, the sample rate must be at least six times the rate at which the beam centroid is changing. Because ground motion is an important source of bunch motion, and because the spectrum drops off rapidly above 10 Hz, the repetition rate of the accelerator must be greater than 60 Hz. Experience at SLAC has shown that a simple factor of six is not sufficient. A substantial offset in the sample frequency is required in order to avoid excessively amplifying the high frequency ground motion. Therefore, we set the repetition frequency to 180 Hz. It could be dropped as low as 120 Hz, however, 90 Hz is probably too low. Substituting the previous parameter values in Eq 2, we find that the maximum number of particles per bunch is \( N_s = 2 \times 10^{11} \).

### 2.1 The Number of Bunches

As discussed in the introduction, the designs for the SLC include the need for many bunches on each cycle of the collider. The purpose of this is, of course, to increase the luminosity bunch by bunching number of bunches. If there were no constraints, the largest luminosity would be obtained by placing all the charge in the bunch into one bunch because in this case there is a quadratic gain with increasing intensity. As discussed in the previous section, the single bunch intensity is limited by the amount of energy that can extract while returning a small relative energy spread. It turns out that this intensity is also consistent with transverse stability and with beam beam effects. Thus, the quadratic gain is stopped by these limits, however, since there is about 98% of the energy left in the structure, it is possible to continue to gain linearly by increasing the number of bunches.

A large number of bunches brings along a host of other complications. Some of these were discussed earlier. The bunches must be stable transversely which means that the structure must be designed in a special way (Section 3.2). The energy spread to bunch must be controlled. This can be done by delaying the filling of the RF structure and by matching the rate of filling of the structure to the rate of energy extraction. The technique is used routinely at the SLC. However, as the number of bunches increases, higher order effects become important and the compensation technique becomes more difficult. This limits the number of bunches to about 10, although the single bunch intensity can be traded off somewhat with the number of bunches. There are other strategies for bunch to bunch energy extraction.
compensation which allow larger energy extraction; however, because of the heavier beam loading, the tolerances are much tighter.

The RF pulse must be of rather high quality. Systematic phase and amplitude variations over the bunch train must be less than about 2% (such tolerances are not realistic with the power sources discussed). Because a significant fraction of the fields felt by the trailing bunches are due to the leading bunches, the intensity of the bunch train must be controlled to a precision less than 2%. The damping rings which produce these trains of bunches must be able to accelerate them without instability. If small position or energy changes occur, a compensation system must be developed to assure that the bunch is over the final focus system in the same trajectory and with the same energy. The final focus system must be designed so that the distant crossings of bunches do not disrupt the primary collisions at the interaction point.

Although the addition of many bunches appears to be "free" in that we simply use energy that would normally be wasted, it introduces complexity into every subsystem of the entire collider. The benefit is an order of magnitude increase in the luminosity.

4.3 The Beam Size

The transverse size of a beam in an accelerator is determined by two basic parameters: the emittance $\varepsilon$ and the beta function $\beta$.

$$\sigma = \sqrt{\varepsilon} \quad (3)$$

The emittance is a parameter that is proportional to the area occupied by the beam distribution in transverse phase space $(x, p_x)$. It is defined by

$$\varepsilon = \frac{1}{\mu_0} \langle x^2 \rangle \langle p_x^2 \rangle - \langle xp_x \rangle$$

where $x$ is the transverse position, $p_x$ is the corresponding transverse momentum, and $\mu_0$ is the central momentum of the bunch of particles. The angle brackets in Eq. 4 indicate an average over the distribution of particles in a bunch. Because the quantity in the square brackets is an additive invariant (in the absence of synchrotron radiation), the emittance decreases inversely with the momentum of the bunch in a linear accelerator.

The longitudinal emittance is defined in a similar way.

$$\varepsilon_z = \frac{1}{\mu_0} \left( < z^2 > - < z \Delta p >^2 \right). \quad (5)$$

where $z$ is the longitudinal deviation from a central position within the bunch, and $\Delta p$ is the deviation of the particle momentum from a central momentum. Once again, the quantity in the square brackets is an additive invariant, which causes $\varepsilon_z$ to increase inversely with the beam momentum in a linear accelerator. In the special case of a high-energy electron beam, the longitudinal distribution and the bunch length are fixed because the particles all travel at essentially the speed of light. In this case, the fractional momentum spread varies inversely with the beam momentum.

The beta function $\beta$ was first introduced by Compton and Snyder in their description of the alternating gradient focusing of particle beams. The parameter not only determines the particle beam size through Eq. 3, it also determines the instantaneous wavelength of the oscillations of particles within the beam envelope as they traverse the focusing magnets (wavelength $= 2\pi\beta$).

The beta function also plays an important role at the interaction point (IP). In a magnet-free region, it has the particularly simple form

$$\beta(x) = \beta + \frac{(x - s_0)^2}{p_x^2} \quad (6)$$

where $\beta^*$ is the minimum value of $\beta(x)$ and $s_0$ is the location of that minimum, the IP in this case. According to Eq. 3, the beam size near the IP is therefore

$$\sigma^2(x) = \sigma^2 + \frac{x}{2}(s - s_0)^2 \quad (7)$$

From this form, it is obvious that $\beta^*$ is the depth of focus because the beam size increases by $\sqrt{2}$ when $x - s_0 = \beta^*$. Thus, the beta function plays two important roles at the IP—it determines both the spot size and depth of focus. In order to achieve the sizes shown in Table 1, we must reduce the emittance as much as possible and preserve it during acceleration, and finally, we must focus the beam down to provide a small $\beta^*$ at the interaction point.

4.4 The Damping Ring

The damping ring serves to reduce the emittance of the bunches of particles in all three degrees of freedom. It is an electron storage ring similar in all essential
features to the storage rings used for colliding beams or synchrotron light production. The particles in an electron storage ring radiate a substantial fraction of their energy on each turn, owing to the interaction between the particle's magnetic moment and the magnetic field of the ring. The particles lose energy from all three degrees of freedom, but the energy is transferred only along one, the direction of motion, the proper amount is supplied at a single RF phase for a particle with the design energy, which leads to damping in all three dimensions. The fact that radiation is emitted as discrete quanta, however, introduces stochastic noise that causes diffusion of particle trajectories.

The competition between these damping and diffusion effects leads to an equilibrium value for the emittance of an electron storage ring. Damping rings are designed to enhance the damping effects using strong magnetic fields (such as those in wigglers magnets), while limiting the diffusion by the special design of the transverse focusing in the ring. In addition, there is a unique feature of electron storage rings that can be used to advantage. Due to the lack of vertical bending, the vertical emittance of the beam is much smaller than the horizontal, typically two orders of magnitude smaller. Such naturally flat beams are a key feature of many SLC designs.

One possible design for a future damping ring is about a factor of five larger and operates at an energy 50 percent higher than that of the SLC damping rings (see Fig. 13). The final emittance of the beam is more than an order of magnitude smaller than that of the SLC beams, which leads to much smaller sizes. In fact, the vertical emittance of a beam emerging from this damping ring would be less than one, or about equal to the final spot size of the SLC interaction point.

![Diagram of a damping ring with wigglers and focusing magnets](attachment:image.png)

2. Another key difference is the simultaneous damping of many bunches of beams. In the SLC, at most two bunches are damped simultaneously, whereas this SLC ring will damp 10 bunches all at once. This feature allows a longer damping time for any given bunch, because we can extract the "oldest" bunch and inject a new "young" bunch while leaving those in their "adolescence" to continue damping undisturbed.

Because the bunches are still bunches in the damping ring, their conditions upon emerging are entirely determined by their behavior in the damping ring. This process special emphasis on the stability of the magnets in the damping ring and extraction system.

3. Although the longitudinal emittance obtained in the damping ring is small enough, the bunch is still much too long for acceleration in a linac. In the SLC and SLC, this problem is solved by a technique called bunch compression, which shortens the bunch while increasing its energy spread. Each bunch passes through an RF accelerating structure placed so that the trailing particles emerge with lower energy than the leading particles. Then the bunch passes through a sequence of magnets that disperses the beam so that particles of different momentum travel on different paths. Particles with higher momentum (at the head of the bunch) travel a larger path than those of lower momentum (at the tail). The end of the bunch can therefore catch up with the head, producing a shorter bunch but at the cost of a greater energy spread.

This type of bunch compression has been used routinely in the SLC, where bunches 7 mm long are compressed to 0.5 mm for acceleration in the linac. Much shorter bunches will be required in the SLC. Short bunches will suffer from transverse wakefields in the linac, and they permit a smaller depth of focus at the IP (about 120 mm for the SLC). In principle, another order of magnitude in compression could be obtained in a single stage; in practice, however, this approach would lead to other deleterious effects due to the large energy spread that would be induced in the beam. For this reason, the extra compression is provided by a second bunch compression operating at a higher energy.

In the SLC, the bunch is first compressed as in the SLC to 0.5 mm in length, after which the beam is accelerated to about 16 GeV. The longitudinal spread of the beam is unchanged by this acceleration, but the relative energy spread decreases linearly with energy. The compression is then repeated, resulting in a bunch length less than 100 nanometers. By separating the compression process into two discrete steps, we can keep the relative energy spread small throughout.

4. Emittance Preservation During Acceleration

During the process of acceleration, we must take care not to dilute the emittance of the beam. There are several effects which can lead to emittance dilution in the next few subsections, we discuss a few of the most important effects.
4.6.1. Chromatic Effects

The filamentation of the central trajectory in a linac can cause dilution of the effective emittance of the beam. If we first consider a coherent betatron oscillation down the linac, then to be absolutely safe, we must require that it be small compared to the beam size. If the spread in betatron phase advance is not too large, then this tolerance is increased to perhaps twice the beam size for the design shown in Table 1.

The chromatic effect of a corrected trajectory is rather different. In this case, it is the distance between an error and a corrector which matters, and the effect can be compensated to yield negligible dilution of the beam emittance. For these techniques to work effectively, it is important that the single-bunch wakefields and beam loading be reasonably small.

4.6.2. Correction Techniques

As the trajectory is corrected the dispersion grows. It is possible, however, to measure this effect and choose a trajectory which is small and also has small dispersion. This type of compensation technique has been extensively investigated. If this correction is provided, then normal alignment tolerances (100-200 µm) can be compensated to yield negligible dilution of the beam emittance. For these techniques to work effectively, it is important that the single-bunch wakefields and beam loading be reasonably small.

4.6.3. Transverse Wakefields and DNS Damping

The wakefield left by the head of a bunch of particles, if it is in effect in the structure, deflects the tail. If the transverse oscillations of the head and tail have the same wave number, the tail is driven on resonance. This leads to growth of the tail of the bunch. This effect can be controlled by a technique called DNS damping. The bunch is given a head-to-tail energy correlation so that the tail is at lower energy. The offset of the head by a small amount induces a deflecting force on the tail away from the axis. This, effect can be controlled by a technique called DNS damping.

The bunch is given a head-to-tail energy correlation so that the tail is at lower energy. The offset of the head by an amount \( \Delta E \) induces a deflecting force on the tail away from the axis. This leads to growth of the tail of the bunch. The bunch is driven on resonance. This leads to growth of the tail of the bunch. If all the quadrupoles in the linac are vibrating in a random way, the effects accumulate down the linac and the orbit error grows \( \propto \sqrt{N_{\text{quad}}} \). This sets the tolerance on the random motion of quadrupoles to be much smaller than the beam size. In the examples in Table 1, the random jitter tolerances are \( \approx 0.01 \mu \text{m} \). However, this phase motion from pulse-to-pulse is unlikely due to the large repetition rate of the collider. More gradual motion, which is larger, can be corrected with feedback, provided the repetition frequency is sufficiently large.

Jitter in RF kicks can cause similar effects. These effects can be reduced by reducing the DC component of the RF kick by eliminating asymmetries in couplers and by careful alignment of structures.

4.6.4. Jitter

In order to maintain collisions at the interaction point, the bunch must not move very much from pulse to pulse. Since the optics of the final focus also demagnify this jitter, the tolerance is always set by the local beam divergence compared to the variation of some angular kick. The jitter tolerance on the damping ring kicker is thus related to the divergence of the beam at that point. At the injection point to the linac, the offset caused by this jitter must be small compared to the local beam size.

If all the quadrupoles in the linac are vibrating in a random way, the effects accumulate down the linac and the orbit error grows \( \propto \sqrt{N_{\text{quad}}} \). This sets the tolerance on the random motion of quadrupoles to be much smaller than the beam size. In the examples in Table 1, the random jitter tolerances are \( \approx 0.01 \mu \text{m} \). However, this phase motion from pulse-to-pulse is unlikely due to the large repetition rate of the collider. More gradual motion, which is larger, can be corrected with feedback, provided the repetition frequency is sufficiently large.

Jitter in RF kicks can cause similar effects. These effects can be reduced by reducing the DC component of the RF kick by eliminating asymmetries in couplers and by careful alignment of structures.

4.6.5. Coupling

Finally, we discuss coupling of the horizontal and vertical emittance. The beam size ratio in the linac is 10:1. The tolerance on random rotations for a flat beam is given by

\[
\Theta_{\text{rms}} \ll \frac{\sigma_T}{\sigma_x} \frac{1}{\sqrt{2N_{\text{quad}}}}.
\]

For the example shown in Table 1, the right-hand side is about 1 mrad; this is straightforward to achieve. If the errors are not random, larger rotations can indeed result; however, because the beam size is so small, the effects are very linear. This means that skew quadrupoles can be used effectively as correction elements. Certainly, in the final focus, skew quadrupoles will be an integral part of the tuning procedure to obtain flat beams.

4.7. Final Focus

At the end of each linear accelerator is a final focus system whose purpose is to compress the tiny bunches to submicron dimensions. To obtain the luminosity desired, the cross-sectional area of each bunch must be only a few hundred square nanometers. In addition, we must focus it to the shape of a flat ribbon (rather than a string) in order to minimize the radiation emitted as the particles in the bunch encounter the intense electromagnetic field of the opposing bunch. These
goals are accomplished by the use of a magnetic focusing system analogous (in reverse) to an optical telescope used to magnify distant objects. This system uses quadrupole magnets as focusing elements in a combination that provides a very large demagnification.

A major problem is the so called "chromatic" effect of the final quadrupole magnets. Two parallel electron beams with different momenta entering a perfect quadrupole magnet are brought to a focus at slightly different horizontal positions because the lower energy beam is bent slightly more than the higher energy beam by the magnetic field (see Fig 14). For it not to affect the spot size, this shift of focal point must be smaller than \( \frac{1}{f} \), the depth of focus of the beam. Due to the requirement of flat beams, this depth of focus is about 100 microns in the vertical dimension.

\[
\begin{align*}
\text{High Energy} & \quad \downarrow \quad \text{IP} \quad \uparrow \quad \text{Low Energy} \\
\text{Central Energy} & \quad \downarrow \\
\text{Magnetic Lens} & 
\end{align*}
\]

Fig. 14: The structure of the final quadrupole focuses is not flat, but bent in two different dimensions.

Such a small depth of focus makes the chromatic effects particularly serious. The chromatic correction of the final quadrupole is in fact the key to the final focus. Upstream of these quadrupoles is a combination of bending magnets that dispenses the beam combined with nonlinear sextupole magnets so that higher energy particles get a bit more focusing than lower energy particles. Where a bunch arrives at the last quadrupole the chromatic effect of the magnet is held upon it is exactly canceled.

The basic principle of the chromatic correction for particle beams have been known and utilized for about 30 years. Their first application was in the SLC where the beams are demagnified by about a factor of 30 yielding spot sizes of about 2 microns. However, the demagnification necessary in the NLC is about a factor of 100. However, the design of its final focus system will be substantially different from that of the SLC.

### 4.6 The FFTB Project

In order to test such a next generation final focus experimentally, an international collaboration including SLAC, INP, KEK, Orsay, DESY and NPI has been formed to design and construct a Final Focus Test Beam (FFTB) at SLAC. This facility will use the SLC beam emerging straight ahead from the linac as its source of electron bunches.

Figure 15 shows a schematic of the location and layout of the FFTB. It is a scaled version of an NLC final focus, and as such, is qualitatively similar to NLC designs. A special feature of the design is that the chromaticity correcting sextupoles are grouped in separate pairs, one for the horizontal dimension and one for the vertical. This pair of magnets is arranged so that the nonlinear aberrations introduced are cancelled, while the chromatic effects are added. The bends shown in Fig. 15 horizontally disperse the different momenta in the beam so that the sextupoles give somewhat more focusing to the higher energy particles. This additional focusing is arranged so as to cancel the lack of focusing of the higher energy particles in the final quadrupoles.

\[
\begin{align*}
\text{Quadrupoles} & \quad \downarrow \quad \text{Dipole Bends} \quad \downarrow \quad \text{Final Quads} \\
\text{IP} & 
\end{align*}
\]

Fig. 15: The location and schematic layout of the Final Focus Test Beam.

The goal of the FFTB is to produce bunches with transverse dimensions of 60 nanometers high by 1 micron wide. Figure 16 shows the vertical beam size plotted versus the vertical \( z \) at the IP in an ideal linear system as discussed in
Section 4.3, the beam size is just proportional to the square root of $x$. This is shown as the dotted line in Fig. 16. If the bunch has finite energy spread and with no correction, this linear decrease is modified by chromatic aberrations so that the beam size reaches a minimum of about 1 mm (the solid line in Fig. 16). Finally, the chromatic-correction sextupoles are powered and the system is properly tuned and adjusted, the vertical beam size follows the linear optics down to a size of about 60 μm before other high-order effects spoil the compensation (the dashed line in Fig. 16).

The FFTB will not achieve the beam size necessary for NLC due to the lack of a suitable low emittance source. In fact, to achieve such low emittance, we need the NLC damping ring and bump. The FFTB will, however, test the optical alignment necessary for NLC. In fact, the design for FFTB is identical to that for NLC. In addition to the primary goal, the collaboration will use this facility to test the alignment, stability, and instrumentation requirements needed to achieve such small spots. The FFTB project is proceeding on schedule. Most of the magnets have been tested and measured and initial tests of the next part of the beam line have just been completed. The initial tests of the FFTB should begin in 1984.

4.9. Beam-Beam Effects

When two oppositely-charged bunches collide at the IP, the intense electromagnetic field generated by the bunches tends to mutually focus them. This leads to disruption of the bunch and to a pinch enhancement of the luminosity. The enhancement factor $H_B$ was given in Eq. 1 for the luminosity. For round bunches, this enhancement can be quite large ($> 5$), for flat bunches, however, it is considerably reduced ($> 2$) because the pinch only occurs in one dimension. If, in addition, the bunches are misaligned relative to each other, the centroids are attracted during the bunch passage. This leads to a two-stream instability which, in moderate disruption, actually helps the collision process; if the bunches are misaligned, they bend toward each other and collide partially anyway.

The combination of very high electromagnetic fields and high particle energy yields substantial amounts of synchrotron radiation known as beamstrahlung. The average energy loss due to beamstrahlung ranges from 1 to 20 percent in various NLC designs. In extreme cases, many of these photons can subsequently generate electron-positron pairs in the intense electromagnetic fields present. The radiated photons or charged particles can strike detector components, causing undesirable backgrounds.

The train of bunches on each cycle also presents a problem at the final focus. In order to have a separate channel for the outgoing disrupted bunch, collisions take place at a small angle. As the bunches approach the collision point, they feel the field from those bunches which are exiting and have already collided. This sequence of bunches can induce a multibunch kink instability which can cause tails in bunches.
estimates suggest that although this may not be serious at 0.5 TeV, it will definitely be a problem at 1 TeV. This effect will push designs towards more bunches per cycle with less charge per bunch. However, there are tradeoffs between detector design and linear collider design which have yet to be explored to help solve this problem.

5. Outlook

There has been a tremendous amount of progress in the accelerator physics and technology of linear colliders during the past five years. This has come from a combination of experience with the SLC, and recent research and development around the world. The NLC design at SLAC has evolved towards more conservative parameters, taking into account the real experience with the SLC. We are very close to having the key components of the technology under our belt, an NLC-style final focus with the FFTB and a model of the high gradient fame, the NLC Test Accelerator. The ATP at KEK will provide a test of the next generation of damping rings. The SLC will continue to provide us with useful information about a real working linear collider. In addition, the SLC provides a test bed to perform scaled experiments to test our understanding of the accelerator physics necessary to obtain the luminosity.

A key question is, can we accelerate flat, low-emittance beams while preserving their emittance? We believe the answer is yes, but experiments on the SLC which probe our understanding of this question are essential. The final boost of a factor of 10 in luminosity is obtained by accelerating 10 bunches on each machine cycle. The choice of multiple bunches impacts every system of the collider. Although we believe we have solved all the key problems associated with the multiplicity issue, there is a significant risk remaining with this approach. Because of this risk, we feel it is important to be able to achieve the design luminosity with only very few bunches. The understanding of beam-beam effects and how they affect backgrounds has improved; however, much more work needs to be done on the interaction of backgrounds, detector design and collider design.

To conclude, during the next few years we look forward to the resolution of both the technical and accelerator physics issues important to the design of the Next Linear Collider.

References


5. R. Bane and R. Gluckstern, to be published.


19. V.N. Balakin, V.M. Katkov and V.V. Smirnov, Proc. of 14th International Conf. on High Energy Particles Accelerators, Tsukuba, Japan (1991), and in SLAC-PUB-5354.


DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.