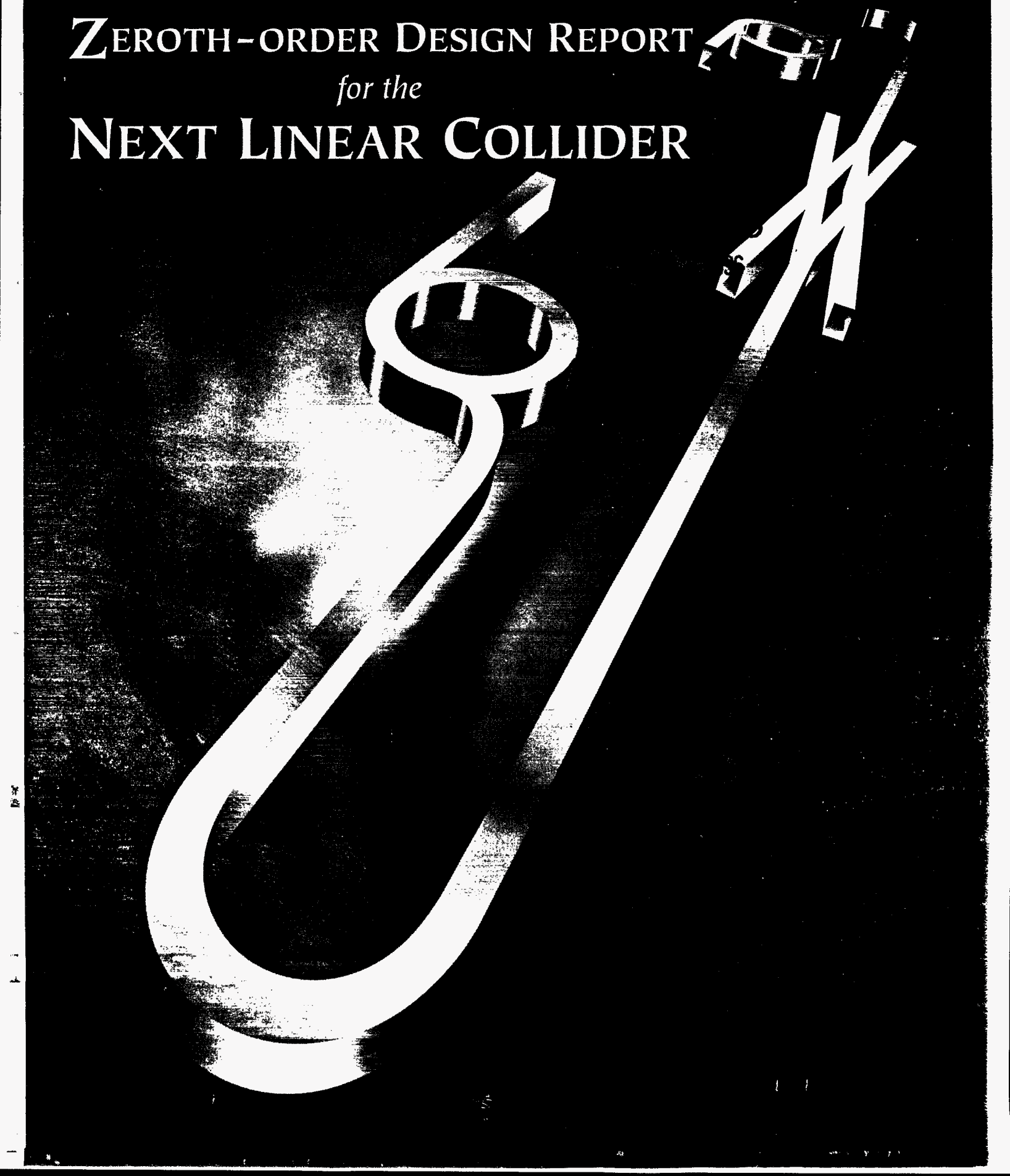
LBNL-PUB-5424
SLAC Report 474
UCRL-ID-124161

ZEROTH-ORDER DESIGN REPORT *for the* NEXT LINEAR COLLIDER



PUB
~~LBNL~~-5424
SLAC-474
UCRL-ID-124161
UC-414

Zeroth-Order Design Report for the Next Linear Collider

Volume II

The NLC Design Group

May, 1996

Prepared for the Department of Energy under contract number DE-AC03-76SF00515 by Stanford Linear Accelerator Center, Stanford University, Stanford, California. Printed in the United States of America. Available from National Technical Information Service, US Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

This document is available on the World Wide Web at http://www.slac.stanford.edu/.

The NLC Design Group

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

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

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

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

L. Rinolfi; *CERN, Geneva, Switzerland*

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

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

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

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

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

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

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

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

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

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

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

[1] Editor-in-chief.

Acknowledgments

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

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

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

Contents

Preface xi

Volume I 1

1 Introduction 1
- 1.1 Goals for the Next Linear Collider . . . 1
- 1.2 Accelerator Design Choices . . . 3
- 1.3 The Next Linear Collider . . . 8
- 1.4 Outlook for the Next Linear Collider . . . 15
- 1.5 NLC Systems Overview . . . 15

2 Electron Source 25
- 2.1 Introduction . . . 27
- 2.2 Polarized Electron Gun . . . 30
- 2.3 NLC Laser System . . . 37
- 2.4 NLC Electron Injector Beam Dynamics . . . 41
- 2.5 Buncher Cavities . . . 51
- 2.6 Positron Drive Linac . . . 51
- 2.7 Diagnostics . . . 53
- 2.8 Operation . . . 56
- 2.9 Conclusion . . . 57
- 2.A Polarized e^- Beam Photocathode RF Gun Development for the NLC . . . 58
- 2.B Charge Limit and its Implications on High-Polarization Long-Pulse Charge ... 74

3 Positron Source 83
- 3.1 Introduction . . . 85
- 3.2 Positron Yield Simulation . . . 90
- 3.3 Drive Electron Accelerator . . . 92
- 3.4 Positron Production Target . . . 92
- 3.5 Positron Collection System . . . 97
- 3.6 Beam Dynamics and Transport . . . 99
- 3.7 Positron Linac . . . 101
- 3.8 Radiation Control Issues . . . 103
- 3.9 Magnets . . . 105
- 3.10 Diagnostics and Instrumentation . . . 106
- 3.11 Feedback and Stability . . . 108
- 3.12 Operations and Tuning Procedures . . . 109
- 3.13 Control System Needs . . . 110
- 3.14 Other Considerations . . . 110
- 3.15 Summary . . . 111

4 NLC Damping Rings 115
4.1 Introduction . . . 117
4.2 System Overview and Parameter Determination . . . 118
4.3 Lattice Design . . . 133
4.4 Collective Limitations . . . 166
4.5 RF Systems . . . 195
4.6 Vacuum Systems . . . 211
4.7 Feedback and Feedforward Systems . . . 217
4.8 Vibration and Stability . . . 223
4.9 Alignment and Supports . . . 226
4.10 Magnet Design and Power Supplies . . . 229
4.11 Instrumentation for the Main Damping Rings . . . 246
4.12 Conclusions and Comments . . . 249

5 Bunch Compressors and Prelinac 257
5.1 Introduction . . . 258
5.2 Requirements and Design Options . . . 260
5.3 System Design . . . 265
5.4 Longitudinal Dynamics . . . 281
5.5 Transverse Dynamics . . . 298
5.6 Tolerances . . . 310
5.7 Conclusions and Comments . . . 314

6 Low-Frequency Linacs and Compressors 319
6.1 Introduction . . . 320
6.2 Beam-Loading Compensation . . . 320
6.3 Numerical simulation of beam-loading compensation . . . 324
6.4 Klystrons . . . 333
6.5 Dipole Wakefields . . . 335

7 Main Linacs: Design and Dynamics 343
7.1 Introduction . . . 345
7.2 Parameters and Specifications . . . 346
7.3 System Design . . . 349
7.4 Emittance Dynamics . . . 366
7.5 Alignment, Stability, and Feedback . . . 392
7.6 Simulation Studies . . . 406
7.7 Tolerances . . . 429
7.8 Operations and Machine Protection System . . . 431
7.9 Linac Components . . . 432
7.10 Instrumentation . . . 440
7.11 Discussion . . . 449

8 RF System for the Main Linacs 455
8.1 Introduction . . . 457
8.2 Accelerator Structure . . . 466
8.3 RF Pulse Compression and Power Transmission . . . 493
8.4 High-Power Klystrons . . . 502
8.5 Klystron Pulse Modulator . . . 517
8.6 RF Drive and Phasing Systems . . . 530
8.7 RF Protection and Monitoring Systems . . . 546

Volume II 555

9 Collimation Systems 555
9.1 Introduction to Beam Delivery Systems . . . 556
9.2 Post-Linac Collimation . . . 560
9.3 Pre-Linac Collimation . . . 631
9.4 Bunch Length Collimation . . . 634

10 IP switch and big bend 643
10.1 Introduction . . . 644
10.2 The IP Switch . . . 644
10.3 The Big Bend . . . 648

11 Final Focus 659
11.1 Introduction . . . 661
11.2 Parameters and Specifications . . . 662
11.3 Skew Correction and Diagnostic Section . . . 665
11.4 Beta-Matching Section . . . 669
11.5 Chromatic Correction and Final Transformer . . . 676
11.6 The Final Doublet . . . 743
11.7 Crossing Angle, Crab Cavity, and Solenoid . . . 767
11.8 The Beam Extraction and Diagnostic System (The Dump Line) . . . 774
11.9 Conclusions and Comments . . . 800
11.A The Beam Dumps . . . 802

12 The Interaction Region 815
12.1 Introduction . . . 816
12.2 The Luminosity Spectrum . . . 817
12.3 Detector Background Sources . . . 823
12.4 Detector Issues . . . 836
12.5 Conclusions . . . 851

13 Multiple Bunch Issues 855
13.1 Introduction . . . 856
13.2 Major Impacts of Multibunching . . . 857
13.3 Machine Protection and Operations . . . 864
13.4 Instrumentation Specifications . . . 865
13.5 Experimental tests related to multibunch issues . . . 865
13.6 Summary and Conclusions . . . 865

14 Control System 869
14.1 Introduction . . . 870
14.2 NLC Requirements . . . 870
14.3 Architectural implications . . . 875
14.4 The Control System Model . . . 876

15 Instrumentation 887

16 Machine Protection Systems 889
16.1 Introduction . . . 890
16.2 Single Pulse Induced Failure . . . 890

17 NLC Reliability Considerations 903
17.1 Goals . . . 904
17.2 Reliability and Availability . . . 904
17.3 Target NLC Availability . . . 905
17.4 NLC Machine Availability and System Reliability . . . 906
17.5 A Formal Solution . . . 908
17.6 Three Examples: Klystrons, Power Supplies, and Motors . . . 911
17.7 Summary . . . 911

18 NLC Conventional Facilities 915
18.1 Introduction . . . 916
18.2 Site . . . 916
18.3 Campus . . . 917
18.4 Injectors . . . 917
18.5 Linac . . . 917
18.6 Detectors . . . 918

A An RF Power Source Upgrade to the NLC Based on the Relativistic-Klystron ... 925
A.1 Introduction . . . 926
A.2 A Design for an RK-TBA-Based rf Power Source . . . 927
A.3 TBNLC Physics Studies . . . 937
A.4 TBNLC Engineering Design . . . 947
A.5 RTA Test Facility . . . 956
A.6 Conclusions . . . 964

B A Second Interaction Region For Gamma-Gamma, Gamma-Electron and ... 971
B.1 Introduction . . . 973
B.2 Physics Opportunities at $\gamma\gamma$ Collider: The Higgs Sector and Other New Physics . . . 974
B.3 Major Parameters . . . 978
B.4 CP Issues . . . 980
B.5 IP Issues . . . 987
B.6 Luminosity Calculations . . . 989
B.7 Backgrounds and Other Detector Considerations . . . 999
B.8 Laser Optical Path in IR . . . 1006
B.9 Gamma-Gamma Final Focus System . . . 1015
B.10 Extraction and Diagnostic Line . . . 1019
B.11 Laser Technology I: Solid State Lasers . . . 1020
B.12 Free-Electron Lasers . . . 1030

C Ground Motion: Theory and Measurement 1043
C.1 Introduction . . . 1044
C.2 Theory . . . 1044
C.3 Ground Motion Measurement Devices . . . 1057
C.4 SLAC Ground Motion Measurements and Analysis . . . 1063

D Beam-based Feedback: Theory and Implementation 1075
D.1 Introduction . . . 1076
D.2 Planned NLC Feedbacks . . . 1076
D.3 Feedback System Design . . . 1077
D.4 Performance Questions . . . 1077

D.5 Adaptive Cascade 1078
D.6 Rate Considerations and Corrector Speeds 1078
D.7 Calibrations and Modeling 1079
D.8 Global Performance Characterization 1079
D.9 Summary 1080

Preface

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

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

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

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

9

Collimation Systems

Contents

9.1 Introduction to Beam Delivery Systems 556
9.1.1 Introduction to Collimation Systems 557
9.2 Post-Linac Collimation 560
9.2.1 Specifications 560
9.2.2 Materials Considerations 569
9.2.3 Tail Re-Population Estimates 578
9.2.4 Wakefield Considerations 587
9.2.5 Lattice Description and Analysis 602
9.2.6 Operational Issues 625
9.2.7 Energy Scaling Laws 629
9.2.8 Nonlinear Collimation Systems 630
9.2.9 Summary and Conclusions 630
9.3 Pre-Linac Collimation 631
9.3.1 Pre-Linac Collimation Function: 631
9.3.2 Pre-Linac Collimation System Requirements: 632
9.3.3 Pre-Linac Collimation Optical Design: 632
9.3.4 Pre-Linac Collimation System Bandwidth 633
9.3.5 Summary 633
9.4 Bunch Length Collimation 634
9.4.1 Introduction 634
9.4.2 Collimator location 635
9.4.3 Particles Loss at Bunch Length Collimators 635
9.4.4 Collimation in the Linac 636
9.4.5 Conclusion and Discussion 637

9.1 Introduction to Beam Delivery Systems

Chapters 9 through 11 describe what are referred to as the beam delivery and removal systems. A schematic of these systems is shown in Figure 9-1. These systems begin at the end of the linac and terminate at a post-IP beam dump. They include a main post-linac collimation system (Section 9.2); an IP switch and big bend (Chapter 10); a pre-final-focus diagnostic and skew-correction section (Section 11.3); a final-focus system (Sections 11.4 to 11.7); and a post-IP beam line (Section 11.8). Chapter 9 also includes the description of the pre-linac horizontal and vertical collimation system (Section 9.3), and a pre-linac bunch-length collimation system (Section 9.4). The design of all of these systems has been strongly influenced by experience with similar systems at the SLC.

Figure 9-2 shows the horizontal and vertical beam envelope from the linac to the IP for the 1-TeV-c.m. beam line, and Figure 9-3 shows the corresponding β functions. The total length is 5.2 km. The first 50 meters contain a post-linac diagnostic chicane, which is then followed by a 2.5-km collimation system. The maximum horizontal points of the envelope in the collimation region, for the most part, correspond to the location of horizontal collimators and chromatic correction sextupoles; the maximum points of the vertical envelope are usually the location of vertical collimators and chromatic correction sextupoles. The next horizontal envelope peak, at 2.6 km, marks the location of the IP switch. After that, up to the 3.0-km marker, is a small envelope region that contains the big (10-mr) bend, and following this, up to the 3.4-km marker, is a small envelope region containing the pre-final-focus diagnostic region. In this region beam sizes in all phases and planes can be measured and the presence of coupling detected and corrected. The region from 3.4 km to 3.8 km contains the beta match into the final focus. Following the beta match, the first two peaks in the horizontal envelope are the positions of the horizontal chromaticity compensation sextupoles. The two large vertical envelope points following these, at about 4.4 km and 4.7 km, are the positions of the vertical chromatic compensation sextupoles. The last peak, in each plane, at the end of the beam line, is located at the position of the final-doublet elements.

Figure 9-4 shows the horizontal dispersion function for the same beam line. The peaks at the very beginning of the beam line are in a post-linac diagnostic region. These regions are similar to others along the linac and are described in Section 7.9.5. Following this and continuing to 2.4 km, one sees the dispersion function of the collimation system. The peak at 2.6 km is in the IP switch, and the small dispersion wiggles which follow are located in the big bend. The small bump at 3.4 km, at the beginning of the beta-match region in the final-focus system, is a short region containing two bends which provide an adjustable entry angle into the final-focus region. The net bend here will have three distinct values corresponding to three distinct final-focus systems that are required to cover the energy range from 350-GeV to 1.5-TeV c.m. Each has slightly different internal bend angles, but all reside on the same support bench and have the same IP location. The layouts are shown in Figure 11-9. The remaining dispersion is in the final-focus system.

We have studied all beam delivery systems for center-of-mass energies from 350 GeV to 1.5 TeV for a broad range of assumptions on beam and IP parameters, and have shown that it is possible to meet the specifications for this entire range. Above 1.0-TeV c.m. the elements of the collimation system also must be relocated. The tunnel length allotted to collimation is adequate to collimate energy and the horizontal and vertical planes at both phases only one time at 1.5-TeV c.m. Since the centroid orbit of the collimation system differs only slightly from a straight line, it would be possible to allocate length at the end of the linac tunnel for collimation at energies above 1-TeV c.m. And it should be sufficient to collimate the FD phase only one time in the collimation system. See Section 9.2.3 for details.

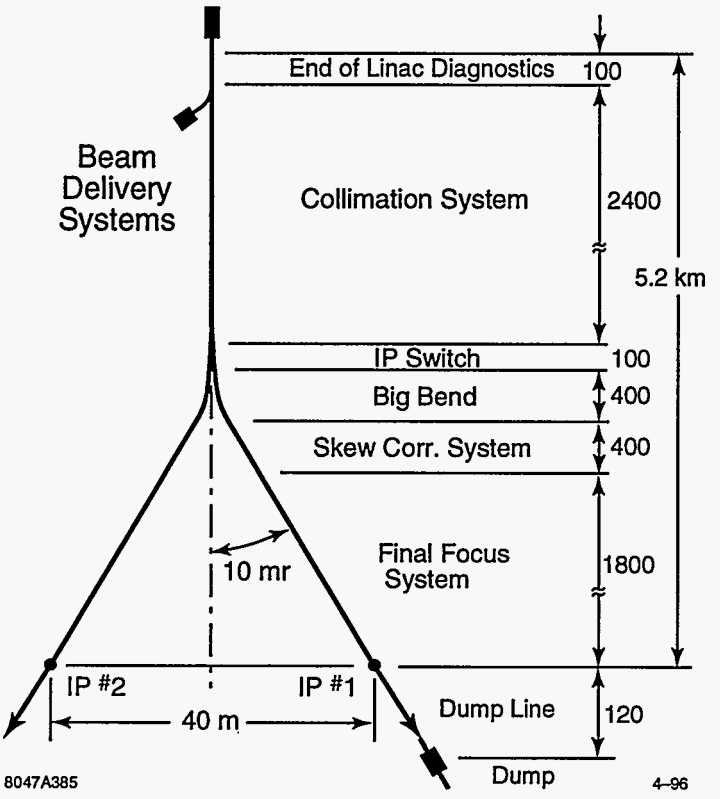


Figure 9-1. *A schematic layout of the beam delivery systems.*

9.1.1 Introduction to Collimation Systems

The post-linac collimation section (Section 9.2) begins with a discussion of system specifications since the design is strongly influenced by assumptions on incoming beam conditions: Of the 10^{12} particles per bunch train, how many particles must be routinely collimated, and how many particles can be safely collimated in the final-focus system? Because the number of particles needing to be collimated is difficult to predict precisely, since it depends on how well the upstream systems have been tuned, we have relied on SLC experience for this estimate. The number of particles that can be collimated in the final-focus system has been determined by edge-scattering and muon transport studies within the final-focus system. These studies are described in Chapter 12. The rough guideline that evolves from these considerations is that there may well be a few times 10^{10} particles in the beam tails at the end of the linac and this number needs to be reduced to a few times 10^6 upon entry into the final-focus system.

Two types of collimation systems have been proposed: linear and nonlinear. In the former, the beam sizes at the collimator are achieved by traditional linear optics methods (quadrupoles); in the latter strong sextupoles are used to blow up the beam. Since the sextupoles are exceedingly strong and system lengths are not reduced in the specific system proposals we have studied, we have opted to look in depth at a linear collimation scheme. It is not precluded that a nonlinear (or combination linear and nonlinear) system could be found that would be operationally superior and have a lower total cost. Our primary objective is to show that at least one collimation system exists that fulfills all functional requirements.

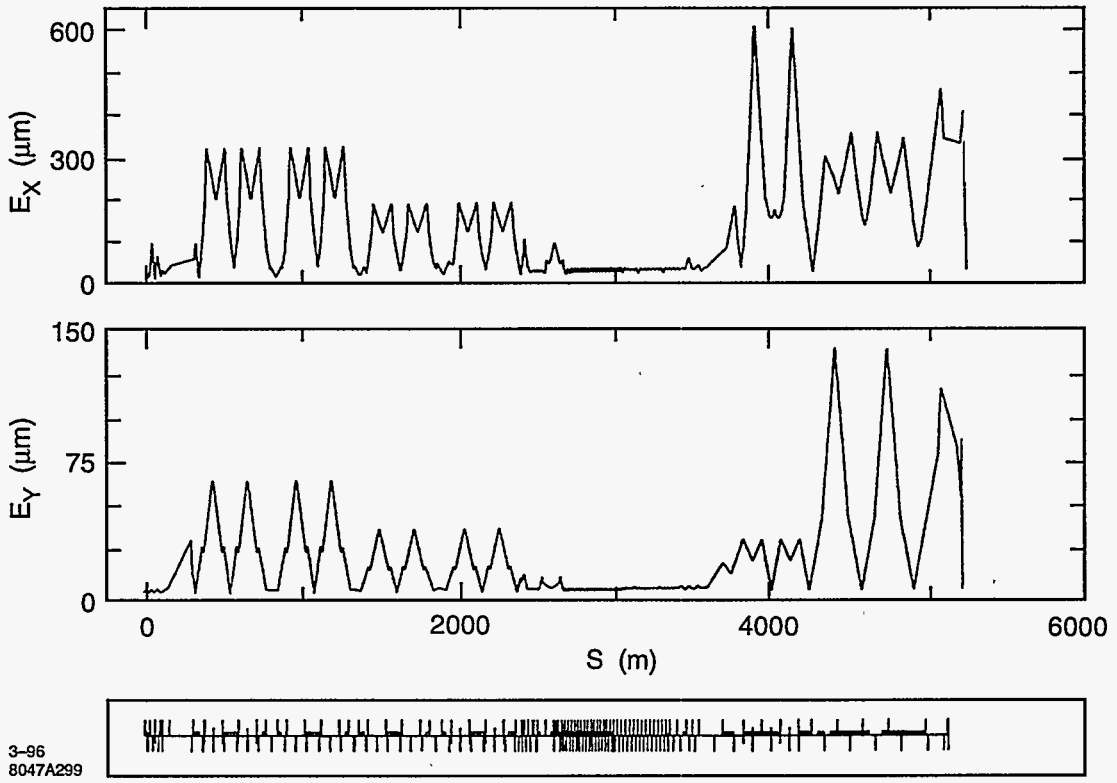


Figure 9-2. *The horizontal and vertical beam envelope from the linac to the IP for the 1-TeV-c.m. beam line.*

Since a small perturbation in upstream conditions could cause a complete bunch train of 10^{12} particles to be incident on the collimators of the collimation system, it is necessary in both the linear and nonlinear systems to rely on a primary collimator that is a spoiler, followed by a secondary collimator which is the absorber. The spoilers must be thermally rugged and very thin ($\leq 1/4$ radiation length). The best material we have found for spoilers is a titanium alloy plated with pure titanium or titanium nitride (TiN) for improved electrical conductivity. The function of the spoilers is to increase the angular divergence of the beam, so that when the beam arrives at the absorber it has a much larger size (millimeters). The absorber, on the other hand, must be able to routinely absorb and remove the energy in the tail of the beam. For the 1-TeV c.m. parameters, 1% of the time-averaged beam power is 84kW. The preferred material for absorbers is copper.

The wakefields of the collimators can have a very deleterious effect on the beam core. To minimize the wakes, the beam pipe must be tapered before and after the collimator. Even for an on-axis beam core, a parallel-jaw collimator will have a quadrupole wake, which can influence focusing of the core and the trajectory of particles in the tails. And beams that have been mis-steered close to the wall can experience very large wake-induced kicks. All of these wake effects are described in Section 9.2.4.

We consider the geometric wakefields for tapered collimators to be uncertain. There are theoretical results which we will discuss below, but the small tapers and short bunch lengths have made these collimators very hard to simulate with existing numerical modeling codes. The geometric wake for the parallel-plate geometry has changed as this document was going to press. We have incorporated notes in the text describing the change. If the new result stands, because the geometric-part of the wake is now larger than the cylindrical

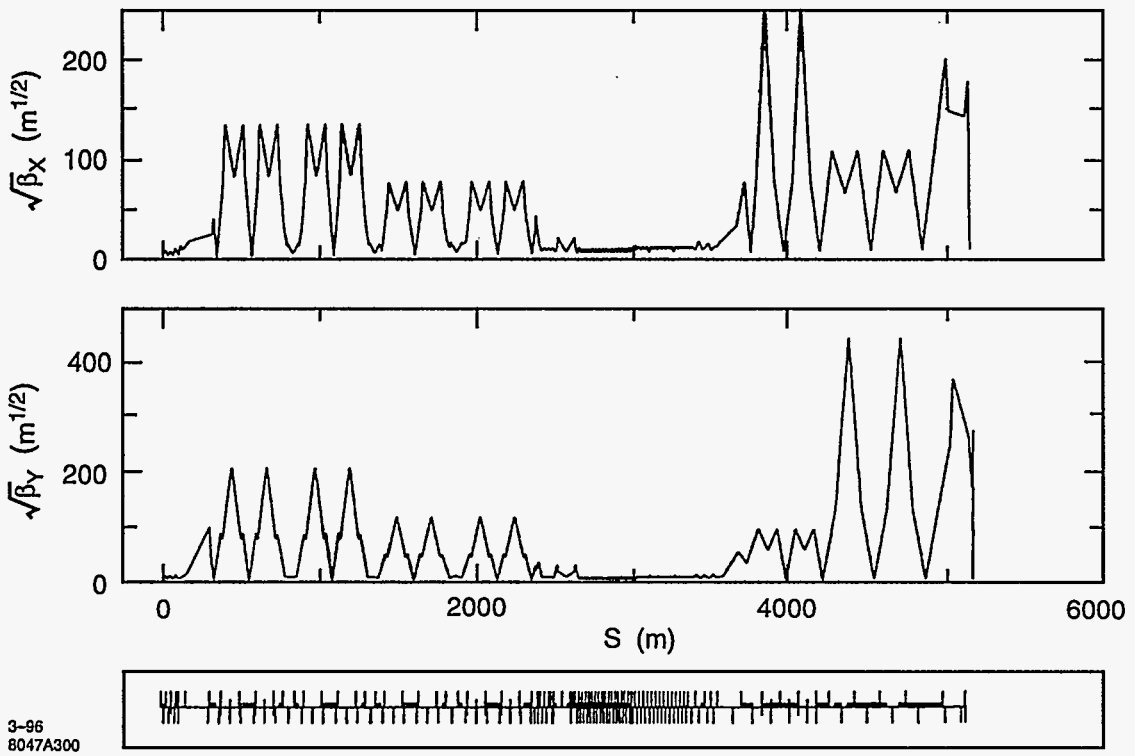


Figure 9-3. *The horizontal and vertical β function from the linac to the IP for the 1-TeV-c.m. beam line.*

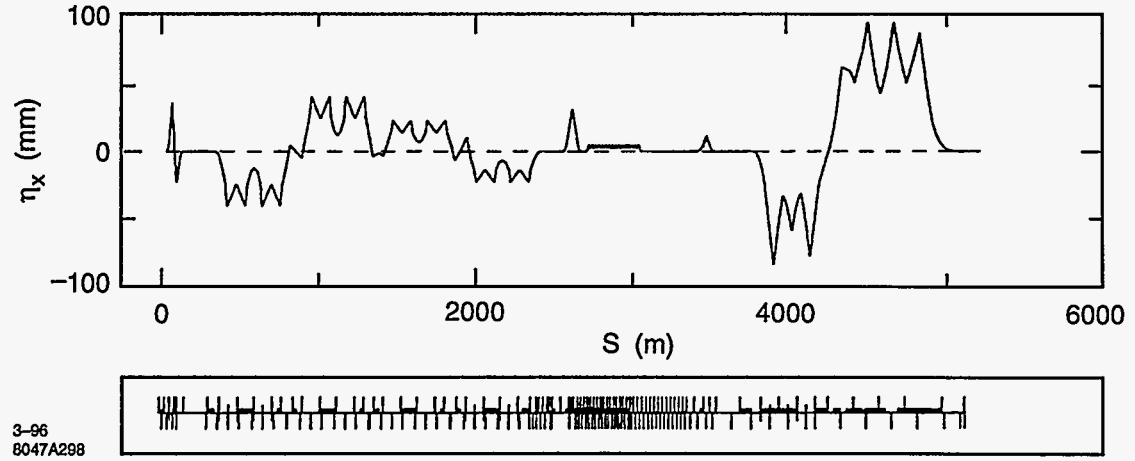


Figure 9-4. *The horizontal dispersion function from the linac to the IP for the 1-TeV-c.m. beam line.*

wake by a factor of about 4, there would be an impetus to use the less-convenient cylindrical geometry collimators.

There has also been some doubt cast on the correctness of the resistive-wall wake formulae we have used for the tapered collimators. A recent measurement of collimator wakes at the SLC has given a wakefield kick that is stronger than expected. The situation indicates a need for clarifying experiments with tapered collimators.

Because of the large β-functions and strong focusing that arise when the beam is blown up with linear optics, there are important chromatic effects to compensate with sextupole pairs. And because it is necessary to collimate each transverse phase at least one time, there are very large R_{12} and R_{34} functions within the system. Large R_{12} and R_{34} functions lead to important tolerances described in Section 9.2.5. Stability tolerances, that must be held between tunings of the waist knobs in the skew correction system, are looser than those within the final-focus system. However the particle backgrounds in the collimation system preclude the beam-based stabilization methods contemplated for the final-focus system. Alternatives are discussed in Section 9.2.6.

Vibration tolerances that must be met between orbit adjustments are greatly simplified by the ground motion studies described in Appendix C which, when applied to the collimation system, show that seismic ground motion has a negligible impact on beam collision offsets at the IP. As a result, vibration tolerances become tolerances between beam-line elements and the ground beneath them, or tolerances on ground motion coming from cultural sources.

9.2 Post-Linac Collimation

9.2.1 Specifications

Incoming Beam Parameter Range

The important parameters from the point of view of the collimation system are:

i) $Nn_b/\sqrt{(\varepsilon_x^N \varepsilon_y^N)}$, total bunch-train charge divided by the square root of the normalized emittance product, for single pulse-train spoiler survival;

ii) $N/(\varepsilon_x^N \sigma_z^{2/3})$ and $N/(\varepsilon_y^N \sigma_z^{2/3})$, for determination of minimum n_x^2 and n_y^2 that can be collimated in each plane (see Eq. 9.45); and

iii) the worst-condition population in tails, for average power loads on absorbers.

The strange units result from omission of a dependence on the surface resistivity of the collimators.

The range of IP parameters under consideration is given in Chapter 12, Table 12-1. Here one sees a maximum charge per bunch of $N = 1.25\ 10^{10}$, a maximum charge per bunch-train of $Nn_b = 1.125\ 10^{12}$, a minimum bunch length of $100\,\mu$m, a minimum horizontal normalized emittance of $\varepsilon_x^N = 4\ 10^{-6}$ m-rad, and vertical normalized emittances which vary from $\varepsilon_y^N = 8$ to $17\ 10^{-8}$ m-rad. The emittances at the end of the linac can be somewhat smaller and, depending on linac alignment conditions, could be as small as $\varepsilon_x^N = 4\ 10^{-6}$ m-rad and $\varepsilon_y^N = 5\ 10^{-8}$ m-rad.

Limit ii) implies that $n_x\sigma_x$ and $n_y\sigma_y$ are independent of energy and emittance. Thus if the downstream beta functions do not change, the physical apertures that may be collimated are independent of energy and incoming beam emittance.

The most sensitive parameter is that of item i), because the required value of the β functions at the spoilers depends on this ratio, and the system length increases as this β function becomes larger.

For evaluating worst conditions we assume the values $\varepsilon_x^N = 4\ 10^{-6}$ m-rad, $\varepsilon_y^N = 5\ 10^{-8}$ m-rad, and $N = 1.25\ 10^{10}$. The worst ratio $N/\sigma_z^{2/3}$, in the parameter range of Table 12-1, occurs when $N = 1.25\ 10^{10}$ and $\sigma_z = 150\,\mu$m. We also assume that a maximum of 1% of the beam may need to be collimated in any phase and any plane, or in energy. The total beam fraction collimated could be as high as 5%. This assumption is discussed in the next section.

Incoming Beam Halo Specification

Experience with the SLC has identified several sources of halo particles: those created

i) during extraction from the damping ring,

ii) in the bunch compressor,

iii) by wakefields within the linac,

iv) by mismatches, misalignment and steering errors in the linac,

v) from acceleration of dark current,

vi) from hard Coulomb and/or bremsstrahlung scattering within the linac, and

vii) from faulty multi-bunch energy compensation.

The first two items may be addressed by the pre-linac collimation system (Section 9.3). A bunch-length collimation system (Section 9.4) can reduce contributions from item iii). Linac diagnostics are intended to reduce iv). Since the injection energy into the linac is at 10 GeV, dark current will be considerably off-energy and presumably not accelerated far.

Item vi) cannot be eliminated, but the number of particles in the tail from this source can be derived [Yokoya 1991] beginning with the scattering angle

$$\Delta\theta = \frac{2Zr_e}{\gamma b} \tag{9.1}$$

of an electron passing at distance b from a nucleus with charge Z. For an azimuthal orientation of an angle ϕ specifying initial conditions, the kick in the horizontal direction would be $\Delta\theta_x = \Delta\theta\cos\phi$. For $\Delta\theta_x \geq n_x\sigma_x'$ we must have $b \leq b_n$ with

$$b_n(\phi) = \frac{2Zr_e}{\gamma n_x \sigma_{x'}}\cos\phi \quad . \tag{9.2}$$

Hence the cross-section for scattering beyond $n_x\sigma_x'$ is

$$\sigma_n = \int_0^{2\pi} d\phi \int_0^{b_n(\phi)} bdb = \pi b_n^2(0) \quad . \tag{9.3}$$

If ρ is the density of nuclei, then the fraction of particles scattered beyond $n_x\sigma_{x'}$ in length ds is given by

$$\frac{\Delta N}{N} = \pi\rho \int b_n^2(0)ds = \frac{2\pi r_e^2 \rho Z^2}{n_x^2 \varepsilon_x^N} \int \frac{\beta_x(s)ds}{\gamma(s)} \tag{9.4}$$

This can be integrated over the length of the linac. The β function oscillates between minimum and maximum values that scale roughly as the square root of energy (See Chapter 7, Figure 7-1 and 7-2. This oscillation is much more rapid than the change of γ so it may be replaced by an appropriate average, $\xi\beta_x^{\max}(s) = \xi\beta_x^{\max}(0)\sqrt{[\gamma(s)/\gamma(0)]}$. We may take $d\gamma/ds$ to be a constant and carry out an integral over γ to obtain

$$\frac{\Delta N}{N} = \frac{4\pi r_e^2 \sum_i \rho_i Z_i^2}{n_x^2 \varepsilon_x^N} \frac{\xi\beta_x^{\max}(0)L}{\sqrt{\gamma(L)\gamma(0)}} \quad . \tag{9.5}$$

We have included a sum over nuclear species, ρ_i being the density of nuclear species with charge Z_i. If the partial pressure associated with a nuclear species is P_{Gi} then the number of nuclei per unit volume is given by $\rho_i \approx 3\ 10^{22} n_i P_{Gi}\ m^{-3}$ where n_i is the number of nucleons per atom. If we take the major composition of gas within the beam pipe to be N_2 with nuclear charge $Z = 7$, and take $P_G = 10^{-8}$ Torr, $\gamma(0) = 2\ 10^4$, $d\gamma/ds = 10^2$, and $\beta_x^{\max}(0) = 10$ m, we obtain the estimate

$$\frac{\Delta N}{N} \approx \frac{1.4\ 10^{-7}}{n_x^2} \quad \text{or, in a similar manner} \quad \frac{\Delta N}{N} \approx \frac{1.4\ 10^{-5}}{n_x^2} \quad . \tag{9.6}$$

For $N = 10^{12}$, $\Delta N \approx 1.4\ 10^5/n_x^2$ or $\Delta N \approx 1.4\ 10^7/n_y^2$. For collimation at $n_x = 5$ or $n_y = 35$, $\Delta N < 10^4$. We will see below that this number of particles could be safely collimated in the final-focus system.

Bremsstrahlung interactions with beam gas can result in particle energy loss and create tails. The cross section for this process is given in Eq. 11.118. For more than 1% energy loss the cross section is estimated to be about $\sigma_{brems} \approx 6$ barn. This gives the estimate

$$\frac{\Delta N}{N} \approx 2\ 10^{-5} P_G L \tag{9.7}$$

where P_G is the gas pressure in Torr and L is the linac length in meters. For $N = 10^{12}$, $L = 10^4$, and $P_G = 10^{-8}$, we have $\Delta N \approx 2\ 10^3$, which is a smaller number than from Coulomb scattering.

Hence the principal source of concern is item iii), the tail particles generated by wakes in the linac. Compared to the SLC, the bunch-length collimation and pre-linac collimation should help reduce this number, but the number of particles in the longitudinal tails under typical (and worst!) operating conditions remains uncertain. Estimates indicate that tail populations should be less than 10^{-3} or even 10^{-4} [Raubenheimer 1995]. On the other hand, at the SLC it has been necessary on occasion to collimate 10% of the beam particles. When the systems are tuned this number is an order of magnitude smaller.

Assuming that the pre-linac collimation systems and other efforts to minimize tail particles will have a positive impact, we somewhat arbitrarily have adopted the specification that at most 1% of the beam will require collimation at any particular phase, transverse plane, or energy. The system to be described will actually have somewhat higher totals, up to 5%, depending on how beam tails are proportioned. The main impact of choosing a larger collimation percentage is in the size of the absorbers, and they could be designed to remove a higher (or lower) average power load if it was deemed necessary.

Required Collimation Depths

Final-doublet (FD) Phase Depth. The collimation depth at the final-doublet (FD) phase, $\pi/2$ from the interaction-point (IP) phase, is determined by beam emittances and the apertures and β functions in

the final-doublet quadrupoles. The final-doublet β functions are determined by the IP β functions and the final-doublet design parameters: the free space from the last quadrupole to the IP and the strength of the quadrupoles.

Final-doublet designs get more compact and have smaller chromaticity for larger quadrupole gradients, so there is a preference for small apertures. But as we have mentioned in Section 9.2.1 above and shall show below in Section 9.2.4, given β in the doublet, there is a minimum beam aperture, $n_x\sqrt{(\varepsilon_x\beta_x)}$, that can be collimated which is determined by the ratio $N^{1/2}/\sigma_z^{1/3}$. Hence the final-doublet aperture is chosen as small as possible subject to the limits of the collimation system, the range of IP β functions which is to be accommodated, and resistive-wall and geometric wake considerations.

The IP β functions are adjusted to an optimum during operation. If the beam is free from tail particles, the optimum will be determined by the maximum luminosity, or equivalently the minimum IP spot size that can be obtained with the given beam emittance. As the IP β functions are decreased, the linear spot size gets smaller, but the strengths of aberrations from nonlinearities in the final-focus system get larger. The values of the IP β functions shown in the parameter sets of Chapter 12, Table 12-1 will be close to the optimum. Synchrotron radiation, Oide effect and beamstrahlung control are also involved in finding the optimum.

Tail particles create problems in two ways (see Chapter 12 for details): i) by impacting an element of the final doublet, or ii) radiating photons which impact the final doublet. Studies of shower particles and their trajectories through the detector show that only tens of particles per bunch train can be allowed to impact the final doublet. Studies of synchrotron radiation lead to the guideline that the horizontal final-doublet aperture should be twice the collimated aperture, and the vertical aperture should be 1.5 times the collimated aperture.

The answer to the question, "What should be the final-doublet phase collimation depth?" turns out to be "As small as possible." As we shall see below, for the beam parameters under discussion, the minimum collimation apertures are $5\,\sigma_x$ and $35\,\sigma_y$. There needs to be some margin on this number, since the apertures must be gradually enlarged in successive collimation stages, and by the final doublet these minima may increase to $7\,\sigma_x$ and $40\,\sigma_y$. Following the synchrotron radiation guidelines, the final-doublets apertures should be chosen to be equal or greater than $14\,\sigma_x$ and $60\,\sigma_y$. Final-doublet designs are discussed in Section 11.6 and particle trajectories are shown in Section 11.6.5.

Interaction Point (IP) Phase Depth. The collimation depth for the interaction-point (IP) phase can be larger than the final-doublet (FD) phase since in the final-focus system the IP phase is demagnified, and apertures for this phase are more than $45\,\sigma_x$ and $200\,\sigma_y$. Figure 9-5 shows the dynamic aperture of the horizontal and vertical phases of the final-focus system determined from tracking.

There are several advantages of collimating the IP phase at a larger aperture than the FD phase.

i) The wakes will be smaller, and since these wakes give rise to jitter in the final-doublet phase, the beam jitter is reduced in the final-focus system, which is helpful for tuning.

ii) The β functions in the second stage IP-phase collimation can be chosen smaller. This means that the IP-phase collimation section can be shorter and the tolerances less severe.

The smaller β functions of item ii) can only be achieved in the first IP collimation stage if one abandons the principle that the spoiler should be able to survive the impact of a full bunch train. This needs to be carefully evaluated in the context of the machine protection system.

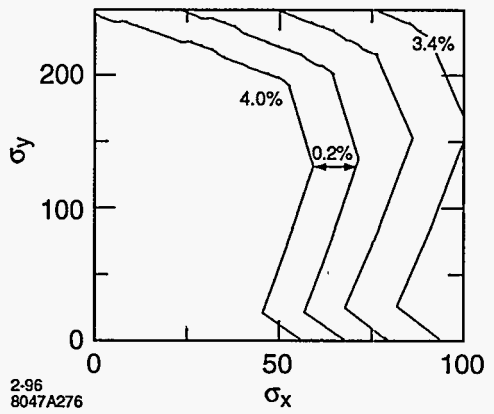


Figure 9-5. *The dynamic aperture of the final-focus system for the IP phase. The FD phase particles have been restricted to be within a diamond region in the $x - \delta$ plane defined by the collimation system, with $|x| \leq 7\sigma_x$ and $|\delta| \leq 0.04$.*

Just how large the collimation in the IP phase can be also depends on beam-line transport within the collimation system and from the collimation system to the IP. The beam line from the collimation system to the final-focus system must be achromatic, so that off-energy IP-phase particles do not migrate into the FD phase. It also requires that there not be significant residual beam-line aberrations from sextupoles or interleaved sextupoles, which distort the phase space of large amplitude particles. These question are addressed in Section 9.2.5.

There is a further constraint on IP-phase apertures that comes from machine-protection considerations. Mis-steered beams that come close to collimator apertures experience very large wake kicks, and can be steered into unprotected downstream absorbers. As we will see below, it is necessary to have two stages of IP collimation in the collimation system. The machine protection considerations alluded to will require that the second stage collimation have an aperture that is much larger than the first stage. These issues are discussed in detail in Sections 9.2.4, "Quadrupole and Large Amplitude Wakes of Parallel Plate Collimators", and "Consequences of Quadrupole and Near-Wall Wakes". We assume throughout that the collimators in the first IP stage are protected by spoilers that can withstand the impact of a full pulse train.

Energy Collimation Depth. The required energy collimation depth also depends on the properties of the lattice between the collimator and the final doublet. If an off-energy beam is able to pass through the spoilers, it should not be able to impact any downstream collimator in the collimation system or in the final-focus system. Figure 9-6 shows the result of tracking particles through the beam line from the end of the collimation system to the IP for an early beam-delivery lattice. The initial coordinates are chosen at the collimator edges of $6\,\sigma_x$ and $40\,\sigma_y$, in both the IP and FD phase, and tracked to the location of the horizontal and vertical collimators of the final-focus system. One sees in Figure 9-6 the image of these particles when they are off-energy by various amounts from 0 to -6%. Based on these and similar studies described in Section 9.2.5 we have chosen 4% as the energy collimation depth.

Because of the large β functions in the collimation system, there must be a chromatic correction scheme of sextupoles placed in dispersive regions [Brown 1979]. It is natural to take advantage of this dispersion function to collimate energy at the same locations where the horizontal plane is collimated. Though the magnitude of the dispersion function is somewhat flexible, 4% energy collimation turns out to be a possible and convenient energy depth.

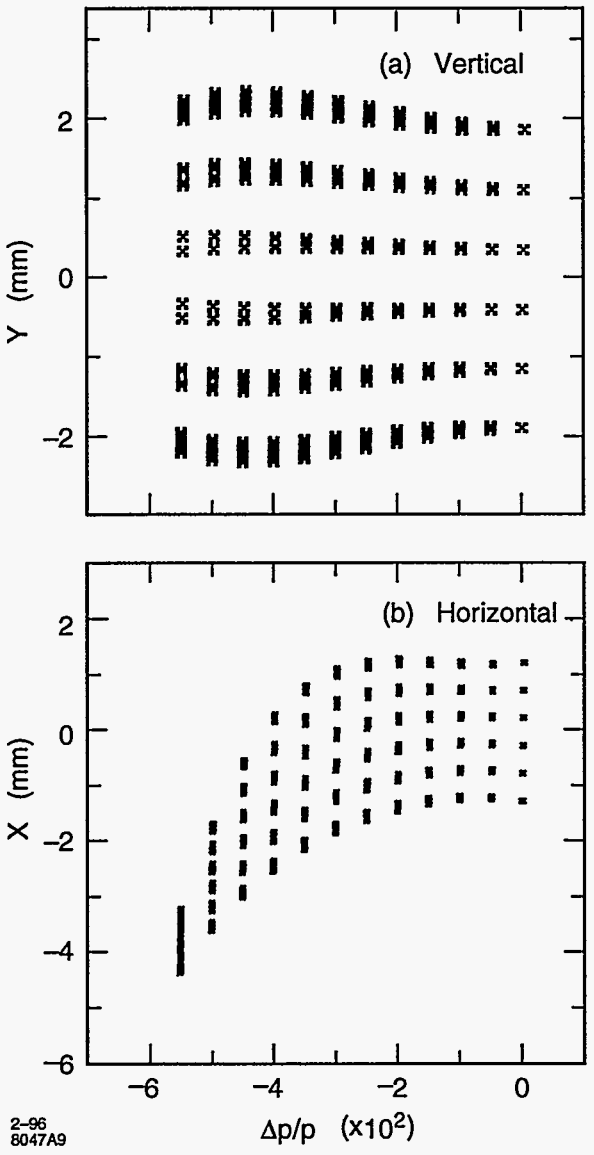


Figure 9-6. *The position of particles at the horizontal collimator in the final-focus system for particles launched in the collimation system on the boundary of the collimation aperture for several energy offsets.*

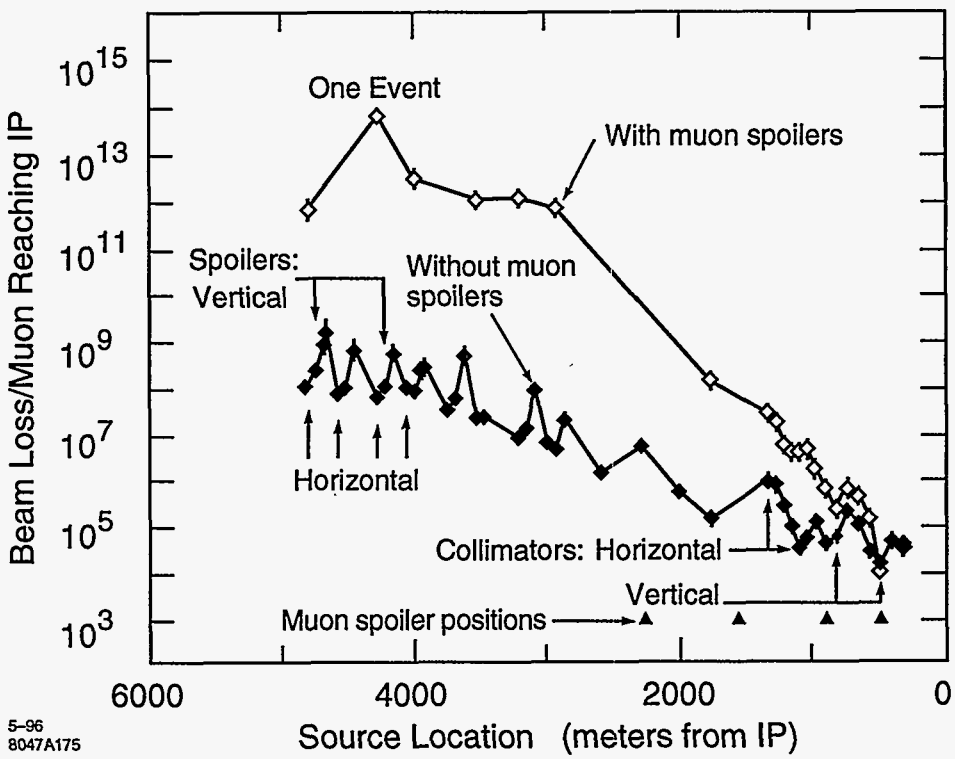


Figure 9-7. *Graph showing electrons that may be collimated if only 1 muon is to reach the detector.*

Limits on Collimation in the Final-focus System

The CCX and CCY chromatic correction sections in the final-focus system provide excellent locations for collimators. However collimation here can create unacceptable muon backgrounds. Figure 9-7 shows the number of particles that may be collimated, and still have just one muon pass through the detector. If muon toroids are placed in the final-focus tunnel, one finds that 10^5 particles may be collimated in the CCY and 10^6 may be collimated in the CCX.

Figures 9-8 and 9-9 show the results of edge scattering within the final-focus system. These studies indicate that less than 1 particle in 10^5 of the particles that re-enter the beam at the collimator edge actually impact the final doublet. This simulation was done assuming the angular and energy edge-scattering profiles presented in Section 9.2.3. In fact no particles in this study impacted the final doublet. Either the particles impacted the beam line between the collimator and the doublet (at the latest in the bend in the final telescope) or they passed through the IP. Particles that are close to full energy will miss the doublet because the collimators are π from the final doublet and have a smaller aperture than the final doublet. About 10% of the particles are in this category. Particles that are off-energy don't make it through the bends. Since only one particle re-enters the beam for every 10 particles per micron incident on the edge of the collimator, this study would say that there is an allowed density of 10^6 particles per micron on these collimators. Since the muon studies conclude that only 10^5 to 10^6 particles may be incident on the whole collimator, the muon studies set the limits on the number of particles that may be collimated in the final-focus system.

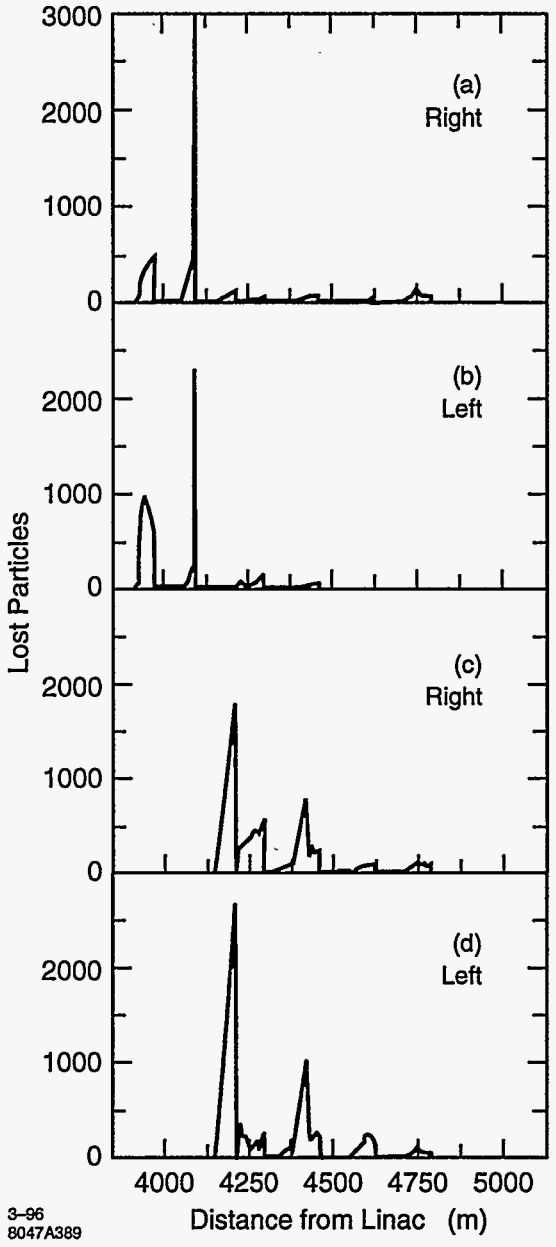


Figure 9-8. *Edge scattering results from the right and left edge of the two horizontal collimators in the CCX section of the final-focus system. In each case 100,000 particles were released from the edge according to angular and energy distributions determined from EGS studies (see Section 9.2.3). The curves give the number of particles which impact elements (or the beam pipe) along the beam line. In both cases no particles impact the final doublet or final-doublet region. About 10% of the particles pass through the IP.*

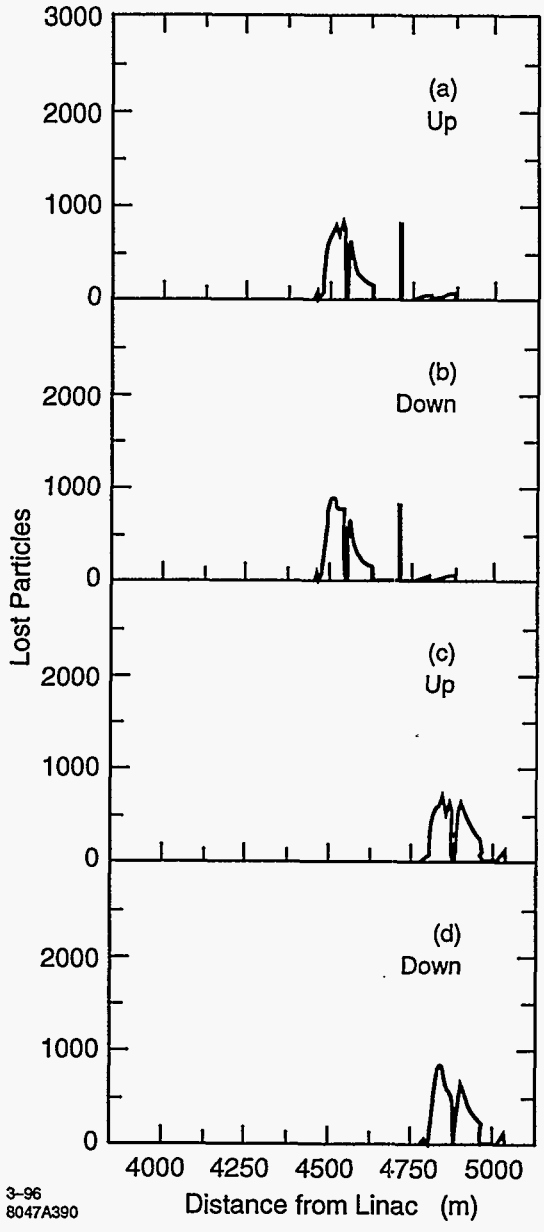


Figure 9-9. *The distribution of impact position for particles edge-scattered from the edge of the vertical collimators at the sextupoles in the CCY region of the final-focus system. A total of 100,000 particles were released from the collimator edge according to the angle and energy distributions shown in Figures 9-19–9-21. None of these particles impact the final doublet. About 10,000 pass through the IP. "Up" and "down" are from the top and bottom edges, respectively.*

Studies described in Chapter 12 indicate that tens of particles can be incident on the final doublet per bunch train without saturating the vertex detector. We conclude that the principal source of such particles will be from scattering on gas particles in the final-focus system. See Section 11.5.5, Eq 11.68 for an estimate of this source.

9.2.2 Materials Considerations

Machine Protection Issues

Single-Bunch Passive Protection. The machine protection strategy (see Chapter 16) consists of a two-level scheme. At level 1 the beam line should be able to passively survive the transport of a single bunch, no matter the state of any device in the beam line. At level 2, which assumes that a single bunch is known to successfully travel through the beam line to a beam dump, the emittance of the single bunch can be reduced and the number of bunches can be increased. It is now the responsibility of the machine protection system to detect any state changes in the system that could lead to the beam not being safely transported to a dump. Upon detection of such a state change it must be possible to turn off the beam before any component is damaged.

For a single bunch, the area of the beam when encountering a copper element must have a radius of about $100\,\mu$m (see Table 9-2). Since the divergence angle following a spoiler is about $15\,\mu$r (see Figure 9-3), the distance from the spoiler to any copper element must be equal or larger than 7 m or 10 m for a 500 or 750-GeV beam energy, respectively. The beam radius at a spoiler, assuming it is composed of 1/4 r.l. titanium, must satisfy $\sqrt{(\sigma_x\sigma_y)} \geq 10\,\mu$m. Since there are many locations where $\sqrt{(\sigma_x\sigma_y)} < 10\,\mu$m, the single-bunch beam emittance must be enlarged for level 1 beam-line check out. This is accomplished by blowing up the beam emittance in or at the exit of the damping ring.

Level 1 objectives are achieved by placing spoilers throughout the system, so that no matter what the trajectory of the bunch, it impacts a spoiler before any other element. Figure 9-10 shows a possible single-bunch trajectory, assuming a worst-case scenario of a shorted quadrupole with an on-axis field equal the design pole-tip field followed by a quadrupole which is defocusing in the plane of the kick. In this example the distance between quadrupoles is taken to be 30 m, and the focal length of the quadrupoles taken to be 20 m. If a spoiler, with a gap of 3/5 a, is placed 10 m before each element, then assuming that the distance between quadrupoles is $d = 3/2f$, where f is the focal length of the quadrupoles, the worst case trajectory has a displacement of 9/10 a at the second quadrupole and would have a displacement of 3 a at the third quadrupole. Thus if the beam pipe between quadrupoles has a radius greater than 3 a, all beam-line elements are protected at level 1.

We take as a guideline that the aperture at the beam-line elements is 2.5 times the gap of any nearby spoiler or absorber of the collimation system. Since the largest aperture of the spoilers or absorbers in the first collimation stage is 2 mm (see Section 9.2.5), it is consistent with this guideline to assume the element aperture is 5 mm. It follows that the beam-pipe inner radius between beam-line elements must be larger than 15 mm.

Under the conditions envisaged in Figure 9-10, when a bunch passes close to a spoiler edge it can receive a wake-field kick from the spoiler. The magnitude of these kicks is calculated in Section 9.2.4. The maximum kick goes like $1/\sigma^2$ where σ is the largest axis of the beam ellipse at the spoiler. Assuming the smallest major axis within the first stage of the collimation system is $100\,\mu$m, this maximum kick is about 150 nr. Between collimation sections, the beam size is much smaller and the kick can be 100 times larger, or $15\,\mu$r.

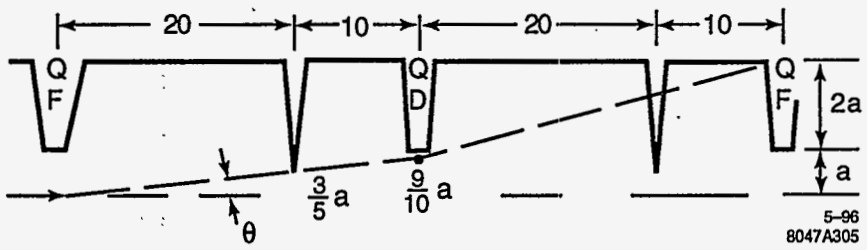


Figure 9-10. *Possible trajectories of a single bunch through a beam line as a result of completely or partially shorted quadrupole legs. Every trajectory encounters a spoiler at least 10 m before reaching the wall or another element. The quadrupoles are assumed to be spaced at 30 m and have a focal length of 20 m in this example.*

This is still much smaller that the kick angles from the quadrupoles, which in Figure 9-10 are as large as 250 mr. Thus these wakefield kicks will not compromise the level 1 protection scheme we have described.

A complete treatment of this problem would involve a study of each beam-line segment for a complete list of assumptions on failure modes. These sections on machine protection are intended as a guide to the nature of a solution, and are not intended to represent a complete study.

Since the machine-protection spoilers have apertures that are not much larger than the system spoilers and absorbers, it will be important that they are tapered to minimize the wakes. (See Section 9.4.5.) The machine-protection spoilers will look very much like the system spoilers. Since the effect of the wakes from a taper go as $\beta/g^{3/2}$, where g is the spoiler gap, only spoilers at large β will have a significant impact. The wake from the largest-β machine-protection spoiler, if constructed according to the above guidelines, would be roughly 0.3 of the wake from the main system spoiler. All machine-protection spoilers taken together can add an estimated additional 10% ($t'/t = 0.07$) to the jitter amplification budget (see Section 9.2.4). This is large enough that it should be reviewed carefully and taken into account when choosing the vertical collimation depth.

Multiple-Bunch Active Protection. For a full bunch train the machine must be turned off before the beams collide with any of the machine protection system (MPS) spoilers, since the full train would easily destroy them. Let us suppose as a worst-case scenario that a quadrupole leg is suddenly shorted. Let us suppose that the on-axis field has reached the fraction ξ of the pole tip field. This will give an angular deflection $\theta = \xi a/f$, where a is the quadrupole aperture and f is its focal length. The beam displacement downstream will be $\Delta x = \theta_x R_{12} \leq \xi R_{12} a/f$. Since the R_{12} values can be the order of $2\,10^4$ m, for $f = 20$ m we would have $R_{12}/f = 10^3$. This means that the machine must turn off before $\xi = 10^{-3}$. Assuming the fastest rate at which the magnet can change field values is the order of 100 milli-seconds, it could change its value by 10^{-3} in 0.1 ms. Since the pulse-repetition rate is 2 ms, the magnets could change an unacceptable amount between beam pulses. Thus this error must be detected with a magnet-monitoring system, and can not be prevented by a BPM system monitoring the beam trajectory. We will assume such a magnet-monitoring system is in place.

If the magnet-monitoring system were to fail, a spoiler could be destroyed. In that case, the beam would not be sufficiently enlarged upon reaching the next beam-line element to protect it from damage. It may therefore be worthwhile to place sacrificial copper absorbers in the beam line in front of each element whose purpose would be to absorb the impact of the full beam. They would need to be replaced after a protection system failure event, but would limit damage to the system.

Property	Unit	Carbon	Copper	Ti-6Al-4V	Tungsten-Rhenium
Density	(gm/cm^3)	2	8.95	4.5	19.3
Radiation length	(cm)	18.9	1.44	3.77	0.344
Resistivity	(μohm-cm)	1000	1.7	180	5.5
Specific heat	Joules/(gm°C)	0.97	0.385	0.54	0.134
Thermal conductivity	Watt/(cm°C)	2.0/0.025	3.9	0.17	1.26
Stress limit	(°C)	2500	180	770	700
Melting point	(°C)	3600	1080	1800	3100
Vaporization temperature	(°C)	4200	3000	3260	5000

Table 9-1. *Properties of materials considered for spoilers and absorbers.*

Multiple-Bunch Passive Protection for Collimators. There are many upstream system-state changes that could cause the beam in the collimation system to impact the main collimators. Major changes will be handled by upstream machine-protection systems, but because the β functions must be large enough to collimate particles that could impact elements in the final-focus system, the change in orbit will be larger in the collimation system than in any upstream system. And since everything in the final-focus system will be in the shadow of the collimation system, the collimation system will be the most vulnerable and important to protect.

The strategy of relying on monitoring all possible state changes in upstream systems becomes a very complex and potentially unreliable task. It seems unwise to assume the collimation system can be actively protected if collimation is to occur at apertures as small as $5\,\sigma_x$. Thus we have assumed that the collimation system collimators must be passively protected. (This decision needs to be carefully reviewed as the machine protection system becomes more mature. The cost impacts are large.)

Passive protection has major consequences, and implies that whatever the material, it must be able to withstand the full impact of at least one full bunch train. In the system to be described below, this is accomplished by having a sufficiently large linear beam size at the spoiler so that it can withstand this impact. Another scenario, which uses sextupoles to blow up the beam size, is discussed in Section 9.2.8. Other schemes, such as liquid-metal collimators and Compton collimators have been proposed. Though we have discussed such systems, we have not pursued them in great depth, preferring a more conservative solution. Since the system we describe is quite expensive, thorough investigations of alternative systems are encouraged.

Properties of Collimator Materials

For the 1-TeV c.m. parameters, each beam contains an average power of 8.4 MW. If routine collimation of 1% of the beam is adopted as a system specification, the energy intercepted at any phase could be as high as 84 kW. Half of this amount could be incident on a single absorber. Devices capable of absorbing and dispersing such heat loads are most often made from copper because of its high heat conductivity, compatibility with water, reasonable size, and general adaptability. See Table 9-1 for a list of some relevant physical properties for copper and some other materials we have considered for spoilers and absorbers.

Figure 9-11 gives the energy-density deposition in three of these materials from a beam of 10^{12} particles with initial cross-sectional area given by $\sqrt{(\sigma_x\sigma_y)} = 100\,\mu$m.

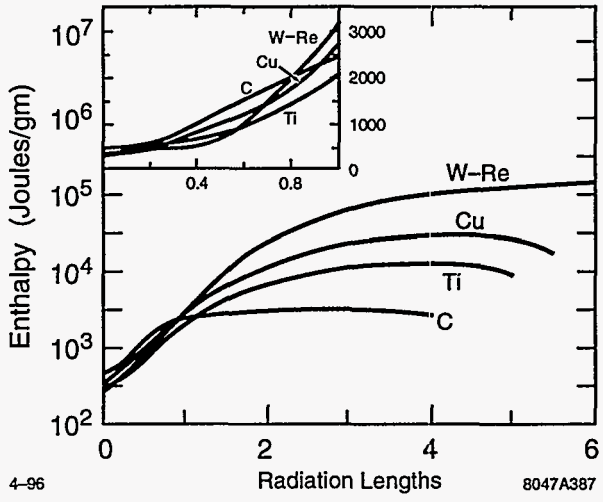


Figure 9-11. *Energy density deposition in four materials for an incident beam of* 10^{12} *particles with cross section given by* $\sqrt{(\sigma_x \sigma_y)} = 100\,\mu m$.

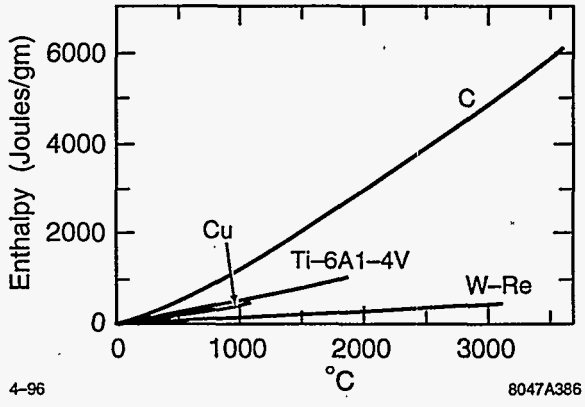


Figure 9-12. *Temperature rise versus enthalpy for four materials.*

For other beam intensities and areas the curves of Figure 9-11 may be scaled by $N/(\sigma_x \sigma_y)$ because in the distances under consideration, showers have little chance to broaden beyond the initial spot size. The multiple-scattering formula (see Eq. 9.8) yields about $30\,\mu r$ at $E = 500\,\text{GeV}$ and $t = 1$ r.l., so the initial spot size of $30\,\mu m$ would not significantly change in several radiation lengths.

Since the energy is deposited instantaneously without any opportunity for the volume to change, the energy-density deposition represents a change in enthalpy. Figure 9-12 gives the temperature rise versus enthalpy for the same four materials, incorporating the changes of specific heat with temperature.

For beams in the regime of 1-mm^2 cross section, it has been found experimentally that when a sudden local temperature rise creates local thermal stresses near the surface which exceed the tensile limit of the material, the material will probably fracture. If the temperature of copper rises suddenly by a mere 180°C, the tensile limit is exceeded (see Table 9-1). Hence for a single pulse of 10^{12} particles to give a temperature rise of less than 180°C anywhere within 2 cm of the surface, Figures 9-11 and 9-12 indicate the beam area must satisfy $\sqrt{(\sigma_x \sigma_y)} \geq 2.2\,\text{mm}$. These limits for the remaining materials in Table 9-1 are shown in Table 9-2

	Fracture 2 cm from a Surface			Melt Anywhere within Volume	
Material	Enthalpy (Joules/ gm)	Incident Surface Minimum Sigma (mm)	Parallel Surface Minimum Sigma (mm)	Enthalpy (Joules/ gm)	Minimum Sigma (mm)
Carbon	3870	0.036	0.090	6000	0.072
Titanium alloy	405	0.14	0.56	960	0.36
Copper	63	1.0	2.2	470	0.82
Tungsten-rhenium	91	4.2	42.	410	2.0

Table 9-2. *Table of minimum $\sqrt{(\sigma_x\sigma_y)}$ to avoid fracture near the surface or melting of a collimator. It is assumed that there are 10^{12} particles per pulse train, and for the melt limit a particle energy of 500 GeV.*

[Walz 1973]. A beam incident on an absorber will always be within 2 cm of the collimation surface, so the temperature of the entire shower within the material must remain less than 180°C. The numbers for the incident surface apply only if the device is a dump. The fracture limit does not depend on beam energy. The melt limit is derived assuming a beam energy of 500 GeV.

To achieve 1-mm beam areas with the contemplated design emittances would require $\sqrt{(\beta_x\beta_y)} > 10^6$ m. Such β functions are impractical, and the usual strategy to achieve passive protection of collimator systems is to have a spoiler followed by an absorber [DeStaebler-Walz].

Properties of Spoilers

Spoiler-Absorber Strategy. The spoiler is fabricated from a thermally-rugged, optically-thin low-Z material so that a full shower does not develop in the material. The requirement is that the temperature rise caused by a full bunch train incident on the spoiler not lead to fracture or melting. The spoiler increases the angular divergence of the incident beam so that down-beam copper absorbers can withstand the impact of one full bunch train. Additionally the absorber must be able to continuously absorb all particles hitting the spoiler even when this is as much as 1% of the beam.

Linear modules may be used to create the large β functions to achieve the required beam sizes at the spoilers, but β enlargement has two limitations: i) creating large β functions takes a long system length, and ii) wakefield kicks from the spoilers become large compared to decreasing angular beam divergences. We show below that because of ii) there is a limit to which a beam may be collimated with a linear β enlargement.

The large β functions create chromaticity which will increase the beam emittance unless locally compensated. This suggests the use of sextupoles in $-I$ configurations in each collimation phase and each transverse plane. Energy must be collimated first, since the bends employed to generate dispersion at the sextupoles could steer an off-energy beam into the wall of the beam pipe.

Single-Train Spoiler Survival. The minimum sigma required for a 0.25 r.l. spoiler are shown in Table 9-3. The best material we have found for use as a spoiler is pyrolytic graphite. However because

Material	Minimum Sigma (μm)
Carbon	50
Copper	280
Titanium	100
Tungsten-Rhenium	225

Table 9-3. *Table of minimum $\sqrt{(\sigma_x\sigma_y)}$ to insure survival of a 1/4 r.l. spoiler for a single pulse of a beam containing 10^{12} particles.*

it is porous and has a high resistivity, it must be plated. Even thin platings will separate from the plated material or melt if the beam size is not larger than the values which are derived from considerations of the energy deposition and temperature rise curves of Figure 9-11 and 9-12.

We have concluded that plated pyrolytic carbon spoilers are a bad idea for three reasons:

i) Carbon has not been used before, and there are questions about its performance under vacuum.

ii) Though it reportedly can be plated, details and experience are scarce.

iii) The plating, even if quite thin, will indeed be damaged unless the entry spot size is large enough to prevent melting and stress damage of the plating material.

From item iii) one concludes one might as well make the spoiler from the materials considered for the plating.

The next best material is titanium. Titanium also has a rather large resistivity when compared to good conductors (the resistivity of pure titanium is 45 $\mu\Omega$-cm, and the usual titanium alloy has a resistivity of 180 $\mu\Omega$-cm, compared to 1.7 $\mu\Omega$-cm for copper), but the wake of an angle-optimized taper goes as the 1/4th power of the resistivity (see Section 9.2.4), so it is not out of the question to use titanium, especially if one uses the alloy plated with the pure titanium. TiN which has a resistivity of 22$\mu\Omega$-cm may also prove useful as a plating material. The flat section of these spoilers will be very short since we want the longest path of particles through the material to be 0.25 r.l. (< 1 cm). In fact, as can be deduced from formulae presented in Section 9.2.4, the wake from one of these titanium spoilers is identical to the wake from a 20 r.l. Cu tapered absorber for a gap of 1 mm.

Because as indicated in Table 9-3, the entry spot size required for titanium must satisfy $\sqrt{(\sigma_x\sigma_y)} \geq 100\,\mu$m, whereas for carbon the requirement was $\sqrt{(\sigma_x\sigma_y)} \geq 50\,\mu$m, the β function product $\sqrt{(\beta_x\beta_y)}$ at the spoilers must be larger by a factor of four. The collimation system length increases by about 500 m for titanium spoilers as compared to carbon.

To keep the path length within a tapered spoiler to 0.25 r.l. radiation length, we propose fabricating it from 50-μm thickness rolled titanium which is then supported by a titanium honeycomb structure. Assuming a taper angle of about 20 mr, the path length in the incoming and outgoing taper will be 2.5 mm each. The rest of the 1-cm 0.25 r.l. thickness in titanium, can be in the honeycomb support structure. The honeycomb should be oriented at 45° to the line-of-flight of the particles.

The longest path length will occur at the tip of the spoiler. If the radius of curvature at the tip is 0.5 m, then the longest path in the titanium will be 1/4 r.l. Edge scattering is not an issue for spoilers, so the radius of curvature can be this small. A sketch of such a spoiler is shown in Figure 9-24(b).

To obtain the required spot size at the spoiler we must have $\sqrt{(\beta_x\beta_y)} \geq 10^{-8}/\sqrt{(\varepsilon_x\varepsilon_y)}$. At 500-GeV c.m., the smallest possible emittance product (see Section 9.2.1) at the end of the linac is $\sqrt{(\varepsilon_x\varepsilon_y)} = 10^{-12}$, and thus we must have $\sqrt{(\beta_x\beta_y)} \geq (100)^2\,\mathrm{m}$. For 1-TeV c.m., the smallest possible emittance product, $\sqrt{(\varepsilon_x\varepsilon_y)} = 0.5\ 10^{-12}\,\mathrm{m\text{-}rad}$, implies $\sqrt{(\beta_x\beta_y)} \geq (140)^2\,\mathrm{m}$. For 1.5-TeV c.m., the smallest possible emittance product, $\sqrt{(\varepsilon_x\varepsilon_y)} = 0.33\ 10^{-11}$, implies $\sqrt{(\beta_x\beta_y)} \geq (170)^2\,\mathrm{m}$.

Spoiler Heating. Each electron passing through the spoiler will deposit an energy of $1.8\,\mathrm{MeV/cm^2}$. For the titanium alloy with a density of $4.5\,\mathrm{g/cm^3}$, this amounts to an energy deposit of 8.1 MeV/ cm per electron. Since the path length of each electron in the titanium is in fact about 1 cm, this is an energy deposit of 8.1 MeV per electron. For 10^{10} electrons per train and 180 trains per second this amounts to 2.3 W. If we assume that the spot hitting the titanium has a cross-section of a semi-circle with a radius of $100\,\mu\mathrm{m}$, then the surface of the spoiler, if straightened into a plane would be intersected in an area of width $200\,\mu\mathrm{m}$ and length equal to $d = 1\,\mathrm{cm} + 2(100\,\mu\mathrm{m}/\theta) \approx 3\,\mathrm{cm}$, where θ is the taper angle. Heat can flow both directions toward the walls, which we will assume is a heat sink at room temperature about 5-mm away. Since the length of the spot is six times larger than the distance to the sink, we assume a simple one-dimensional parallel heat flow. Using the thermal conductivity for titanium of $0.17\,\mathrm{W/(cm\text{-}^\circ C)}$, we find a temperature elevation at the center of 230°C. If the 10^{10} electrons had been distributed between the two jaws, or the semi-circle had a radius of 0.2mm, the temperature rise would be only 115°C. In any case, these are acceptable temperature rises in titanium.

Spoiler Transmission. Figure 9-13 shows an EGS simulation [Nelson 1993] of the angular distribution for several 1-GeV energy bins of particles that have passed through a 1/4 r.l. spoiler. The angular distributions are remarkably similar for all energies, peaking at about $14\,\mu\mathrm{r}$. The solid line through the particles with energy in the 499–500-GeV bin is an expected distribution based on multiple scattering theory given by

$$\frac{dP}{d\Omega} = \frac{1}{2\pi\theta_0^2} e^{-\frac{\theta^2}{2\theta_0^2}} \qquad \text{with} \qquad \theta_0 = \frac{13.6}{E(\mathrm{MeV})}\sqrt{t}\,(1 + 0.038\ln[t]) \quad . \tag{9.8}$$

$\theta_0 = 14\,\mu\mathrm{r}$ is, respectively, 10^3 or 10^4 times larger than the horizontal or vertical angular beam divergence at the spoilers. Therefore in order for particles coming from the spoilers to be incident on the edges of downstream collimators, they must have $\theta \ll \theta_0$.

Figures 9-14 and 9-15 show the energy distribution of exit electrons and exit photons, respectively, from 500-GeV electrons incident on a 1/4 r.l. spoiler. The energy distribution of the photons is fit very well by a 1/k law. The energy distribution of the electron is well fit by a formula based on single-scattering which is [Rossi 1952]

$$\frac{1}{N}\frac{dN}{dE} = \frac{1}{E_0\Gamma\left(\frac{t}{\ln 2}\right)}\left[\ln\left(\frac{E_0}{E}\right)\right]^{\left(\frac{t}{\ln 2}-1\right)} \quad . \tag{9.9}$$

The diamonds in Figure 9-14 are calculated using this equation. Figure 9-14 has an insert which is a blow-up of the region from 450 to 500 GeV. The dots in the insert are also calculated with the theoretical formula (Eq. 9.9). For small values of $\delta = \Delta E/E$, Eq. 9.9 becomes

$$\frac{1}{N}\frac{dN}{d\delta} \approx \frac{1}{\Gamma(1.44t)}\,\frac{1}{\delta^{1-1.44t}} \quad . \tag{9.10}$$

With the energy and angular distribution known, and the angular distribution independent of energy, it is straightforward to initialize distributions of particles that have passed through spoilers for the purpose of tracking them through the lattice. See Section 9.2.5, Figures 9-39 and 9-41.

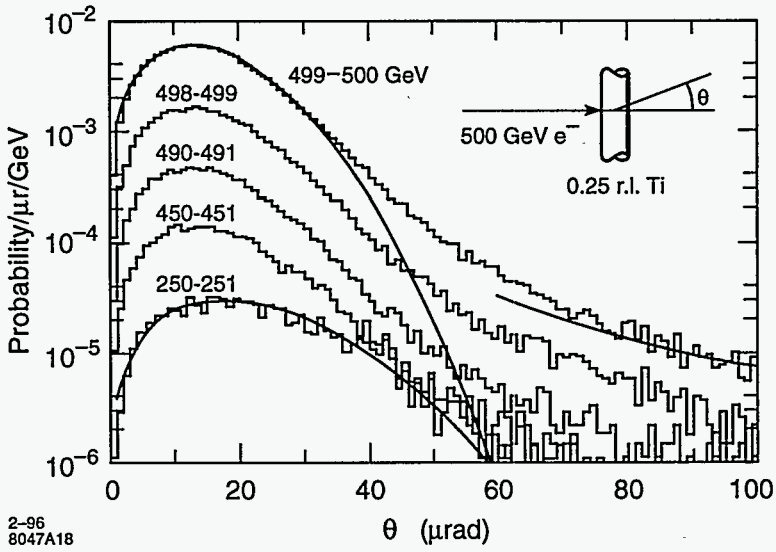


Figure 9-13. *Angular distribution of 500-GeV electrons that have passed through a 0.25 r.l. spoiler. The solid line is the theoretical prediction given in Eq. 9.8.*

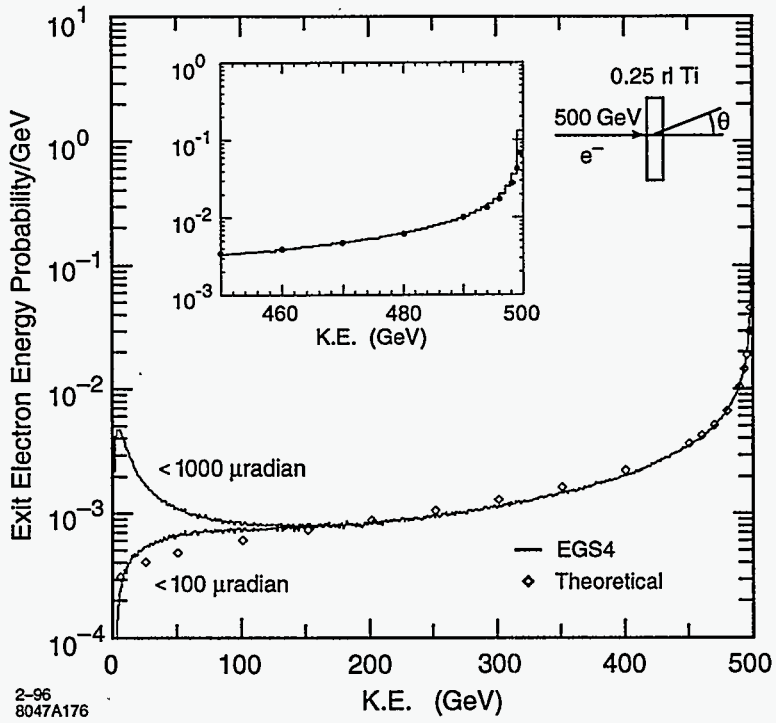


Figure 9-14. *The energy distribution of electrons from a 500-GeV beam that have passed through a 0.25 r.l. spoiler.*

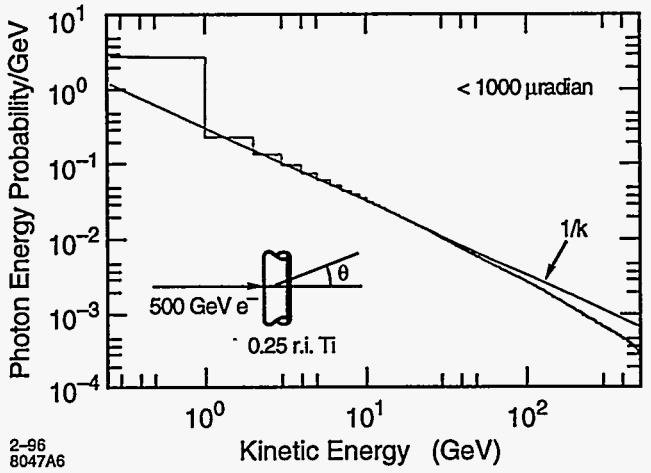


Figure 9-15. *The energy distribution of photons created by a 500-GeV beam of electrons passing through a 0.25 r.l. spoiler.*

Required Spoiler-Absorber Separation

If a beam is completely mis-steered and impacts a spoiler, the beam sigma product at the down-stream copper absorber should, according to the results of Section 9.2.2 and Table 9-2, be greater than $(2.2\,\mathrm{mm})^2$. The shape of the image on a down-beam absorbers of the particles exiting a spoiler will be an ellipse. The horizontal and vertical axes will have Gaussian distributions with rms width equal to $\theta_0 R_{12}$ and $\theta_0 R_{34}\,\mathrm{m}$, respectively. Thus we have the condition $\theta_0^2 R_{12} R_{34} \geq (2.2\,\mathrm{mm})^2$. For $\theta_0 = 14\,\mu\mathrm{r}$ we deduce $\sqrt{(R_{12}R_{34})} \geq 160\,\mathrm{m}$.

This criteria can be eased somewhat by realizing that, on the average, the beam has lost one quarter of its energy when passing through the 1/4 r.l. spoiler, so that when passing off-axis through a quadrupole between spoiler and absorber the beam is spread out. This effect is not as large as the average energy loss would indicate since the distribution is peaked at small energy loss. For example, only 25% of the particles have lost more than one-third of their energy. The exact magnitude of this effect will depend on lattice details, but if we assume that 25% of particles are not contributing to the core heating within the absorber, the condition of the last paragraph is modified to $\sqrt{(R_{12}R_{34})} \geq 140\,\mathrm{m}$.

If there is a quadrupole of inverse focal length f at a distance L_1 from the spoiler, followed by a drift of length L_2 to the copper absorber, then the $R_{12}R_{34}$ product is given by

$$R_{12}R_{34} = (L_1 + L_2)^2 - \left(\frac{L_1 L_2}{f}\right)^2 \quad . \tag{9.11}$$

Using Eq. 9.11 the current lattice does not satisfy $\sqrt{(R_{12}R_{34})} \geq 140\,\mathrm{m}$, since this lattice was designed for a lower beam intensity and a less stringent fracture criteria. A next-generation lattice under development does satisfy this condition. We will discuss future research directions in Section 9.2.9.

9.2.3 Tail Re-Population Estimates

Edge Scattering from Absorbers

Edge scattering from a spoiler is not a concern since very few particles will be incident on the edge (which, we will see below, has an effective width of about 1 μm) and, in any case, all the particles which pass through the spoiler are scattered at angles the order of 14 μr, which is similar to edge scattering. However at the absorbers, where one is hoping to clean up the beam, it is the edge scattering which re-populates the beam.

Figure 9-16 shows the results of an EGS simulation [Nelson 1993] to determine the fraction of particles that re-enter the beam as a function of distance form the edge of a flat titanium collimator edge (Note: We are expecting to use copper, but the curve can be readily scaled for all materials as shown in Figure 9-17). Only particles losing less than 4% of their energy have been counted because particles losing more energy will be collimated elsewhere. Later we present distributions with energy loss up to 20%. There is an abrupt edge in penetration at about 0.1 μm (note logarithmic scale) establishing this as the relevant edge-depth scale for titanium. To insure that the surface of an absorber with a length of a several cm be placed parallel to the beam with an accuracy of better than 0.1 μm would require an unlikely angular-placement accuracy of 1 μr. Without this placement accuracy particles in the incident beam would hit only a corner of the absorber and a larger number would be scattered back into the beam.

To avoid this effect, it has been usual to use collimators with curved surfaces [von Holtey]. Figure 9-16 also contains a curve showing the penetration probability for a curved titanium spoiler with a 100-m radius of curvature, which would require an easier 0.2-mr orientation accuracy. The transmission of the 100-m curved surface is increased by about a factor of 7, with a transmission edge at 0.7 μm. A third curve on this graph shows the edge scattering for a surface with a radius of curvature of 20 m, which would require an orientation accuracy of 1 mr. This has a transmission edge at about 3 μm, with a transmission that is a factor of 4 worse than the 100-m curvature. In the SLC collimator radii are about 10 m. The eventual flattening of these transmission curves at large offsets occurs because the EGS study was performed for a 2 r.l. absorber. If it had been thicker the curves would continue to fall.

Figure 9-17 compares the penetration probability of a curved titanium collimator with a curved carbon collimator, with the abscissa now chosen to be micro-radiation lengths (which we will denote by μrl) rather than microns. In terms of micro-radiation lengths the collimators perform identically. Since the radiation length of titanium is 3.77 cm, whereas copper's radiation length is 1.44 cm, the knee for a 100-m curved copper absorber occurs at 0.3 μm. For a 20-m curved copper absorber the knee would be at 1 μm and for a tungsten absorber it would be 0.3 μm. It is interesting to note that one can use the results of Figure 9-17 to take into account the roughness of an absorber by degrading its radiation length [DeStaebler 1994].

Figure 9-18 shows the angular distributions for edge-scattered particles re-entering the beam with energy losses of less than 20%. The angle in this distribution is the cone angle measured from the particle direction. The particles incident on the absorber were uniformly distributed from 0 to 10 μm from the edge of a Cu absorber with a radius of curvature of 100 m. Note the remarkable similarity to the angular distributions from spoiler transmission shown in Figure 9-13. The total number of particles scattered back into the beam (with energy loss less than 20%) is 0.1 times the number of particles incident on the edge per micron. In other words the number re-entering the beam is equal to the number that fall on the edge at a distance of 0.1 microns or less. The theoretical distribution of Eq. 9.8 fits this graph rather well for an rms angle of $\theta_0 = 20\,\mu$r. The effective thickness of the material (using Eq. 9.8) would be $t_{eff} \approx 0.5$ r.l.

Figure 9-19 shows the angular distribution when the angle is chosen to be the angle from a plane of the collimator surface, for the same condition as those in Figure 9-18. This distribution is rather surprising in

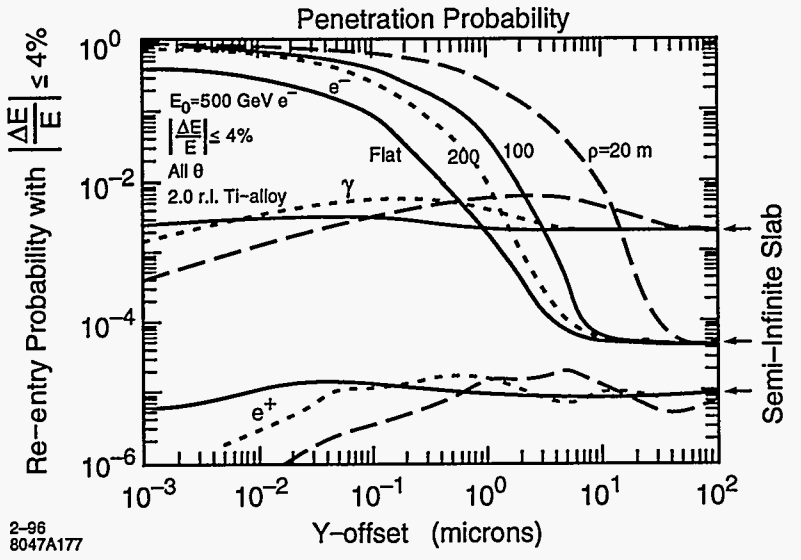


Figure 9-16. *The fraction of particles that re-enter the beam as a function of distance from the edge of a flat and curved (200-m radius of curvature) titanium scraper.*

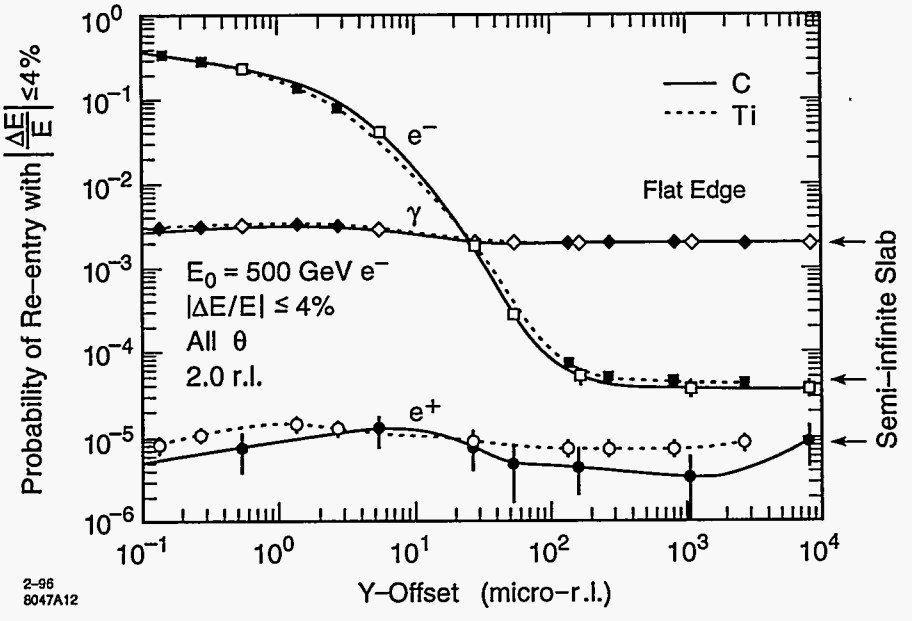


Figure 9-17. *The fraction of particles that re-enter the beam as a function of distance from the edge of a curved (200-m radius of curvature) titanium and carbon collimator. Abscissa is in micro-radiation lengths.*

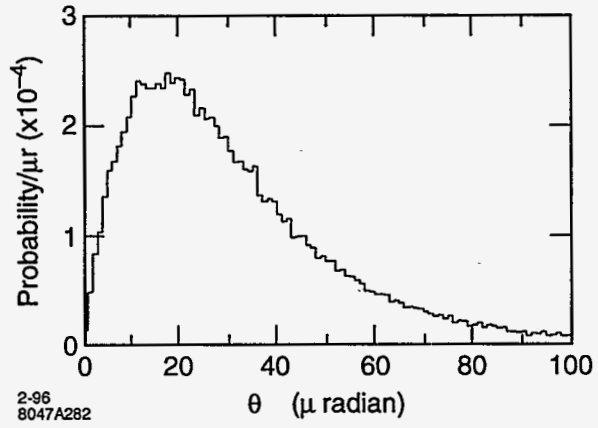


Figure 9-18. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the cone angle measured from the initial particle direction. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

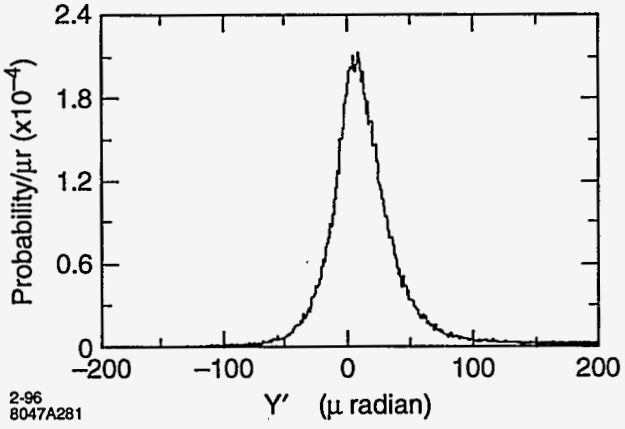


Figure 9-19. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber surface. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

that the shape is still Gaussian. A quick calculation shows that the rms angle of this distribution is achieved in a distance of $20\,\mu\text{r} \times 100\,\text{m} = 2\,\text{mm}$, which is about one fifth of a radiation length for Cu, so indeed the particles can be headed back toward the plane if they are close enough to the edge that they pass through two-fifths of a radiation length or less in the collimator. We note however that there is a mean of this distribution away from the surface of about $12\,\mu\text{r}$.

Figure 9-20 shows the angular distribution when the angle is chosen to be the angle from a plane containing the line of flight of the particles and perpendicular to the surface of the absorber. If the absorber where a vertical collimator, then this angle is the horizontal angle. This distribution should be symmetric because there is no physical way to distinguish left from right.

Figure 9-21 shows the energy distribution for the particles which re-enter the beam. The particles with energy loss less than 4% is similar to Figure 9-14. The distribution for small energy losses fits a power law

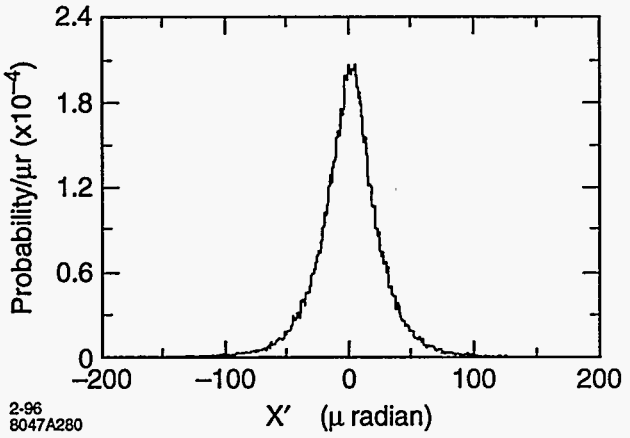


Figure 9-20. *The angular distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from a plane which is perpendicular to the absorber surface containing the particle line-of-flight. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

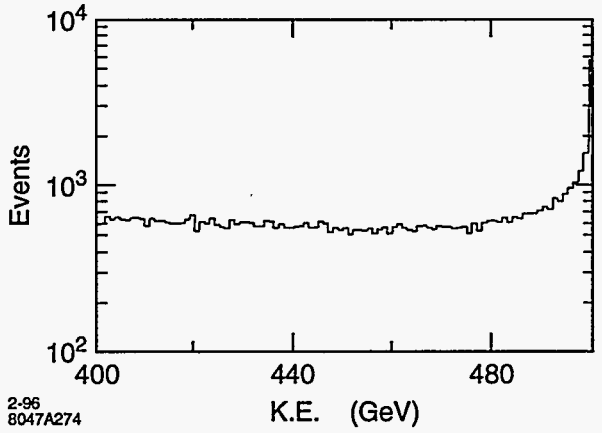


Figure 9-21. *The energy distribution of particles that re-enter the beam (with less than 20% energy loss) for a Cu collimator with a 100-m radius of curvature. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

as in Eq. 9.9 with $t_{eff} = 0.4$. For energy losses greater than 4% the distribution is quite flat. A number which will be important for us is the total number of particles that re-enter the beam. For a uniform density distribution of 1 particle per micron incident on the edge of this 100-m curved copper absorber there are a total of 0.1 particles re-entering the beam with an energy loss of less than 20%. Of this number 30% (or 0.03 particles) have an energy loss less than 4%, and 0.07 have energy loss between 4% and 20%.

Figure 9-22 shows the distribution of Figure 9-19 for four energy bins: the top 1-GeV bin (0.2%), the next 1-GeV bin, a 1-GeV bin with energy 2% below the maximum energy, and a 1-GeV bin with energy 10% below the maximum. There is a tendency for the larger energy loss particles to have a wider angular distribution. Figure 9-22b contains the same distributions as shown in Figure 9-22a but with a radius of curvature equal to 20 m rather than 100 m.

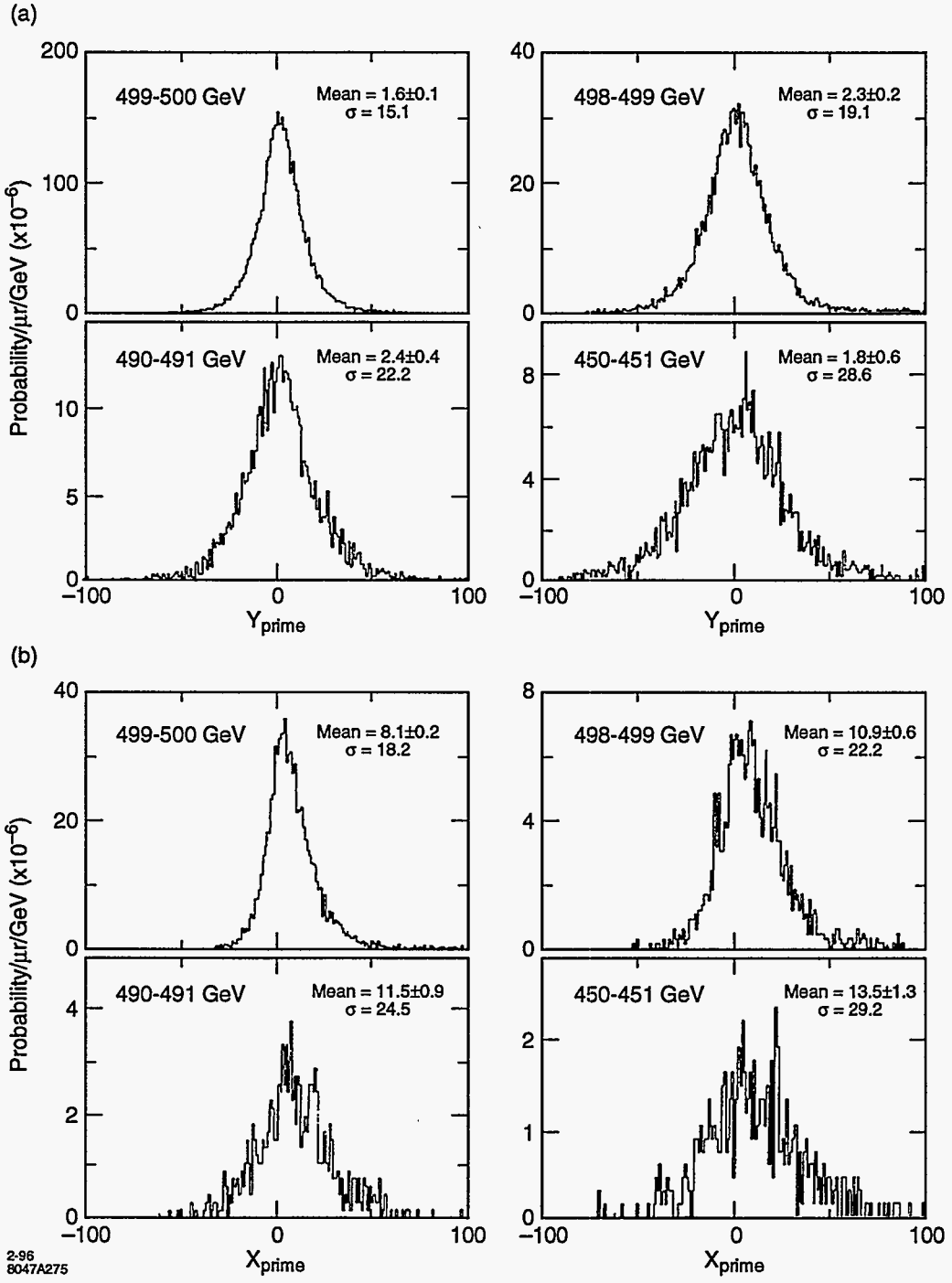


Figure 9-22. a. *The angular distribution of particles that re-enter the beam for four different energy losses for a Cu collimator with a 100-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber. The incident beam is uniformly distributed at the edge for 0 to 10 cm.* b. *The angular distribution of particles that re-enter the beam for four different energy losses for a Cu collimator with a 20-m radius of curvature. The angle is taken to be the angle measured from the plane of the absorber. The incident beam is uniformly distributed at the edge for 0 to 10 cm.*

From the distributions of this section we can arrive at the following conclusions:

- It is primarily particles very close to the edge ($\leq 0.3\,\mu$m) that are scattered back into the beam, so the distributions will not depend on the incoming particle distribution (as long as it is constant over a 1-μm distance). The total number re-entering the beam will depend only on the number of particles per μm incident on the absorber at its edge.
- The angular distributions are very similar to the angular distribution of particles transmitted through a 0.5 r.l. spoiler. The tails of the distribution will be spread due to low energy particles transmitted with larger angular kicks.
- The energy distribution, for particles with less than 4% energy loss, follows a power law distribution characterized by an effective thickness of 0.4 r.l.
- The total number of particles with energy loss less than 20% that reenter the beam is 0.1 times the number of incident particles per micron. The total number of particles with energy loss less than 4% is 0.03 times the number of incident particles per micron.

These distributions can be used to track re-scattered particles through the beam line, and the results can be used to estimate the number of particles that will be present in the particle tails as the beam travels along the beam line from the end of the linac to the IP.

Same-Section Absorbers. After the spoilers there are absorbers in the beam line to actually absorb the bulk of the energy of the tail particles. One percent of the average beam power at 1-TeV c.m. is 84 kW, so this function is substantial.

We have argued above that we must have $\sqrt{(R_{12}R_{34})} \geq 160\,\text{m}$ (or the equivalent when allowing for energy spread) so that the beam particles that have passed through the spoiler are spread over an area with $\sqrt{(\sigma_x\sigma_y)} \geq 2.2\,\text{mm}$. This implies that at the edge of the absorber there will be a particle density of about $n_\mu \approx 10^{10}/\,\text{mm} = 10^7/\,\mu\text{m}$.

Since the first spoilers are located at a β maximum, they are adjacent to a quadrupole which is focusing in the plane being collimated. The next quadrupole is invariably defocusing in that plane, so at the absorber, which is located just prior to the second downstream quadrupole, the beam distribution will be elliptical with the major axis in the plane being collimated. In other words, the estimate of $10^7/\,\mu$m can be taken as an upper limit.

In the previous section it was determined that the number of particles re-scattered from a copper absorber with energy loss less than 4% is about 0.03 times the number of particles incident per micron at the edge. Using this fact and the above estimate of $n_\mu = 10^7/\,\mu$m we have the estimate for the number re-scattered (with less than 4% energy loss) as $N_{sc} \leq 0.3\,10^6$.

Of the 10^{10} tail particles coming from the spoiler only 1/2 will impact the first absorber, since 1/2 of the particles have an angle away from the collimated plane.

We have established in Section 9.2.2 that the β functions at the spoiler must satisfy $\sqrt{(\beta_x\beta_y)} \geq 100^2\,\text{m}$. From this we can conclude that the rms angle ($\theta_0 \approx 14\,\mu$r) of particles emerging from the spoiler is much larger than the beam divergence. Indeed we have $\hat{\sigma}_{x'} = \sqrt{(\varepsilon_x/\beta_x)} \leq 25$ nr and $\hat{\sigma}_{y'} = \sqrt{(\varepsilon_y/\beta_y)} \leq 2.5$ nr. [Note: There can be, and often is, a large non-zero $\alpha = -1/2\,d\beta/ds$. The relevant angular spread within the beam is $\sqrt{(\varepsilon/\beta)}$, not the rms of the angular distribution in phase space, $\sigma' = \sqrt{[(1+\alpha^2)(\varepsilon/\beta)]}$. Therefore we have introduced the notation $\hat{\sigma}' = \sqrt{(\varepsilon/\beta)}]$. So $\theta_0/\hat{\sigma}_{x'} \geq 600$ and $\theta_0/\hat{\sigma}_{y'} \geq 6{,}000$. However, in the transport

system from the collimation system to the IP, the β functions are about 50 m, so that beam-stay-clears are also large, approximately 300 σ_x and 3,000 σ_y. Collimation must be inserted somewhere in this beam line, since 10^9 particles can not be collimated in the final focus. Though they are originally in the IP phase, because of beam-line nonlinearities, chromaticity and high order dispersion, they are sure to migrate to other phases because of their extremely large amplitudes. (See the tracking studies in Section 9.2.5. Actually only 3.5 10^8 particles reach the final focus, but of these 10^8 impact elements in the final-focus system, which is 100 times the acceptable number.)

Suppose collimation is introduced so that only 10^6 particles remain in the tails. In Section 9.2.2 we integrated the distribution to find the number of particles with angles less than some small angle θ_1 to be $\Delta N/N = 1/2(\theta_1/\theta_0)^2$. For this ratio to be 2 10^{-4}, we must have $\theta_1/\theta_0 \leq 2\ 10^{-2}$. Hence we have the estimates $\theta_1/\hat{\sigma}_{x'} \sim 10$ and $\theta_1/\hat{\sigma}_{y'} \sim 100$. We do not have to collimate to apertures that are this small, since these particles are at the IP phase, and such small amplitude particles could be transported through the IP to the dump. However we can conclude that after spoiling (and absorbing) the final-doublet phase, we must collimate the IP phase at least to an amplitude that will be transported safely through the IP. According to the dynamic aperture studies of the final-focus system (see Section 9.2.1 and Figure 9-5) this must be less than 45 σ_x and 200 σ_y.

Considerations of this section lead us to conclude that at a minimum we must either collimate the IP phase at least two times, once before the FD-phase collimation (so that the FD-phase collimation can clean up tail particles generated in the IP-phase collimation) and once after the FD-phase collimation, or collimate the FD phase twice and the IP phase once. The latter option seems attractive because the effect of gas scattering from the end of the collimation system to the FD phase of the final focus is smaller (R_{12} and R_{34} are smaller.) However the gas scattering with the IP phase last is acceptable, and it has the advantage that the last IP phase can have a larger aperture. This turns out to be important when we consider machine protection issues that arise from wake effects on badly mis-steered beams.

Though in this chapter we describe an option in which there is a second FD phase as well as a second IP phase collimation, our present opinion is that the second FD phase is unnecessary.

Next-Phase Absorbers. The R_{12} and R_{34} between phases is given by $R_{12} = \sqrt{(\beta_{x1}\beta_{x2})}$ and $R_{34} = \sqrt{(\beta_{y1}\beta_{y2})}$. Between the first IP-phase and FD-phase collimation (assuming they are collimated to the same depth), we have, from our estimates on the minimum size of the β functions, that both R_{12} and R_{34} are greater than 8 km. This implies that the size of the distribution coming from the first IP-phase spoiler will be about $R\theta_0 \approx 10$ cm when reaching the FD collimators, 100 times larger than the 1-mm size at the absorber immediately following the spoiler.

These estimates can be made more precise by writing the distribution from the spoiler as (see Eq. 9.8)

$$\frac{dN}{N} = \frac{1}{2\pi\theta_0^2} e^{-\frac{\theta_x^2}{2\theta_0^2}}\, e^{-\frac{\theta_y^2}{2\theta_0^2}}\, d\theta_x d\theta_y \quad , \tag{9.12}$$

then integrating over one of the planes to find the distribution in the other plane as

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}\,\theta_0} e^{-\frac{\theta_x^2}{2\theta_0^2}}\, d\theta_x \quad . \tag{9.13}$$

For small θ_x the particle density is given by

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}}\,\frac{d\theta_x}{\theta_0} \quad . \tag{9.14}$$

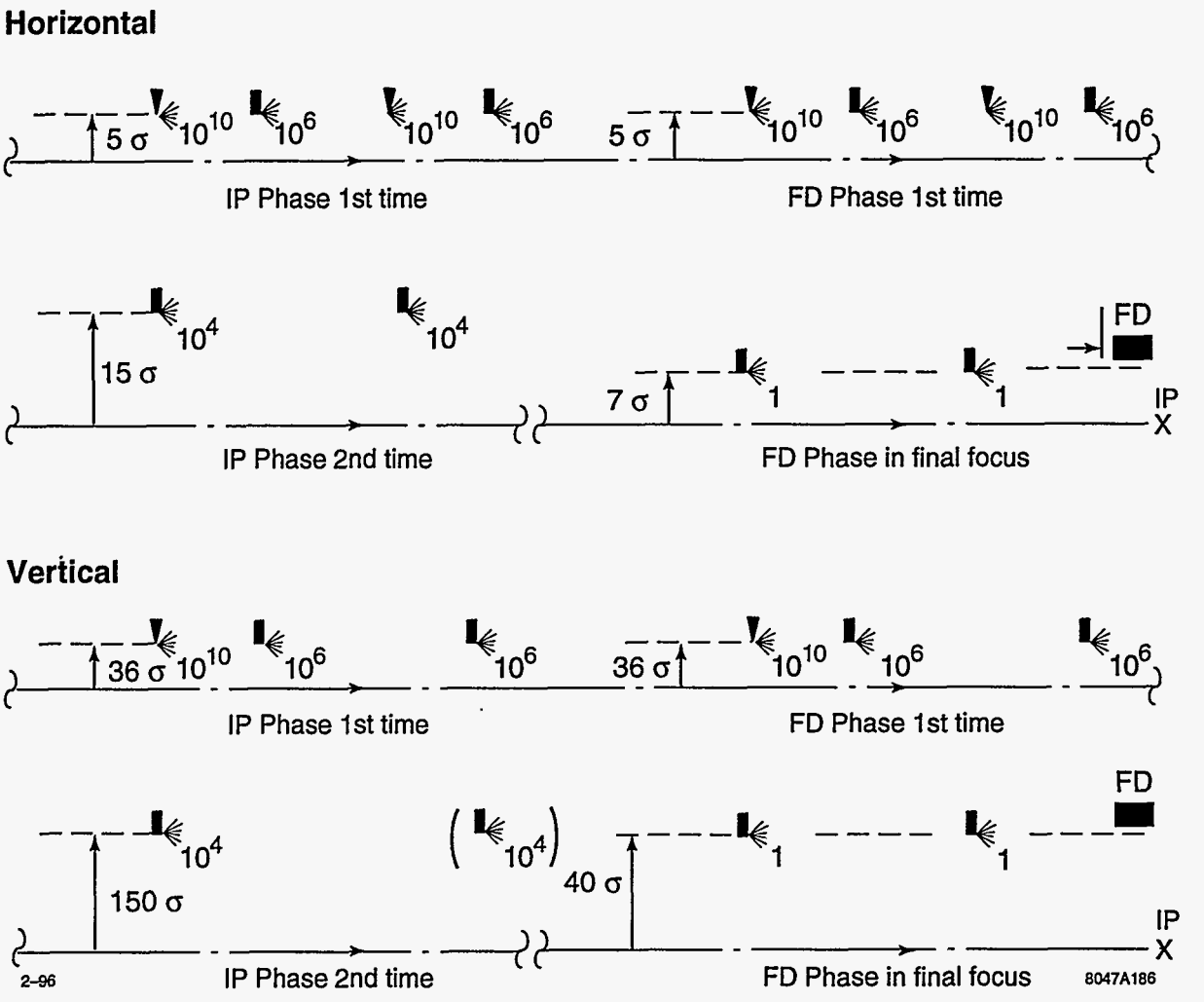


Figure 9-23. *A diagram showing the spoilers and absorbers of a collimation system that contains a single FD phase and two IP phases. The estimated particle numbers emerging from each collimator is shown.*

The distribution of Eq. 9.14 coming from the spoiler implies that the spatial distribution of particles hitting the downstream absorber edge is

$$\frac{dN}{N} = \frac{1}{\sqrt{2\pi}} \frac{dx}{R_{12}\theta_0} = \frac{1}{\sqrt{2\pi}} \frac{n_{1,x}n_{2,x}}{g_1 g_2} \frac{\varepsilon_x}{\theta_0}\, dx \quad . \tag{9.15}$$

We have put the expression in this form because, if the later phases are collimated at a different beam aperture, still the gaps g will be mostly between 1 or 2 mm, whereas β-functions are varied to achieve the appropriate collimator size. If the beam aperture (number of sigma) of the second stage is larger, we see that the estimate for the number of particles at the edge increases linearly with the aperture.

Figure 9-23 shows the spoilers and absorbers of a collimation system having two IP stages and a single FD stage and indicates the re-scattering that is expected at each stage.

Tail Re-population From Gas Scattering

Gas scattering must also be considered as a potential source of particles that will hit downstream collimators. The cross section for Coulomb scattering is given by

$$d\sigma = \left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 d\phi \frac{dq}{q^3} \quad . \tag{9.16}$$

Depending on the nature of the downstream aperture, one can find an expression for the minimum transverse momentum $q_{\min}$ which an electron must receive in order to impact that aperture. If the aperture is round the expression is given by

$$\frac{q_{\min}(s,\phi)}{p_0}\left(R_{12}(s)^2\cos^2\phi + R_{34}(s)^2\sin^2\phi\right)^{\frac{1}{2}} = a \tag{9.17}$$

where p_0 is the beam momentum, a is the aperture radius, and ϕ is the azimuthal angle at which the particle is scattered. If the aperture is a pair of horizontal planes, with half gap g, the expression for $q_{\min}$ will be

$$\frac{q_{\min}(s,\phi)}{p_0} R_{12}(s)\cos\phi = g \quad . \tag{9.18}$$

The cross section for a scattered particle to hit the final doublet is

$$\begin{aligned}
\frac{d\sigma}{d\phi} &= \left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 \int_{q_{\min}(s,\phi)}^{\infty} \frac{dq}{q^3} = \frac{1}{2}\left(\frac{e^2 Z}{2\pi\varepsilon_0 c}\right)^2 \frac{1}{q_{\min}^2(s,\phi)} \\
&= 2\left(\frac{r_e Z}{\gamma a}\right)^2 \left(R_{12}(s)^2\cos^2\phi + R_{34}(s)^2\sin^2\phi\right) \ , \quad \text{circular aperture, or} \\
&= 2\left(\frac{r_e Z}{\gamma g}\right)^2 R_{12}(s)^2\cos^2\phi \quad \text{for a flat aperture.}
\end{aligned} \tag{9.19}$$

This may now be integrated over ϕ. The probability of a scatter in distance ds is given by $\rho\sigma ds$. Thus the number scattered along the beam line which will exceed the aperture limit is given by

$$\begin{aligned}
\frac{\Delta N}{N} &= 2\pi\rho_N\left(\frac{r_e Z}{\gamma a}\right)^2 \int ds\left(R_{12}(s)^2 + R_{34}(s)^2\right) \ , \quad \text{or} \\
&= 2\pi\rho_N\left(\frac{r_e Z}{\gamma g}\right)^2 \int ds R_{12}(s)^2 \quad .
\end{aligned} \tag{9.20}$$

If we assume that the major composition of gas within the beam pipe is a diatomic molecule like N_2, the number of nuclei per unit volume is approximately $\rho_N \approx 6\ 10^{22} P_G m^{-3}$ where P_G is the pressure measured in Torr. If we assume a value of $Z = 7$ and a gas pressure of 10^{-8} Torr, these equations become (for 1-TeV-c.m. energies)

$$\begin{aligned}
\frac{\Delta N}{N} &= 1.5\,10^{-24}\frac{1}{a^2}\int ds\left(R_{12}(s)^2 + R_{34}(s)^2\right) \quad , \quad \text{or} \\
&= 1.5\,10^{-24}\frac{1}{g^2}\int ds R_{12}(s)^2 \quad .
\end{aligned} \tag{9.21}$$

Also of interest is the number density at the edge. This is given by

$$\begin{aligned} \frac{dN}{da} &= 3\,10^{-12}\frac{1}{a^3}\int ds\left(R_{12}(s)^2+R_{34}(s)^2\right) \quad , \qquad \text{or} \\ \frac{dN}{dg} &= 3\,10^{-12}\frac{1}{g^3}\int ds R_{12}(s)^2 \end{aligned} \tag{9.22}$$

where we have set $N = 10^{12}$.

First we can make an estimate of how many particles will hit the absorber in the first FD phase from gas scattered particles in the first IP phase. Since the two beam-line sections are separated by $\pi/2$ in phase we can set $R_{12}^2(s) = \beta_{x1}(s)\beta_{x2}$. We get the estimate

$$\frac{dN}{dg} = 3\,10^{-12}\,\frac{\beta_{x,2}}{g^3}\int ds\beta_{x,1}(s) = 3\,10^{-12}\,\frac{\beta_{x,2}}{g^3}L\,\langle\beta_{x,1}\rangle \tag{9.23}$$

where L is the length of the first phase collimation section. If we take $L = 400\,\text{m}$ (see the lattices in Section 9.2.5), $<\beta>= 1/2\,\beta_{\max}$, and $g = 2\,\text{mm}$, we get $dN/dg = 5\,10^6$. The number per micron would be $n_\mu = 5$, and hence negligible. The total number hitting the collimator is estimated to be $N \approx 2\,10^4$. These numbers are small compared to other backgrounds within the first stage of the collimation system.

The same approach can be used to find an estimate for number of particles that are gas-scattered in the last IP phase of the collimation system (assuming it is an IP phase) that are incident on the collimators of the final-focus system. The $\beta_{x2} = 4\,10^4$ is larger by a factor of about 2, but g is larger by a factor of about 3. Thus estimates are $dN/dg \approx 3\,10^6 (n_\mu = 3)$ and $N \approx 6\,10^3$. These numbers are now not negligible, but they are safely within the number of particles that can be collimated there.

There are two other important estimates to make: the number of tail particles gas scattered in the entire beam line from the collimation system to the final-focus system, and the number of particles incident on the final doublet from particles that are gas scattered within the final telescope of the final-focus system.

To estimate the number from the beam line, we will assume the value of $\sin^2\phi_{12}(s)$ can be taken to be $1/2$. This should be true since the phase advances of these beam-line modules are not locked into the final-focus system phases. We can thus use the same formula derived for the IP-to-FD phase multiplying by a factor of $1/2$. Now $L = 1.3\,\text{km}$, and $<\beta_{1,max}>\approx 40\,\text{m}$. Taking $\beta_{x,2} = 4\,10^4$ and $g = 4\,\text{mm}$, we get the estimates $dN/dg \approx 5\,10^4 (n_\mu = .05)$ and $N \approx 10^2$. These are much smaller than the particles from the last collimation section.

Within the final telescope it is important to use the correct R_{12} function and the round aperture formula. The numbers are still quite acceptable. See Chapter 11, Eq. 11.68.

9.2.4 Wakefield Considerations

Wakefields of Tapered Collimators

The wakefield kick from a spoiler or absorber consists of a geometrical part and a resistive-wall part. To reduce the geometrical wakefield, the spoiler and/or absorber may be tapered as shown in Figure 9-24.

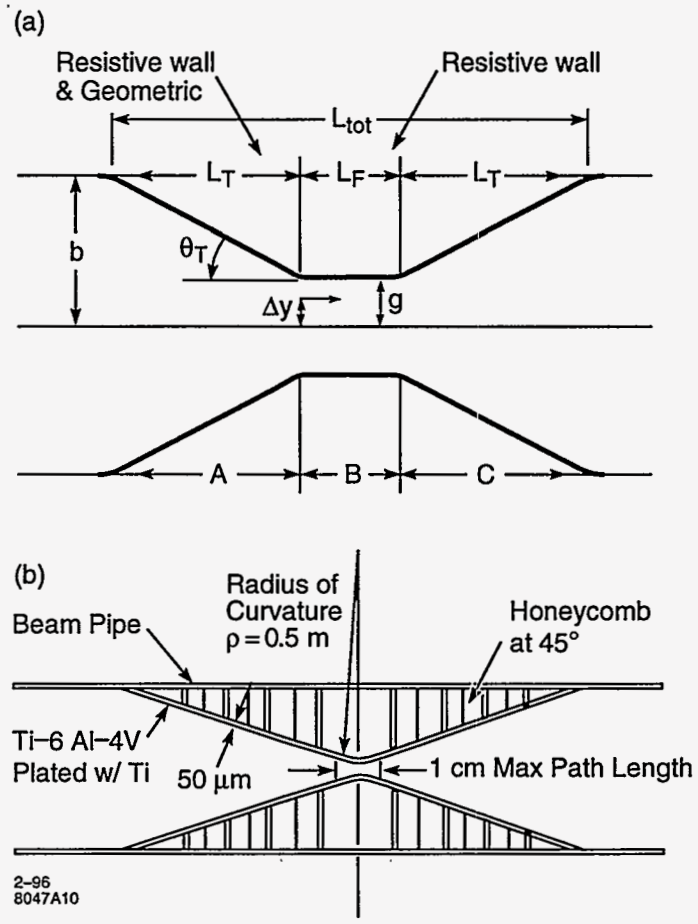


Figure 9-24. a. *Cross-sectional profile of a tapered absorber.* b. *Cross-sectional profile of a tapered spoiler.*

The geometric wake from a tapered spoiler was first studied by K. Yokoya for a cylindrical beam pipe geometry [Yokoya 1988]. The resulting formula may be written as

$$\Delta y_G'^T = A\,[2 I_1 y_0] \quad \text{where} \quad A = \frac{N r_e}{\gamma \sigma_z} f_G(\tau)\,, \quad I_1 = \int \frac{b'^2}{b^2} ds\,, \quad \tau = \frac{z}{\sigma_z} \quad , \text{ and } \quad f_G(\tau) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} \quad . \tag{9.24}$$

Here $b(s)$ is the half-height of the beam pipe as a function of longitudinal position s. y_0 is the displacement of the source particle from the axis. There is no dependence on the position of the test particle for small displacements. The integral I_1 is especially easy to evaluate for a constant b', by taking one factor of b' outside the integral and writing $b'ds = db$ inside the integral. The integral evaluates to $(1/g - 1/b)$, where b is the maximum half-height and g is the collimator half-gap.

The geometric wake for a parallel-plate taper has been studied by G. Stupakov [Stupakov 1996]. The initial result, which is the formula used in many of the computations in this chapter, was

$$\Delta y_G'^T = A\,[1.85\, I_1 y_0 + 1.43\, I_1 y] \tag{9.25}$$

where there is now a dependence on y, the position of the test particle. This term is called the quadrupole wake which is present even for on-axis $y_0 = 0$ source particles. For $y = y_0$ this wake is 64% larger than the cylindrical geometry wake.

The current theoretical result, which was deduced by Stupakov as this text was going to press, may be written

$$\Delta y'^T_G = A\left[(2\pi w I_2 - 2I_1)y_0 + 2I_1 y\right] \quad \text{where} \quad I_2 = \int \frac{b'^2}{b^3} ds \quad . \tag{9.26}$$

Now a dependence on w, the width of the collimator, is present. For a constant b', the integral I_2 evaluates to $1/2(1/g^2 - 1/b^2)$. For $y = y_0$ the I_1 terms cancel, and the dependence on gap width changes to $1/g^2$. The resulting wake is larger than the one used in this chapter by a factor of $(\pi/3.82)(w/g)$, which for $w = 5g$ is about 4.

Since the result of the last paragraph is new and requires confirmation, we have decided to use a modified cylindrical collimator wake in the following sections. For the near-wall and quadrupole-wake effects we will use the original parallel-plate results. This is unfinished and somewhat unsatisfactory, but is a compromise which allows us to illustrate the range of physical effects that are important.

Evaluating the integral I_1 and inserting $b' = (b-g)/L_T$ where L_T is the length of the taper, the taper will have the wake equation

$$\Delta y'^T_G(\tau) = \alpha_G \frac{Nr_e}{\gamma\sigma_z} \frac{2(b-g)^2}{bgL_T} f_G(\tau)\Delta y \qquad \text{where}$$
$$\tau = \frac{z}{\sigma_z} \qquad \text{and} \qquad f_G(\tau) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} \quad . \tag{9.27}$$

α_G is a geometric factor that is unity for a cylindrical beam pipe (Eq. 9.24), equal $3.28/2 = 1.64$ for parallel plate geometry if Eq. 9.25 is valid, and under conditions where w was assumed to scale with g (large g case), would be $(\pi/2)\, w/g$ where Eq. 9.26 holds.

To get the resistive-wall kick from a taper we have integrated parallel-plate resistive-wall wake formula [Chao 1992].

$$\Delta y'^T_R(\tau) = \alpha_R \frac{Nr_e}{\gamma\sigma_z} \frac{2}{r^3(s)} ds\sqrt{\lambda\sigma_z} f_R(\tau)\Delta y \qquad \text{where}$$
$$\tau = \frac{z}{\sigma_z} \qquad \text{and} \qquad f_R(\tau) = \frac{\sqrt{2}}{\pi} \int_0^\infty \frac{d\tau'}{\sqrt{\tau'}} e^{-\frac{(\tau'+\tau)^2}{2}} \quad . \tag{9.28}$$

The result is

$$\Delta y'^T_R(\tau) = \alpha_R \frac{Nr_e}{\gamma\sigma_z} \frac{(b+g)}{b^2 g^2} L_T \sqrt{\lambda\sigma_z} f_R(\tau)\Delta y \quad . \tag{9.29}$$

The function f_R is shown in Figure 9-25. α_R is a geometric factor that is 1 for a cylindrical beam pipe and equals $\pi^2/8$ for parallel plate geometry. Note: This formula is also somewhat uncertain, and SLC measurements are indicating a larger wake [SLC].

$\lambda \equiv \rho/(120\pi)$ is referred to as the resistive depth (ρ is the resistivity in Ω-m). This formula can be taken as valid over the entire bunch if $\sqrt{(\sigma_z\lambda)} \ll g \ll (\sigma_z/\lambda)\sqrt{(\sigma_z\lambda)}$. These inequalities are true for all combinations of λ, σ_z, and g we will consider. For Cu, with $\lambda = 0.045\,\text{nm}$, this translates to $.08\,\mu\text{m} \ll g \ll 0.2\,\text{m}$, and for Ti, with $\lambda = 1.2\,\text{nm}$, this becomes $4\,\mu\text{m} \ll g \ll 0.4\,\text{m}$.

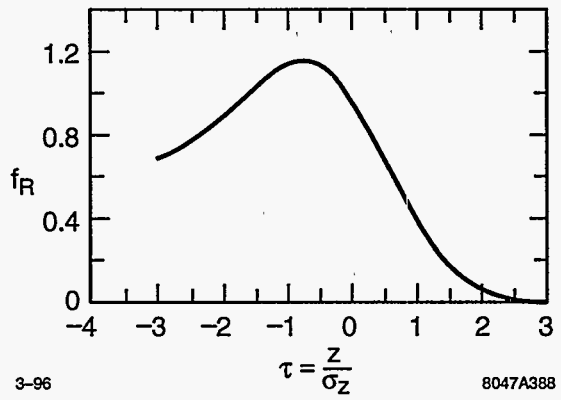


Figure 9-25. *The resistive-wall wake strength as a function of longitudinal position within the bunch.*

We will need averages over the bunch length of f_R, f_R^2, f_G, f_G^2, and $f_G f_R$, whereby the average of any function, h, over the bunch length we mean

$$\langle h \rangle \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\tau' \, e^{-\frac{\tau'^2}{2}} h(\tau') \tag{9.30}$$

These averages are $< f_G >= .282$, $< f_G^2 >= .092$, $\sigma_{f_G} = .111$, $< f_R >= .816$, $< f_R^2 >= .774$, $\sigma_{f_R} = .330$, and $< f_G f_R >= .25$ where by σ_h is meant: $\sigma_h \equiv \sqrt{(< h^2 > - < h >^2)}$. Note that both for f_G and f_R, $\sigma_f / < f > \approx 0.4$. This suggests that emittance growth effects will be smaller than centroid change effects. We also note for later reference that for both f_R and f_G the maximum values $f_{\max} = 1.4 < f >$.

To find the taper length L_T for which the total kick to the centroid is a minimum we must find the minimum of

$$\frac{A \langle f_G \rangle}{L_T} + B \langle f_R \rangle L_T \tag{9.31}$$

which occurs at $L_T = \sqrt{(A < f_G > /(B < f_R >))} = .58\sqrt{(A/B)}$ where both kicks are equal, and the total centroid kick is $< \Delta y'^T >= 2\sqrt{(< f_G >< f_R > AB)} \approx \sqrt{(AB)}$. Plugging in the expressions for A and B we have a minimum kick equal to (now for both incoming and outgoing taper)

$$\begin{aligned} \Delta y'^T &= 2\sqrt{\alpha_G \alpha_R} \frac{N r_e}{\gamma \sigma_z} \left[2 \langle f_G \rangle \langle f_R \rangle \frac{(b^2 - g^2)(b-g)}{b^3} \right]^{\frac{1}{2}} \left(\frac{\lambda \sigma_z}{g^6} \right)^{\frac{1}{4}} \Delta y \\ &\approx 2\sqrt{2} \sqrt{\alpha_G \alpha_R} \frac{N r_e}{\gamma \sigma_z} \left(\frac{\lambda \sigma_z}{g^6} \right)^{\frac{1}{4}} \Delta y \qquad \text{for} \qquad g \ll b \end{aligned} \tag{9.32}$$

and

$$\begin{aligned} L_T^{opt} &= (b-g) \left(\frac{2b \langle f_G \rangle}{(b+g) \langle f_R \rangle} \right)^{\frac{1}{2}} \left(\frac{g^2}{\lambda \sigma_z} \right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \\ &\approx 0.8(b-g) \left(\frac{g^2}{\lambda \sigma_z} \right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \qquad \text{for} \qquad g \ll b \; . \end{aligned} \tag{9.33}$$

For the usual case where $g \ll b$ the optimum taper angle is given by

$$\Theta_T^{opt} \approx 1.1 \left(\frac{\lambda \sigma_z}{g^2} \right)^{\frac{1}{4}} \sqrt{\frac{\alpha_G}{\alpha_R}} \tag{9.34}$$

Taking $g = 1\,\mathrm{mm}$, $\sigma_z = 100\,\mu\mathrm{m}$, $\alpha_G = \alpha_R$ and $\lambda = \{0.045, 0.15, 1.2\}\,\mathrm{nm}$ for {Cu, W, Ti} respectively yields $\Theta_T^{opt} = \{9, 12, 21\}\,\mathrm{mr}$. For $b = 5\,\mathrm{mm}$ $L_T^{opt} = \{55, 42, 24\}\,\mathrm{cm}$. The tungsten-rhenium alloy referred to here can be plated with copper to improve its surface conductivity. The titanium, because it will be used as a spoiler, cannot be plated, except with a material that has survival characteristics as good or better than itself.

The emittance of the beam is also enlarged by the kicks. Since these kicks are in one direction in phase space, they cause the normalized phase-space distribution to depart from a circle into something resembling an ellipse. We will introduce the notation that the IP- and FD-phase axis lengths of this ellipse are $r_{y,\mathrm{IP}}$ and $r_{y,\mathrm{FD}}$. For kicks that are small compared to the original radius of this distribution we would have the estimate for the final radius: $\tilde{r}_{y,\mathrm{IP}}^2 = r_{y,\mathrm{IP}}^2 + \langle \Delta r_{y,\mathrm{IP}}{}^2 \rangle$, or equivalently

$$\tilde{r}_{y,\mathrm{IP}} \approx r_{y,\mathrm{IP}} \left(1 + \frac{1}{2} \frac{\langle \Delta r_{y,\mathrm{IP}}{}^2 \rangle}{r_{y,\mathrm{IP}}^2} \right) \quad . \tag{9.35}$$

The ratio in the parenthesis way be determined from

$$\frac{\langle \Delta r_{y,\mathrm{IP}}{}^2 \rangle}{r_{y,\mathrm{IP}}^2} = \frac{\langle \Delta y'^2 \rangle - \langle \Delta y' \rangle^2}{\hat{\sigma}_{y'}^2} \quad . \tag{9.36}$$

Hence to minimize the emittance growth we must minimize

$$\begin{aligned} &\left\langle \left(\frac{A f_G}{L_T} + B f_R L_T \right)^2 \right\rangle - \left\langle \left(\frac{A f_G}{L_T} + B f_R L_T \right) \right\rangle^2 \\ &= \frac{A^2 \sigma_{fG}^2}{L_T^2} + B^2 \sigma_{fR}^2 L_T^2 + 2AB \left(\langle f_G f_R \rangle - \langle f_G \rangle \langle f_R \rangle \right) \end{aligned} \tag{9.37}$$

The minimum emittance growth occurs at $L_T = \sqrt{(A\sigma_{fG}/(B\sigma_{fR}))} \approx 0.53\sqrt{(A/B)}$, (compare $0.58\sqrt{(A/B)}$ for the minimum centroid kick), and the emittance growth at this minimum is given by

$$\frac{\langle \Delta r_{y,\mathrm{IP}}^2 \rangle}{r_{y,\mathrm{IP}}^2} = 2 \left(\sigma_{f_g} \sigma_{f_R} + \langle f_G f_R \rangle - \langle f_G \rangle \langle f_R \rangle \right) \frac{AB}{\sigma_{y'}^2} = 0.11 \frac{AB}{\sigma_{y'}^2} \approx 0.11 \frac{\langle \Delta y_T' \rangle^2}{\sigma_{y'}^2} \quad . \tag{9.38}$$

For a resistive-wall kick or geometric kick alone we would have

$$\frac{\langle \Delta r_{y,\mathrm{IP}}{}^2 \rangle}{r_{y,\mathrm{IP}}^2} = \frac{\sigma_{\Delta y_R'}^2}{\sigma_{y'}^2} = 0.16 \frac{\langle \Delta y_T' \rangle^2}{\sigma_{y'}^2} \quad . \tag{9.39}$$

Since $\frac{\Delta \sigma_{\mathrm{IP}}}{\sigma_{\mathrm{IP}}} = \frac{\Delta r_{y,\mathrm{IP}}}{r_{y,\mathrm{IP}}} = \frac{1}{2} \frac{\langle \Delta r_{y,\mathrm{IP}}^2 \rangle}{r_{y,\mathrm{IP}}^2}$, the luminosity loss from the taper emittance growth is

$$\frac{\Delta L}{L} = -\frac{\Delta \sigma_{\mathrm{IP}}}{\sigma_{\mathrm{IP}}} = 0.055 \frac{\langle \Delta y_T' \rangle^2}{\sigma_{y'}^2} \quad . \tag{9.40}$$

The luminosity loss from centroid kicks can be derived from the expression $(\Delta y_R - \Delta y_L)^2/(4\sigma^2)$. Taking the right and left kick to be statistically independent, this becomes $\Delta y_R^2/(2\sigma^2)$. Disruption, even at 1/2 design current, allows for twice the mis-steering for the same luminosity loss. Hence with disruption we have the estimate:

$$\frac{\Delta L}{L} = -\frac{1}{8} \frac{\Delta y_R^2}{\sigma_{\mathrm{IP}}^2} = -\frac{1}{8} \frac{\langle \Delta y_T' \rangle^2}{\sigma_{y'}^2} \quad . \tag{9.41}$$

We see that for an optimized taper the emittance growth effect is 2.3 times smaller than the centroid kick effect, even when including disruption. In the horizontal plane, without disruption, it is 9.2 times smaller.

We are expecting an incoming vertical jitter from the linac of 0.2 σ. If we assume an $\frac{\langle \Delta y'_T \rangle}{\sigma_{y'}} \leq 0.7$ for the kicks from both the collimation and final-focus system, the luminosity loss from additional jitter would be about $(1/8)[(0.2)(0.7)]^2 = 0.25\%$, and the luminosity loss due to emittance growth would be 0.1%.

System Wake Equation

If the beam-centroid displacement Δy is due to beam jitter, then the wakefield $\Delta y'$ contributes to beam jitter 90° out of phase from and proportional to the source. The kick from the collimator 'i' will give a kick to the beam, which measured in terms of $\hat{\sigma}_{y'} = \sqrt{\varepsilon_y/\beta_y}$ is

$$t'_i = \frac{\Delta y'^T_i + \Delta y'^F_i}{\hat{\sigma}_{y',i}} = 2\alpha_R \frac{N r_e}{\gamma \sigma_z} \left[\sqrt{2} \frac{(\lambda_i \sigma_z)^{1/4}}{g_i^{3/2}} \sqrt{\frac{\alpha_G}{\alpha_R}} + \langle f_R \rangle \frac{L_{F,i} (\lambda_i \sigma_z)^{1/2}}{g_i^3} \right] \frac{\Delta y_i}{\hat{\sigma}_{y',i}} \quad . \tag{9.42}$$

If we write the incoming jitter $\Delta y_i = t_i \sigma_{y,i}$ and use $\beta_{y,i} = \sigma_{y,i}/\hat{\sigma}_{y',i}$ then the above equation contains a dependence on the β function at the collimator, the collimator gap, and two parameters which depend on the material and dimensions of the collimator, λ and L_F. The β function can be written in terms of the gap and the number of σ collimated, through $\beta_y = g^2/(\varepsilon_y n_y^2)$, to obtain

$$\frac{t'_i}{t_i} = 2\alpha_R \frac{N r_e}{\varepsilon_y^N \sigma_z} \left[\sqrt{2}\, g_i^{1/2} (\lambda_i \sigma_z)^{1/4} \sqrt{\frac{\alpha_G}{\alpha_R}} + \langle f_R \rangle \frac{L_{F,i} (\lambda_i \sigma_z)^{1/2}}{g_i} \right] \frac{1}{n_{y,i}^2} \quad . \tag{9.43}$$

There is an optimal gap, $g = g_0$, which minimizes the bracket in the above equation,

$$g_{0i} = 0.54 (\lambda_i \sigma_z)^{1/6} L_{F,i}^{2/3} (\alpha_R/\alpha_G)^{1/3} \tag{9.44}$$

which depends only on σ_z and the collimator parameters. If the gap is taken to be optimal, then $t'_i = (\overline{n}_i/n_i)^2 t_i$ where $\overline{n}_i$ is equal to

$$\overline{n}_{y,i}^2 \equiv 2\alpha_R \frac{N r_e}{\varepsilon_y^N \sigma_z} [\ldots]_{opt} \approx 3.2 \frac{N r_e}{\varepsilon_y^N} \frac{(\lambda_i L_{F,i})^{1/3}}{\sigma_z^{2/3}} \alpha_G^{1/3} \alpha_R^{2/3} \tag{9.45}$$

When the collimation aperture is set to $n_i = \overline{n}_i$ and g is chosen optimally, then the ratio $t'/t = 1$. In terms of the parameters g_{0i} and $\overline{n}_i$, Eq. 9.43 can be written

$$\frac{t'_i}{t_i} = \frac{\overline{n}_{y,i}^2}{n_{y,i}^2} \left[\frac{2}{3} \sqrt{\frac{g_i}{g_{oi}}} + \frac{1}{3} \left(\frac{g_{oi}}{g_i} \right) \right] = \frac{\overline{n}_{y,i}^2}{n_{y,i}^2} \phi_i \tag{9.46}$$

For convenience we have introduced the function

$$\phi(x) \equiv \frac{1}{3} \left(2\sqrt{x} + \frac{1}{x} \right) \quad \text{and defined} \quad \phi_i \equiv \phi \left(\frac{g}{g_{0i}} \right) \quad . \tag{9.47}$$

It has the values $\phi(1) = 1$, $\phi(2) = 1.1$, $\phi(4) = 1.42$, $\phi(8) = 1.93$. We see that g can be a factor of 8 larger than g_0 before the kick from a collimator has doubled.

Equation 9.46 has the advantage that the parameters have an intuitive meaning and in addition the dependence on g will be quite weak if it is chosen near g_{0i}. The optimal gaps g_{0i} and the $\overline{n}_i$ parameters

Attribute	Symbol	Titanium Plated Titanium Alloy Spoiler (1/4 r.l.)	Copper Absorber (20 r.l.)	Copper-plated Tungsten-Rhenium Absorber (20 r.l.)
Resistive Depth	$\lambda = \rho/(120\pi)$ (nm)	1.2	0.045	0.045
Flat length	L_F (cm)	1.0	30.	7.0
Optimum gap	g_o (mm)	0.23	1.3	0.5
Horizontal $\overline{n}$	$\overline{n}_x$	1.42	1.45	1.14
Vertical $\overline{n}$	$\overline{n}_y$	11.2	11.5	9.0

Table 9-4. *Important parameters for spoilers and absorbers. The parameter depends on the emittance, bunch length and particle number of a bunch. This table uses the values $\alpha_R = \alpha_G = \pi 2/8$. π varies as $\alpha_G^{1/6}$.*

for the three types of collimators we will use are shown in Table 9-4. For the titanium spoiler, we take the length of the curved section at its apex ($L_F = 1$ cm) to be the appropriate length of the flat section.

The half-gaps, g, cannot be too large, or else the apertures of nearby beam-line elements become uncomfortably large, and they can not be too small, or else the collimator becomes impossible to control, align and adjust. It is remarkable that the optimum g, depending only on properties of materials and the bunch length, are actually reasonable values for collimator gaps.

Now we can write the equation for a system of collimators. The kicks from each collimator at any particular phase will add monotonically. If we assume that throughout the system the jitter in the phase being collimated does not increase appreciably, then setting $t' = \Sigma t_i'$ and all $t_i = t$ we have the result

$$\frac{t^i}{t} = \sum_i \left(\frac{t_i'}{t_i}\right) = \sum_i \left(\frac{\overline{n}_i}{n_{y,i}}\right)^2 \phi_i \quad , \tag{9.48}$$

where $n_{y,i}$ is the setting of the ith collimator. Because of machine protection and off-axis wake requirements to be discussed below, the collimator apertures $n_{y,i}$ will increase through the system. Since the resultant jitter is perpendicular in phase to the source jitter, statistically jitter from the collimation system will add in quadrature to the incoming jitter. A ratio of $t'/t = 0.7$, which would statistically produce a 25% increase in average beam jitter, is taken as an upper limit on acceptable jitter amplification.

Equation 9.48 must be solved iteratively for the required $n_{y,i}$ because the β functions may not be chosen freely to optimize the gap. The first iteration consists of making a list of all collimators that are planned for the system, taking $g = 1$ mm for all gaps, and assuming all $n_{y,i}$ are equal. Equation 9.48 can then be solved for a first guess n_y. With this value of n_y and an estimate of the β functions required at each collimator, a table of estimated gap sizes can be calculated. Any anticipated growth in n_y through the system can also be included. Now Eq. 9.48 can be used to obtain the values of t'/t for this n_y estimate. An iteration in the $n_{y,i}$ may be required to exactly meet the system amplification (t'/t) budget. This procedure is carried out in Section 9.2.5, and the results are listed in Table 9-5.

Quadrupole and Large Amplitude Wakes of Parallel Plate Collimators

Large Amplitude Resistive-Wall Wakes. The potential for the resistive-wall wake of parallel-plate collimators is given by [Piwinski]

$$V(x,y;x_0,y_0) = -\kappa f_R(\tau)\left[\frac{-x_-\sinh x_- + y_+\sin y_+}{\cosh x_- + \cos y_+} + \frac{x_-\sinh x_- + y_-\sin y_-}{\cosh x_- + \cos y_-}\right]$$
$$\text{where} \quad y_+ = \frac{\pi}{2g}(y+y_0), \qquad y_- = \frac{\pi}{2g}(y-y_0), \qquad x_- = \frac{\pi}{2g}(x-x_0), \tag{9.49}$$
$$\text{and} \quad \kappa = \frac{1}{2}\frac{Nr_e}{\gamma\sigma_z}\frac{L}{g}\sqrt{\lambda\sigma_x} \quad .$$

Here x and y are the horizontal and vertical position of the particle experiencing the wake, and x_0 and y_0 are the coordinates of the source of the wake. The collimator planes are set to collimate in y, and are separated by $2g$.

There are several limits of interest. One important limit ignores the difference between the horizontal position of the source and the test particles. If we are only interested in the vertical kicks we can set $x_- = 0$. Then

$$V(x_0,y;x_0,y_0) = -\kappa f_R(\tau)\left[\frac{y_+\sin y_+}{1+\cos y_+} + \frac{y_-\sin y_-}{1-\cos y_-}\right] \quad . \tag{9.50}$$

We can take the derivative of this equation to get the kick anywhere in the beam pipe

$$\Delta y' = \frac{\pi}{2}\frac{\kappa f_R}{g}\left[\frac{y_+ + \sin y_+}{1+\cos y_+} - \frac{y_- - \sin y_-}{1-\cos y_-}\right] \tag{9.51}$$

The kick received by particles in the core of the beam can be found by setting $y = y_0$ in Eq. 9.51.

$$\Delta y' = \frac{\pi}{2}\frac{\kappa f_R}{g}\left[\frac{\sin\hat{y}+\hat{y}}{1+\cos\hat{y}}\right] \qquad \text{where} \qquad \hat{y} = \pi\frac{y_0}{g} \tag{9.52}$$

The wake kick for a mis-steered beam corresponding to Eq. 9.52 is shown in Figure 9-26b.

For small y and y_0 we can expand V in a power series

$$V(x_0,y;x_0,y_0) \approx -\frac{\pi^2}{3}\frac{\kappa f_R}{g^2}\left[y_0 y + \frac{1}{4}y^2 + \ldots\right] \quad , \tag{9.53}$$

which produces the kick

$$\Delta y' \approx \frac{\pi^2}{3}\frac{\kappa f_R}{g^2}\left[y_0 + \frac{1}{2}y + \ldots\right] \quad . \tag{9.54}$$

For particles ($y = y_0$ in the beam)

$$\Delta y' \approx \frac{\pi^2}{2}\frac{\kappa f_R}{g}\frac{y}{g} \quad . \tag{9.55}$$

The coefficient of y/g in this equation would be the extrapolation to the wall of the small amplitude kick. We define this kick at the wall to be

$$\Delta y'_{ref} \equiv \frac{\pi^2}{2}\frac{\kappa f_R}{g} \tag{9.56}$$

For the flat part of a 20 r.l. Cu absorber with $g = 1.8\,\text{mm}$, $N = 10^{10}$, $\sigma_z = 125\,\mu\text{m}$, and $\gamma = 10^6$, $< \Delta y'_{ref} >= 11$ nr. For the flat part of a 20 r.l. copper -plated tungsten-rhenium absorber, $< \Delta y'_{ref} >= 2.6$ nr.

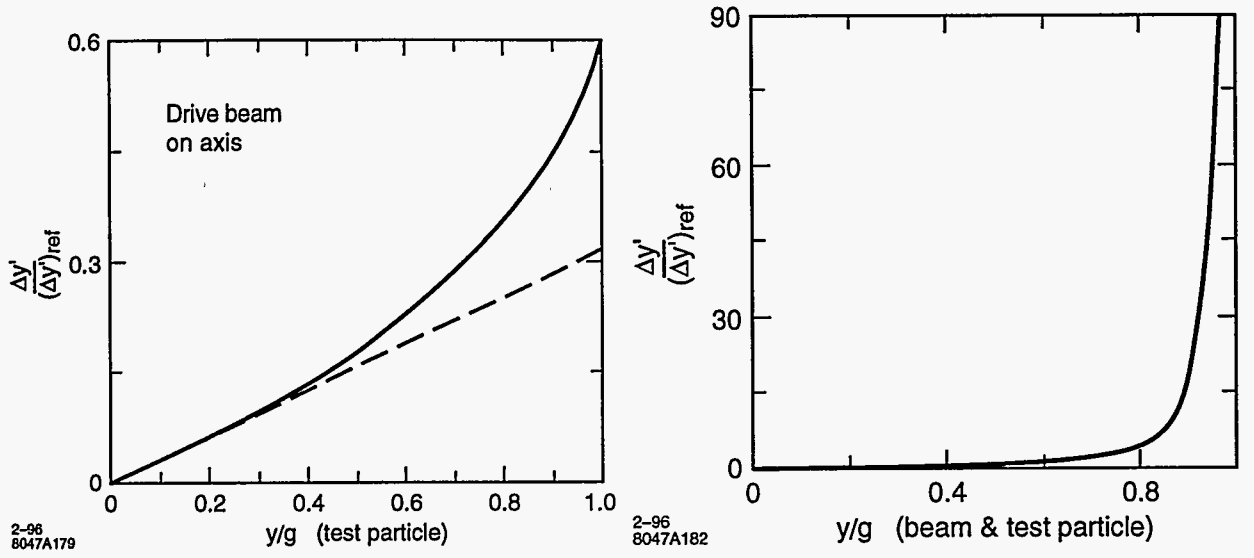


Figure 9-26. *(a) The resistive-wall wake from parallel plate collimators for an on-axis beam. (b) The resistive-wall wake from parallel plate collimators for a mis-steered beam. The wake-kick diverges as the beam approaches the collimator edge.*

Another limit of interest are wakes for on-axis beams. Keeping $x_- = 0$ and setting $y_0 = 0$, we have

$$V(x_0, y; x_0, 0) = \kappa f_R \left[\frac{2\tilde{y}}{\sin\tilde{y}}\right] \quad \text{where} \quad \tilde{y} = \frac{\pi}{2}\frac{y}{g} \quad . \tag{9.57}$$

Taking the derivative of this equation we get the kick

$$\Delta y'_{y_0=0} = \pi\frac{\kappa f_R}{g}\left[\frac{\sin\tilde{y} - \tilde{y}\cos\tilde{y}}{\sin^2\tilde{y}}\right] \quad . \tag{9.58}$$

As $y \to g$, $\tilde{y} \to \pi/2$, hence $\Delta y'_{y_0=0} \to \pi\kappa f_R/g$. The quadrupole term in the multipole expansion of Eq. 9.54 extrapolated to the wall would have given a kick $\Delta y' \to 1/3\Delta y'_{ref}$, which is smaller than the value of $\Delta y'_{y_0=0}$ at the wall by a factor of $\pi/6 \approx 1.9$. Figure 9-26(a) is a graph of the wake for an on-axis beam (Eq. 9.58) from $y = 0$ to $y = g$.

A third limit of interest is when both y and y_0 are close to g. Here it is useful to introduce the variable $\Delta_+ = \pi - y_+$ into Eq. 9.50 to obtain

$$V(x_0, y; x_0, y_0) = -\kappa f_R\left[\frac{\pi\sin\Delta_+}{1-\cos\Delta_+} - \frac{\Delta_+\sin\Delta_+}{1-\cos\Delta_+} + \frac{y_-\sin y_-}{1-\cos 6_-}\right] \xrightarrow{\Delta_+ \to 0} -2\pi\frac{\kappa f_R}{\Delta_+} \quad . \tag{9.59}$$

As $y_0 \to g$ we have the result that

$$\Delta y' \approx \kappa g f_R \frac{1}{(g-y_0)^2} \quad . \tag{9.60}$$

When the entire beam passes near the edge of a collimator the wake kick can be very large. However, as a beam approaches the wall, when $g - y_0 \leq \sigma_x$, the singularity of the potential found in Eq. 9.53 becomes modified by the fact that $x_- \neq 0$. We must return to Eq. 9.49 and expand in x_- , Δ_+, and y_-. The result is

$$V(x, y; x_0, y_0) = -\kappa f_R\left[\frac{\pi\sin\Delta_+}{\frac{x_-^2}{2}+(1-\cos\Delta_+)} - \frac{x_-^2+\Delta_+\sin\Delta_+}{\frac{x_-^2}{2}+(1-\cos\Delta_+)} + \frac{x_-^2+y_-\sin y_-}{\frac{x_-^2}{2}+(1-\cos y_-)}\right]$$

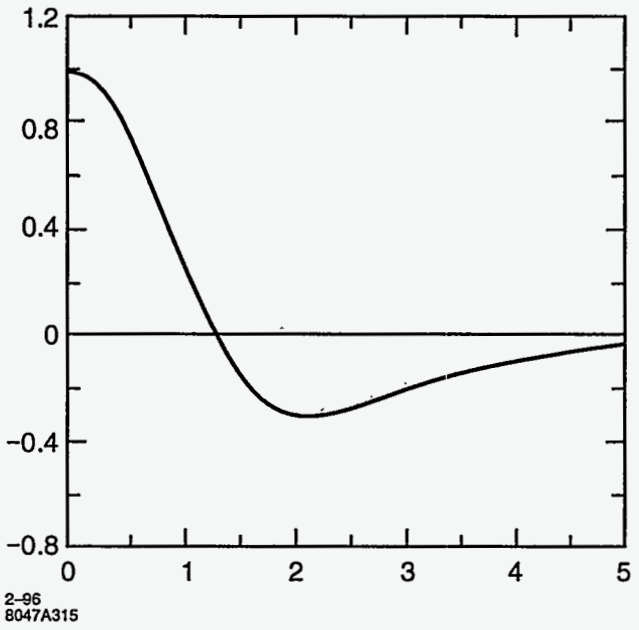


Figure 9-27. *The coefficient of the first power of λ in Eq. 9.65.*

$$\xrightarrow[\Delta_+ \to 0]{} -2\pi\,\kappa f_R \frac{\Delta_+}{(x_-^2 + \Delta_+^2)} \tag{9.61}$$

It is an interesting feature of this potential that for $\Delta_+ > x_-$ the force is toward the wall, while for $\Delta_+ < x_-$, the force is away from the wall. For $x_- = 0$ the force is always toward the wall (positive), and indeed the potential diverges as $1/\Delta_+$ as in Eq. 9.59.

Assuming the beam has a Gaussian shape, we can integrate Eq. 9.61 over the horizontal distribution. We have the result [Stupakov 1995].

$$\lim_{\Delta_+ \to 0} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{dx_0}{\sigma_{x_0}} e^{-\frac{x_0^2}{2\sigma_{x_0}^2}} V(x, y; x_0, y_0) = -\frac{\sqrt{2\pi}\kappa f_R}{\sigma_{x_-}} \left[\pi e^{-\frac{\xi^2}{2}} - \frac{\Delta_+}{\sigma_{x_-}} \sqrt{2\pi}\, h_1(\xi) \right]$$

$$\text{where} \quad \sigma_{x_-} = \frac{\pi}{2g}\sigma_{x_0}, \quad \text{and} \quad \xi = \frac{x}{\sigma_{x_0}} \quad . \tag{9.62}$$

A plot of $h_1(\xi)$ is shown in Figure 9-27. From Eq. 9.62 we may obtain the horizontal and vertical kicks:

$$\begin{aligned} \Delta y'(x, y', y_0) &\approx \frac{4\kappa g f_R}{\sigma_{x_0}^2} h_1(\xi) \\ \Delta x'(x, y', y_0) &\approx -\sqrt{\frac{\pi}{2}} \frac{4\kappa g f_R}{\sigma_{x_0}^2} \xi e^{-\frac{\xi^2}{2}} \quad . \end{aligned} \tag{9.63}$$

Note that these kicks do not depend on the distance to the wall of either the source or the test particle, as would be expected for a plane geometry. Comparing Eq. 9.63 with 9.60 we see, for $\xi = 0$, the kick from Eq. 9.60 limits out at $g - y_0 = \sigma_{x0}/2$.

Geometric Wake Formulae. For the quadrupole term from the geometric wake we will use Eq. 9.24.

When the beam is close enough to the wall that it is appropriate to neglect the contribution of the opposite wall, the expression for the wake is [Stupakov 1996]

$$\Delta y'^{T}_{G}(\tau) = 4A \int \frac{a'^2}{a} ds = 4A\theta \ln\left[\frac{0.4g}{a_g}\right] \tag{9.64}$$

where a is the distance from the beam to the wall as a function of s, and a_g is the nearest approach of the beam to the taper.

If the beam is very close to the wall, the dimension of the beam parallel to the wall be comes important. Then the appropriate formula is

$$\Delta y'^{T}_{G} = 4A \int \frac{aa'^2}{x_2 + a^2} ds = 2A\theta \ln\left[\frac{x^2 + (0.4g)^2}{x^2 + a_g^2}\right] \tag{9.65}$$

where $\theta = b - g/L_T$.

This must now be integrated over the horizontal distribution. This can be performed in the limit of $a_g < \sigma_x$ to give (independent of the exact value of a_g)

$$\Delta y'^{T}_{G} = 2A\theta \ln\left[1.9\left(\frac{0.4g}{\sigma_x}\right)^2\right] \tag{9.66}$$

A table of numerical results for the collimators of the beam delivery lattice is given in Section 9.2.5.

If the Equation 9.26 is valid, the dipole term will be longer than this term, and is the current near-wall limit.

Consequences of Quadrupole and Near-Wall Wakes

Quadrupole-Wake Focusing. We see in Eq. 9.54 that for the resistive-wall wake the quadrupole term is 1/3 of the total wake kick. If the quadrupole wake were written as $\Delta y' = k_R \Delta y$, then $\Delta y'/\delta_{y'} = (k_R\beta)\Delta y/\sigma_y$. Hence we have

$$\text{“}k_{R,i}\beta_i\text{”} = 0.4\,\frac{1}{3}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}} + \frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}} + \frac{g_{0i}}{g_i}}\right] \quad . \tag{9.67}$$

The factor of 1/3 comes from the fact that, for the resistive-wall wake, 1/3rd of the wake comes from the quadrupole wake. t'_i/t_i is the number of σ' from a 1-σ displacement at the ith collimator. The bracket gives the fraction of the wake that is a resistive-wall wake. One-half of the taper wake is resistive-wall. At $g_i = g_{0i}$ the second bracket has the value 2/3 which is just the fraction of the total wake that is resistive-wall at the optimum gap. The value of the bracket lies between 1/2 and 1.

The initial factor of 0.4 comes from the fact that $\sigma_{fR}/ < f > \approx 0.4$. The average quadrupole effect of the wakes will be tuned out when the waist is tuned. It is the spread in focusing that occurs longitudinally along the bunch that is untunable, and will result in a luminosity loss.

As we have seen above in Eq. 9.65, the geometric parallel plate wakefield also has a quadrupole term. In other words, even for on-axis beams there will be a wakefield kick for off-axis particles. For the geometric wake the fraction of the total wake that is a quadrupole wake is 1.43/3.28 = .44. This is somewhat larger than for the resistive-wall wake where this fraction is equal to 0.33. Multiplying this result together with

the fraction of the wake from a collimator that is geometric we have

$$\text{“}k_G\beta_i\text{”} = 0.4\left(\frac{t_i'}{t_i}\right)\left[\frac{.16\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}u}{g_i}}\right] \quad . \tag{9.68}$$

The bracket has a value less than 1/2. Adding the resistive-wall and geometric together we have

$$\text{“}k_{R+G,i}\beta_i\text{”} = 0.4\left(\frac{t_i'}{t_i}\right)\left[\frac{0.77\sqrt{\frac{g_i}{g_{0i}}}+0.33\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \tag{9.69}$$

The value of the bracket lies between 0.39 and 0.33.

We can sum Eq. 9.69 over all collimators, taking the worst case of 0.39 for the bracket to get

$$\text{“}k_{R+G}\beta''_{\text{Total}} \leq (0.4)\left(\frac{t'}{t}\right)[0.39] \leq 0.11 \tag{9.70}$$

The luminosity is decreased by 2% when $k_Q\beta = 0.2$, hence the luminosity loss expected from the quadrupole focusing of collimators is less than 0.5%.

Near-Wall Wakes for On-Axis Beams. For parallel-plate collimators there is a wakefield even when the beam is on-axis. This raises the concern that particles passing near the collimator could be deflected into the tails and cause a problem downstream [Yokoya 1995].

We have shown above that for the resistive-wall wake the value of the wake at the wall for an on-axis beam is 1.9 times the linear extrapolation of the quadrupole wake of the on-axis beam. It is possible, by integrating the resistive-wall wake over the taper, to calculate the wake for a particle passing at the taper minimum. One finds the wake at the taper minimum is 1.4 times the linear extrapolation of the quadrupole wake. Putting these results together we get the following expression for the kick at the wall from the resistive-wall part of the wake:

$$\frac{\Delta y'_{RWX}}{\sigma_{y'}} = \frac{1}{3}n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{1.4\sqrt{\frac{g_i}{g_{0i}}}+1.9\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.71}$$

For the geometric wake we have not formally found an answer, but will assume that it is a factor of 2 times the linear extrapolation of the quadrupole wake. Hence

$$\frac{\Delta y'_{GWX}}{\sigma_{y'}} \cong (0.44)n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{2\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.72}$$

Adding Eq. 9.71 and 9.72 we obtain

$$\frac{\Delta y'_{R+G,WX}}{\sigma_{y'}} = n_{y,i}\left(\frac{t_i'}{t_i}\right)\left[\frac{1.34\sqrt{\frac{g_i}{g_{0i}}}+0.63\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.73}$$

The bracket has a value between 0.67 and 0.63. When we sum over i, it should be over the upstream collimators only. Let us assume that 3/5 of the collimation budget for t'/t is in the first stage. Then Eq. 9.73

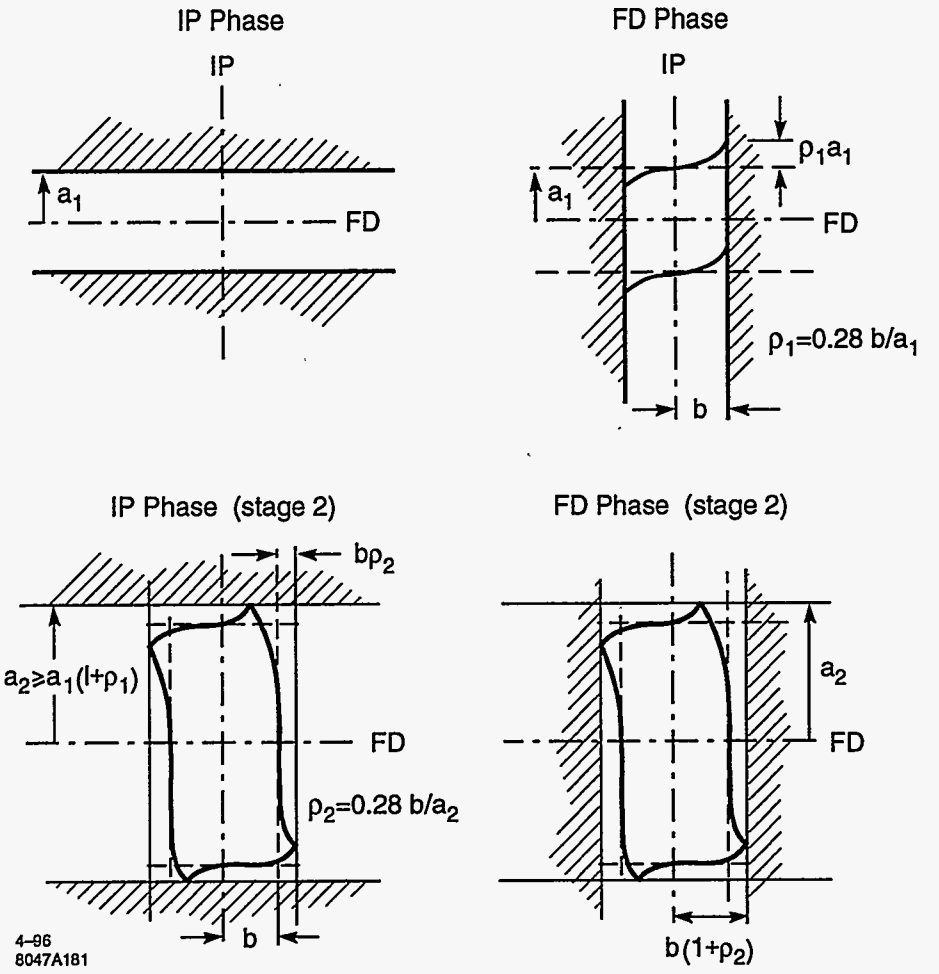


Figure 9-28. *When phase #2 is collimated, the on-axis beam wake which is present for parallel plate collimators, moves particles into a phase space region that was previously collimated in phase #1.*

gives

$$\left.\frac{\Delta y'_{R+G,WX}}{\sigma_{y'}}\right|_{stage1} \leq \frac{3}{5} n_{y1} \left(\frac{t'}{t}\right) [0.67] \leq 0.28\, n_{y1} \tag{9.74}$$

The consequences of near-wall wakes for on-axis beams can be clarified by considering the particle motion in a sequence of normalized phase-space diagrams. See Figure 9-28. Let us suppose that the IP phase is first collimated to some aperture "$a_1\sigma$" and then the FD phase is collimated to an aperture "$b\sigma$". Because of the wake at the wall, the IP boundary will be distorted, as shown in Figure 9-26a, and particles will extend into the previously collimated region by an amount $0.28b\sigma$ (see Eq. 9.74). When the IP phase is collimated again, at a larger aperture "$a_2\sigma$" the distortion will be $0.28a_2(b/a_2)^2\sigma = (0.28b/a_2)b\sigma$. The factor $(b/a_2)^2$ arises from the assumption that b was the appropriate aperture for 3/5 of the 0.7 jitter amplification budget, and the effect of $a_2 > b$ can be determined by using the fact that the kick scales as the aperture squared. In Figure 9-28 we have set $\rho_2 = 0.28\, b/a_2$.

Because of near-wall wake for mis-steered beam, to be discussed in the next section, $a_2/b \approx 3.5$, hence $\rho_2 = 0.28b/a_2 = 0.08$. The stage-2 FD aperture must be enlarged by only 8%: from 36 to 39 in the vertical plane, from 5 to 5.4 in the horizontal plane.

Near-Wall Wakes for Mis-steered Beams. We would like to estimate the near-wall kick. The limit from the resistive-wall part is $(\kappa f_R/g)(2g/\sigma_x)^2 < h_1 > \approx 1/\pi^2 \Delta y'_{ref}(2g/\sigma_x)^2$. Hence we have

$$\frac{\Delta y'_{RWW}}{\sigma_{y'}} = \frac{4}{\pi^2}\left(\frac{g}{\sigma_x}\right)^2 n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.75}$$

The geometric-wake limit can be found from Eq. 9.66, putting $A\theta$ as the extrapolation of the small-amplitude wake to the wall:

$$\frac{\Delta y'_{GWW}}{\sigma_{y'}} = 2\ln\left[1.9\left(\frac{0.4g}{\sigma_x}\right)^2\right] n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{0i}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.76}$$

We have not included any term for the resistive-wall part of the taper. Let us assume that only that length of taper contributes for which the surface is still within $\sigma_x/2$ of the taper minimum, since the wake falls off very sharply beyond that distance. When we integrate the resistive-wall-wake formula over this distance we find that the contribution to the small-amplitude wake for this region as compared to the whole taper is the fraction σ_x/g. Thus we have the result

$$\frac{\Delta y'_{RWW}}{\sigma_{y'}} = \frac{4}{\pi^2}\left(\frac{g}{\sigma_x}\right)^2 n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{\sqrt{\frac{g_i}{g_{0i}}}}{2\sqrt{\frac{g_i}{g_{01}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.77}$$

Combining these results we get (setting $g/\sigma_x = 10$)

$$\frac{\Delta y'_{R+GWW}}{\sigma_{y'}} = n_{y,i}\left(\frac{t'_i}{t_i}\right)\left[\frac{7.4\sqrt{\frac{g_i}{g_{0i}}}+40\frac{g_{0i}}{g_i}}{2\sqrt{\frac{g_i}{g_{01}}}+\frac{g_{0i}}{g_i}}\right] \quad . \tag{9.78}$$

The bracket lies between 1.9 and 40. The resistive-wall contribution is much larger than the taper. If we assume g/g_{0i} is about 2, we get a value for the bracket of about 10. This gives an estimate for the first stage of

$$\left.\frac{\Delta y'_{R+G,WW}}{\sigma_{y'}}\right|_{stage1} \leq \frac{3}{5} n_{y1}\left(\frac{t'}{t}\right) 10 \leq 4.2\, n_{y1} \tag{9.79}$$

For $n = 36$ this gives the estimate for the near-wall kick of $150\sigma'$. The detailed calculation with the exact apertures of the lattice gives $118\sigma'$. A plot of this kick with distance from the wall is shown in Figure 9-29(a). Only the resistive-wall terms, adding up to $90\sigma'$ are included.

While these kicks are large, they are still in the beam pipe and smaller than the dynamic aperture of the final-focus system. Also because of the τ dependence, the spread of these kick is at least $0.4(120)\sigma_{y'} = 48\sigma_{y'}$. Additionally, there is a dependence on x, making the spread about equal to the magnitude of the kick. This can be deduced from the following relationships:

$$< \Delta y' > \;=\; \frac{4\kappa g}{\sigma_{x_0}^2}\langle f_R\rangle\langle h_1(\xi)\rangle = 0.41\frac{4\kappa g}{\sigma_{x_0}^2}$$

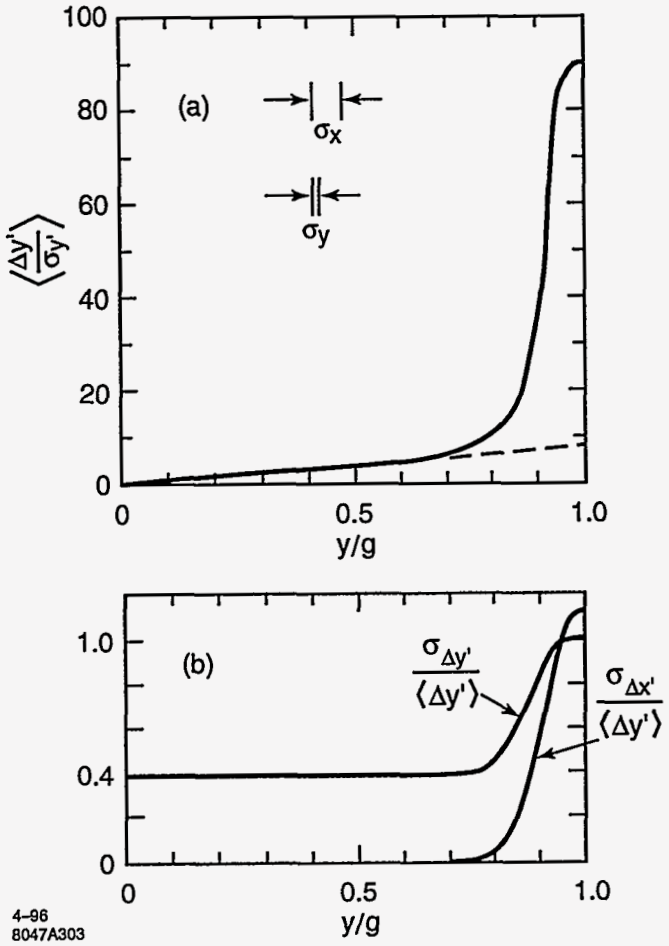


Figure 9-29. *(a) The wake kick as a function of distance to the wall showing the cut-off at the wall when the distance approaches the width of the bunch. (b) An estimate for the variation of the σ of the wake-kicks of (a).*

$$
\begin{aligned}
\sigma_{\Delta y'} &= \frac{4\kappa g}{\sigma_{x_0}^2}\sqrt{\langle f_R^2\rangle\langle h_1(\xi)^2\rangle - \langle f_R\rangle^2\langle h_1(\xi)\rangle^2} = 0.44\,\frac{4\kappa g f_R}{\sigma_{x_0}^2} \\
\langle \Delta x'\rangle &= 0 \\
\sigma_{\Delta x'} &= \sqrt{\frac{\pi}{2}}\,\frac{4\kappa g}{\sigma_{x_0}^2}\sqrt{\langle f_R^2\rangle\langle h_2(\xi)^2\rangle - \langle f_R\rangle^2\langle h_2(\xi)\rangle^2} = 0.49\,\frac{4\kappa g f_R}{\sigma_{x_0}^2}
\end{aligned}
\tag{9.80}
$$

where $h_2(\xi) \cong \xi e^{-\frac{\xi^2}{2}}, \langle h_2(\xi)^2\rangle = 3^{-3/2}, \langle h_1(\xi)\rangle = 0.5, \text{and}\, \langle h_1(\xi)^2\rangle = 0.44$.

It is remarkable that both the sigmas of the horizontal and vertical kick are a bit larger than the average vertical kick. An estimate for the variation of these σ with amplitude is shown in Figure 9-29(b).

We have gone in to some detail here because these kicks are large. Since there can be no spoiler in the second stage of the IP collimation, either the collimators at the next stage need to be set back this far, or it could be that the beam is blown up by the kicks and has a sufficiently large area that it is permissible for it to hit the collimator. As mentioned above, the sigma of the kick near the wall is very close to the magnitude of

the kick itself. And both the vertical and horizontal beam sizes are blown up. However the beam size must be very large for copper. We shall see that a kick of 135σ is required to blow up the beam so that the area of the beam is larger than the mandatory $(2.2\,\mathrm{mm})^2$.

To get an estimate for the enlarged horizontal σ we use the product

$$\begin{aligned} \text{“}\sigma_{y_2}\text{”} &= R_{34}\sigma_{\Delta y_1'} = \frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle}\frac{\langle \Delta y_1' \rangle}{\sigma_{y_1'}} R_{34}\sigma_{y_1'} = \frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}\sigma_{y_2} \\ \text{“}\sigma_{x_2}\text{”} &= R_{12}\sigma_{\Delta x_1'} = \left\langle \frac{R_{12}}{R_{34}} \right\rangle \frac{\sigma_{\Delta x_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}\sigma_{y_2} \\ \text{“}\sigma_{x_2}\text{”}\text{“}\sigma_{y_2}\text{”} &= \left\langle \frac{R_{12}}{R_{34}} \right\rangle \frac{\sigma_{\Delta x_1'}}{\langle \Delta y_1' \rangle}\frac{\sigma_{\Delta y_1'}}{\langle \Delta y_1' \rangle} n_{y_1'}^2\sigma_{y_2}^2 \end{aligned} \tag{9.81}$$

The average $< R_{12}/R_{34} >$ is an rms average weighted by the relative importance of the three collimators in stage 1. For the lattice described in Section 9.2.5 it has a value of about 1/2. Also for this lattice $\sigma_{y2} = 23\,\mu\mathrm{m}$. To get the sigma product to be $4.8\,\mathrm{mm}^2$ we will need $n_{y'} = 135$. This is larger than the $118\text{-}\sigma_{y'}$ maximum kick, and will not occur. Figure 9-30 shows the phase space distortions due to these large kicks when the beam is off-axis.

We might conclude that the aperture of vertical IP stage 2 collimation must be (118+35) $\sigma_{y'} = 153\sigma_{y'}$ because the beam could start out near the collimator edge. But if the beam was already out at 35 σ before receiving the wake kick, it would have gotten a rather large kick in the first stage of the IP collimation. That implies that all of the beam could not have been be near the edge. This effect dies off rapidly though, and by an amplitude of 30 σ the kick was 15 σ' with a spread of about 9 σ'. For an amplitude of 28 σ, the spread is 4.5 σ'. The conclusion is that the required aperture is about $(118 + 30)\sigma_{y'} \approx 150\sigma_{y'}$, not much less.

Fortunately the dynamic aperture of the final-focus system is large enough to accommodate these large amplitude particles. See Figure 9-5.

9.2.5 Lattice Description and Analysis

Collimation System Schematic

Figure 9-31 is a schematic layout of the collimation-system spoilers and absorbers according to the guidelines arrived at in the preceding sections. In this system we have added a second stage for the FD-phase collimation. This is not necessary according to our estimates. The inclusion of a second FD-phase collimation section should be considered tentative. In the diagram of Figure 9-31 we have let a spike represent a titanium spoiler and a rectangular block represent a copper absorber. Further,

- $x + \eta\delta$ indicates collimation outside of the lines $x + \eta\delta = \pm n_x\sigma$ in an (x, δ) plane, with x at the IP phase,
- $x' + \eta\delta$ indicates collimation outside of $x' + \eta\delta = \pm n_x\sigma$ in an (x', δ) plane, with x' at the FD phase, and
- y and y' indicate collimation at $\pm n_y\sigma$ in the IP and FD phase, respectively.

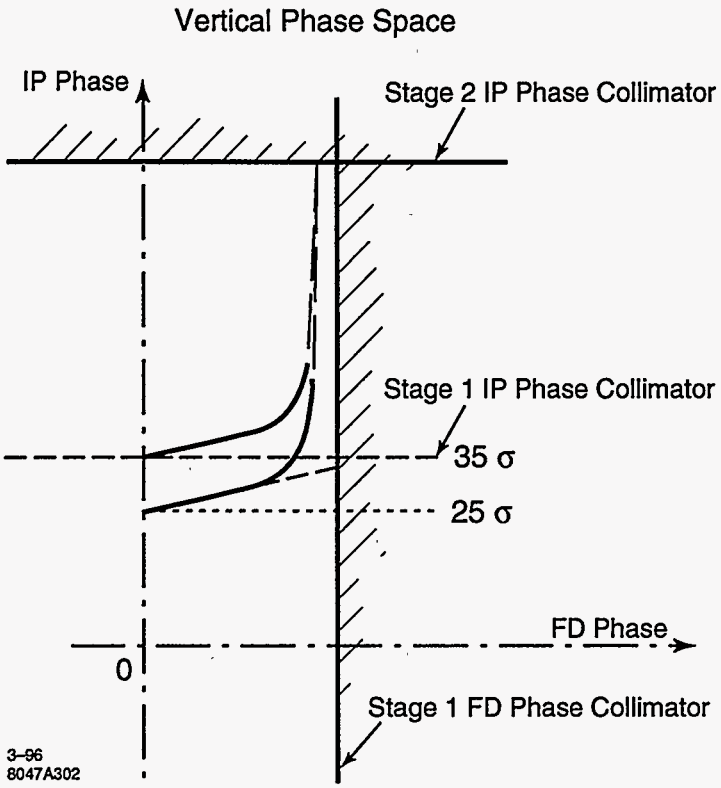


Figure 9-30. *This figure illustrates the enlargement of the aperture that is necessary to avoid impact of a mis-steered beam which narrowly misses upstream spoilers.*

Figure 9-32 shows the volume in (x, x', δ) space collimated by this arrangement. The horizontal and vertical collimation can be interleaved since on the one hand we need large βs in both planes to achieve sufficient spot area, and on the other hand, the chromaticity can be compensated by interleaved sextupoles without problems arising from octupole aberrations.

Each element in the first spoiler and absorber set is located at $-I$ from a corresponding element in the second set. Absorbers with lesser absorption requirements, that will be located downstream of spoilers and absorbers to protect various components and absorb remaining scatter, have not been specifically indicated.

The Lattice

Lattice Functions. Lattice functions for an optical system which satisfies the requirements we have outlined are shown in Figure 9-33 [Helm]. Sextupoles are located at each of the maximum β points.

System Bandpass. Figure 9-34 shows the exit spot size as a function of incident energy for the 1-TeV-c.m. collimation system. At $\delta = \pm 0.5\%$ the spot size has grown by about 4%. This is satisfactory, when one considers that the net result for a beam will be an average over the energy distribution.

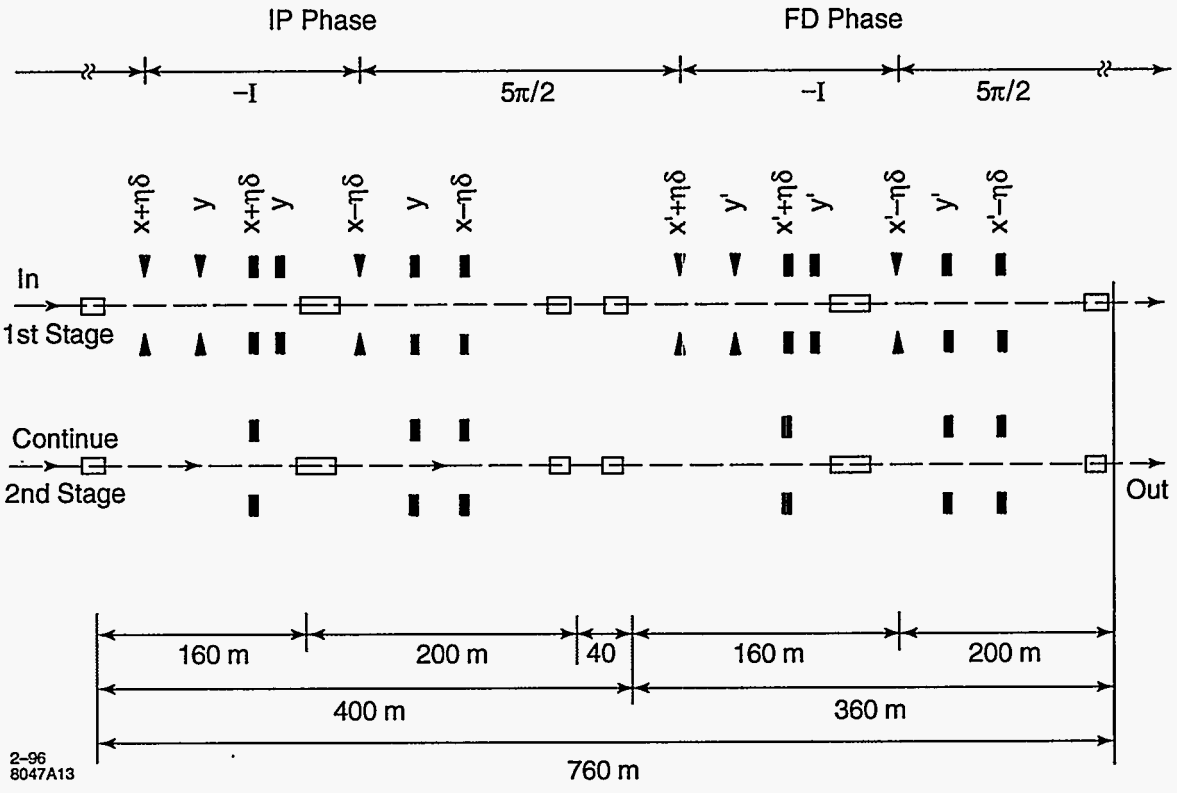


Figure 9-31. *Schematic of the collimation system spoilers and absorbers.*

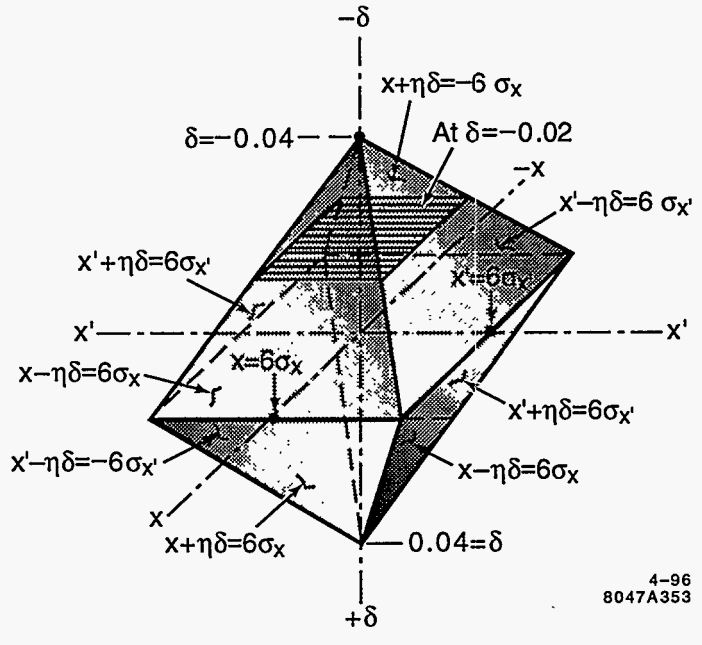


Figure 9-32. *The 3-D volume in x, x', δ space that remains after the beam passes through the spoilers of the collimation system defined in Figure 9-31.*

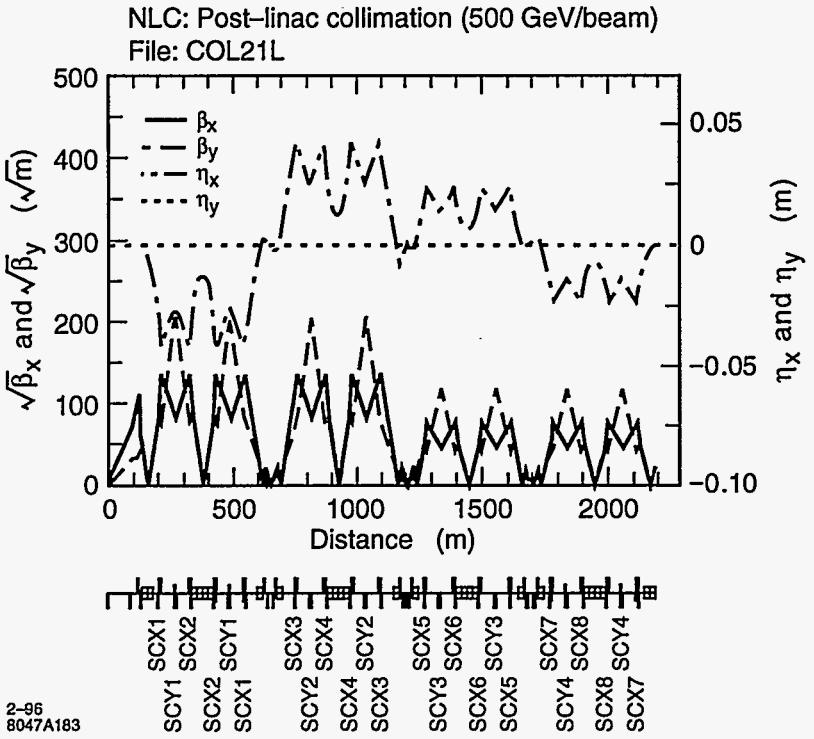


Figure 9-33. *The 1-TeV-c.m. collimation system lattice functions.*

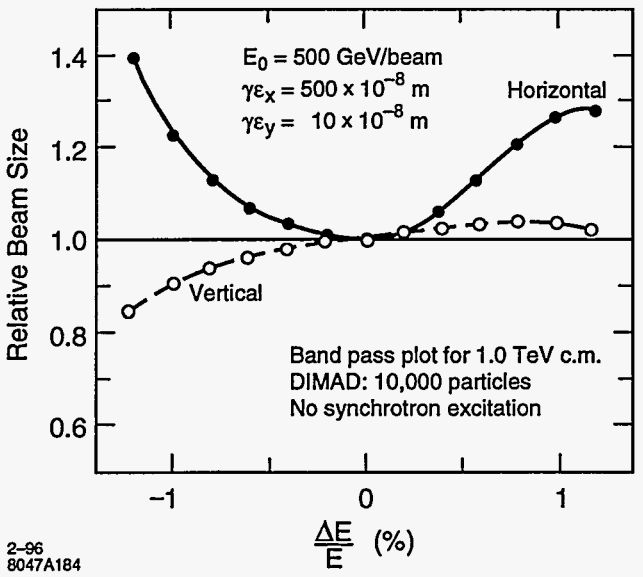


Figure 9-34. *The spot size increase in the 1-TeV-c.m. collimation system as a function of mono-energetic energy offset.*

Collimator Apertures and Wakes

Collimator Apertures and Small Amplitude Wakes. We will now turn to calculating the minimum values of n_x and n_y that may be collimated. The results are presented in Table 9-5.

For this discussion we will concentrate on the FD phase, and assume that the system consists of two stages, one in the collimation system denoted in Table 9-5 by FD_V_1 and FD_H_1 for the vertical and horizontal plane, respectively, and one in the final-focus system, labeled FD_V_3 and FD_H_3. The data on the second stage of FD collimation is also included but the results are not summed into the totals. In the vertical plane we assume there are one Ti spoiler and one Cu absorber in the first stage and two copper-plated W-Re absorbers in the second stage. In the horizontal plane we will assume there are two Ti spoilers and two Cu absorbers in the first stage and two copper-plated W-Re absorbers in the second stage. The spoilers and absorbers are doubled in the first stage of the horizontal plane because energy is also collimated in this plane.

The n_i have been chosen to produce a t'/t value less than 0.7. They have been expanded one stage to the next to allow for near-wall wake effects and alignment inaccuracies.

In the final-focus system the gaps are determined by a complicated set of relationships that determine optimum βs required for minimizing higher order aberrations, tolerances and system length. At 1-TeV c.m. we have (for the 1.5-TeV design operated at 1.0 TeV, see Figure 11-12) $\beta_{x,max} = 10^5$ and $\beta_{y,max} = 3\ 10^5$. βs are a factor of three smaller for the 0.5-TeV-c.m. design, Figure 11-8. They can be taken to scale approximately linearly with energy. Hence we have chosen the worst-case 1.5-TeV-c.m. β values in Table 9-5.

In the first stage the gaps are determined by the requirements that the βs be large enough so that the impact of a full beam train will not destroy the spoiler. For $\varepsilon_x = 5\ 10^{-12}$ m-rad and $\varepsilon_y = 8\ 10^{-14}$ m-rad, we have the condition that at the spoilers, $\sqrt{(\beta_x\beta_y)} \geq 1.6\ 10^4$, where we used the condition from Table 9-3 that $\sqrt{(\sigma_x\sigma_y)} \geq 100\,\mu$m. For Table 9-5 we have used the values from the lattice of Figure 9-33. Because the parameters have changed during the course of design, this does not quite meet the beta-product specification at the horizontal spoilers. For this lattice $\sqrt{(\beta_x\beta_y)} = 1.1\ 10^4$ at the horizontal spoilers and $\sqrt{(\beta_x\beta_y)} = 1.7\ 10^4$ at the vertical spoilers . The $\sqrt{\beta}$s must be 25% larger at the horizontal spoilers. Another change in this lattice must be made to increase the R_{12} and R_{34} between spoiler and absorber. See Section 9.2.2. On the other hand the second stage of FD collimation can be omitted altogether. Present system length will remain adequate.

In Table 9-5 we see that within the collimation system gaps run between a comfortable 0.8 mm and 2.0 mm. The ratio of the wake kick to the kick of the same collimator with an optimum gap, denoted by ϕ_i, has a maximum value of 2.4.

All absorbers have a length of 20 r.l. If necessary this could be reduced to 10 r.l. In this sense we feel there is some margin in the aperture limits derived.

Collimator Focusing and Near-Wall Kick Strengths. Table 9-6 presents the results from calculating the focusing strengths and near-wall kick strengths of all collimators. The focusing totals agree with the estimations of Section 9.2.4 and, because of a waist location at the IP that varies longitudinally within the bunch, will add 0.5% to the luminosity loss.

The near-wall kick strengths for on-axis beams for the 1st vertical FD stage (FD_V_1) add up to 9.8 σ'. The estimate of Eq. 9.79 equal to 0.28 n_y has the value 10 σ'. Because the second-stage IP collimators must be opened up to 150 σ to insure passive protection from mis-steered beams, the 10 σ plays little role here. And because the second-stage collimators are set this wide, the wake effects on the second (or third) stage FD collimators is tiny total of 0.4 σ.

Location	Sect.	Collimator	$\sqrt{\beta}$ ($\sqrt{m}$)	n_i	g_i (mm)	g_i/g_{0i}	ϕ_i	$t_i'/t_i = (\overline{n}_i/n_i)^2\phi_i$
FD_V_1	1	Ti Spoiler	200	36	2.0	11.1	2.3	0.22
		Cu Absorber	40		0.8	0.8	1.0	0.10
	2	W Absorber	200		1.4	3.7	1.4	0.09
FD_V_2	2	W Absorber	140	40	0.9	2.3	1.2	(0.06)
FD_V_3	1	W Absorber	550	42	6.5	17.0	2.8	0.13
	2	W Absorber	550		6.5	17.0	2.8	0.13
Totals:		**6 Collimators**						**0.66**
FD_H_1	1	Ti Spoiler	140	5	1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.7	1.0	0.09
	2	Ti Spoiler	140		1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.7	1.0	0.09
FD_H_2	1	W Absorber	80	6	1.0	2.8	1.2	(0.04)
	2	W Absorber	80		1.0	2.8	1.2	(0.04)
FD_H_3	1	W Absorber	300	7	4.7	12.1	2.4	0.06
	2	W Absorber	300		4.7	12.1	2.4	0.06
Totals:		**8 Collimators**						**0.63**
IP_V_1	1	Ti Spoiler	200	36	2.0	11.1	2.3	0.22
		Cu Absorber	40		0.8	0.8	1.0	0.10
	2	W Absorber	200		1.4	3.7	1.4	0.09
IP_V_2	2	W Absorber	140	150	3.3	3.3	1.3	0.01
Totals:		**4 Collimators**						**0.42**
IP_H_1	1	Ti Spoiler	140	5	1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.5	1.0	0.09
	2	Ti Spoiler	140		1.6	8.6	2.0	0.16
		Cu Absorber	140		1.6	1.5	1.0	0.09
IP_H_2	1	W Absorber	75	15	2.7	2.6	1.2	0.01
	2	W Absorber	75		2.7	2.6	1.2	0.01
Totals:		**6 Collimators**						**0.52**

Table 9-5. *A list of all the collimators in the collimation and final-focus system showing the number of σ collimated at each and the contribution to the total collimation budget. The β functions from the lattice shown in Figure 9-33 have been used to calculate beam sizes at each collimator. The second stage of the FD collimation is optional and has not been included in totals. The large second-stage IP apertures are dictated by machine protection requirements. Calculations are based on $\alpha_G = \alpha_R = \pi^2/8$.*

Location	Sect	Collim.	Quad "$k\beta$" Strength for On-Axis Bm.		Near Wall Kick for On-Axis Bm.		Near Wall Kick for Near Wall Beam			
			RW	G	RW (σ')	G (σ')	RW nr	RW (σ')	G nr	G (σ')
FD_V_1	1	Ti Spoil	0.01	0.02	1.9	3.4	48	34	24	17.0
		Cu Abs.	0.01	0.01	1.4	1.0	124	37	16	5.0
	2	W Abs.	0.01	0.01	0.8	1.3	40	20	12	6.0
FD_V_2	2	W Abs.	0.00	0.00	0.7	0.9	64	18	16	4.0
FD_V_3	1	W Abs.	0.01	0.01	1.3	2.3	53	102	6	11.0
	2	W Abs.	0.01	0.01	1.3	2.3	53	102	6	11
Totals:		**6 Coll.**	**0.05**	**0.06**	**7.4**	**11.2**				
FD_H_1	1	Ti Spoil	0.01	0.01	0.2	0.3	110	0.7	25	1.5
		Cu Abs.	0.01	0.00	0.1	0.1	270	1.7	11	0.7
	2	Ti Spoil	0.01	0.01	0.2	0.3	180	0.7	25	1.5
		Cu Abs.	0.01	0.00	0.1	0.1	270	1.7	11	0.7
FD_H_2	1	W Abs.	0.00	0.00	0.1	0.1	240	0.9	13	0.5
	2	W Abs.	0.00	0.00	0.1	0.1	240	0.9	13	0.5
FD_H_3	1	W Abs.	0.00	0.00	0.1	0.2	6	0.8	6	0.9
	2	W Abs.	0.00	0.00	0.1	0.2	6	0.8	6	0.9
Totals:		**8 Coll.**	**0.04**	**0.02**	**0.9**	**1.4**				
IP_V_1	1	Ti Spoil	0.01	0.02	1.9	3.4	48	34.0	24	17
		Cu Abs.	0.01	0.01	1.4	1.0	120	37.0	16	5
	2	W Abs.	0.01	0.01	0.8	1.3	40	20.0	12	6
IP_V_2	2	W Abs.	0.000	0.00	0.5	0.6	240	69.0	10	3
Totals:		**4 Coll.**	**0.03**	**0.04**	**4.6**	**6.3**				
IP_H_1	1	Ti Spoil	0.01	0.01	0.2	0.3	18	1.1	25	1.5
		Cu Abs.	0.005	0.00	0.1	0.1	27	1.7	11	0.7
	2	Ti Spoil	0.01	0.01	0.2	0.3	18	1.1	25	1.5
		Cu Abs.	0.005	0.00	0.1	0.1	27	1.7	11	0.7
IP_H_2	1	W Abs.	0.00	0.00	0.1	0.1	76	2.7	10	0.4
	2	W Abs.	0.00	0.00	0.1	0.1	76	2.7	10	0.4
Totals:		**6 Coll.**	**0.03**	**0.02**	**0.8**	**1.0**				

Table 9-6. *A list of all the collimators in the collimation and final-focus system showing the number of σ collimated at each and the contribution to the total collimation budget. The β functions from the lattice shown in Figure 9-33 have been used to calculate beam sizes at each collimator. Geometric wake numbers should be considered to be tentative for parallel-plate tapers.*

The second-stage collimators must be opened up because of the large near-wall wakes for mis-steered beams. The total for the resistive-wall and geometric wakes in the second-stage vertical FD phase is 119 σ'. To this we must consider that the beam could have already been at 30 σ in the IP phase without having been too distorted, hence the aperture must be 150 σ. As we have argued in Section 9.2.4, a kick of 135 σ' would imply a beam sigma product of 4.8 mm^2, which would also insure the survival of the collimator from an impact of a full train. In the horizontal plane the 1st stage total is 9.2 σ'. The aperture has been opened to 15 σ to account for the fact that the beam could already have been close to 5 σ. There is no danger of the beam ending up near these apertures and being steered into collimators in the final-focus system.

Tolerances

The lattice has large β functions and some very large R_{12} and R_{34} values which will lead to important tolerances on quadrupole strength and position stability and on sextupole position stability. Since we are designing near the limit allowed by wakefield intensities, we can also expect important tolerances on collimator jaw stability.

There are four important time domains for tolerance specification:

- capture tolerances,
- long term stability,
- short term stability, and
- jitter tolerances.

Capture tolerances are the absolute alignment or strength tolerances required at start-up so that the system can be aligned and tuned using beam-based techniques.

Long term stability tolerances are those limits that when exceeded would require re-execution of these beam-based alignment strategies.

Several tuning knobs will be activated perhaps every hour at the IP to adjust sensitive aberrations affecting the IP beam spot size. Change in waist and skew aberrations can also be observed in the skew correction system (SCS). Short-term stability tolerances are the tolerances that must be held until tuning knobs are reset or aberrations are rechecked and corrected. We estimate that the SCS scans occur ten times per hour, if required, since this process can be non-invasive.

The beam centroid is stabilized with feedback systems based on orbit measurements and a rule of thumb estimate is that they are effective for times longer than or equal to 15 beam pulses, or about 1/10th of a second. Jitter tolerances refer to requirements on the stability of the system for times shorter than this time.

Except for initial collimator jaw alignment, this section discusses only the stability tolerances associated with the last two time scales.

Tolerance Budgets. The alignment of the beam collision at the IP is maintained with a fast feedback system. As noted above, the jitter tolerances refer to beam-line changes on time scales which can not be corrected by this feedback system.

Machine Section	Steering Budget (H/V) σ	Accumulated Jitter (H/V) σ	Luminosity Loss (H/V) %
To end of linac	0.1 / 0.25	0.10 / 0.25	0.5 / 0.8
Collimation system	0.1 / 0.2	0.14 / 0.32	1.0 / 1.3
w/ amplification		0.18 / 0.40	1.6 / 2.0
Big bend	0.05 / 0.1	0.19 / 0.41	1.8 / 2.1
Final focus	0.1 / 0.2	0.21 / 0.46	2.2 / 2.6
Final doublet	0.1 / 0.2	0.23 / 0.50	2.7 / 3.1

Table 9-7. *A proposal for a jitter budget.*

In addition to limiting luminosity loss, a budget is necessary because with a large beam jitter it becomes difficult to carry out beam-based alignment, wakefield effects are enhanced, and synchrotron radiation in the final doublet is increased. (See Section 11.7). Table 9-7 is a proposal for a jitter budget.

There are many aberrations which must be tuned in the final-focus system using the spot size at the IP as the diagnostic tool (see Table 11-9). Once the aberrations are tuned, changes in the beam line can cause an increase in aberration strength, and that aberration must be re-tuned, or compensated based on an SCS measurement. Several of these aberrations can result from changes in the collimation beam line. These are: horizontal and vertical waist, horizontal and vertical dispersion, and one skew aberration. The collimation system has been allotted 0.5% luminosity change for each of these seven aberrations.

Quadrupole Strength Stability. Each FD-phase section has four quadrupoles at particularly large values of the vertical β function, about 30 km. A small change in the strength of any of these quads would cause a shift of the waist at the IP. On the assumption that changes in the beam size can be measured with a relative resolution of 10%, it follows that each independent aberration can be adjusted so as to leave a residual increase in beam size of only 2%. SLC experience indicates this may be less than 0.5%. The waist position is one of these aberrations, thus a waist-stability budget of 2% must be apportioned to the various modules of the collider. Let us assume that 1/2 of this is assigned to the collimator system, and of that is divided further into 1/2 for quadrupole-strength stability and 1/2 for sextupole-position stability. Hence the quadrupole-stability budget is 1/2%. Assuming that the majority of this can be assigned to the sixteen quadrupoles at large vertical β, and that they are on independent power supplies, it can be estimated that

$$\frac{dk}{k} \leq \frac{1}{60}\,\frac{1}{k\beta} \tag{9.82}$$

For our lattice $k\beta = 1,200$ hence $dk/k < 1.3\ 10^{-5}$.

Based on sensitivity calculations for each element, Table 9-8 assigns strength stability tolerances to all collimation system quadrupoles so that the collimation system meets the allotted 0.5% horizontal and vertical waist budget. We have conservatively included both phases in the list of Table 9-8, hence there are in fact 32 quadrupoles that have been assigned a tolerance of $dk/k < 1\ 10^{-5}$. This could be reduced to 16 with a tolerance of $1.8\ 10^{-5}$.

Sextupole Position. The results of this section are summarized in Table 9-9.

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
Q703	17.00	1.00	1.900	0.100	11.00	0.5
Q704	15.00	1.00	0.810	0.050	113.00	1.0
Q704	15.00	1.00	0.820	0.050	12.00	0.5
Q705	18.00	1.00	16.000	1.000	2.90	0.2
Q705	18.00	1.00	58.000	1.000	2.90	0.2
Q706	11.00	1.00	0.710	0.050	25.00	0.5
Q706	11.00	1.00	0.720	0.050	32.00	0.5
Q707	9.50	1.00	2.800	0.100	5.40	0.5
Q707	9.80	1.00	2.900	0.100	5.40	0.5
Q708	10.00	1.00	4.800	0.300	19.00	1.0
Q708	10.00	1.00	4.700	0.300	19.00	1.0
Q709	6.90	1.00	4.100	0.300	23.00	1.0
Q709	6.90	1.00	3.900	0.300	24.00	1.0
Q710	20.00	1.00	2.200	0.100	76.00	1.0
Q710	20.00	1.00	2.200	0.100	66.00	1.0
QCL1	94.00	1.00	13.000	1.000	230.00	1.0
QCL2	12.00	1.00	16.000	1.000	13.00	0.5
QCL3	0.15	0.01	10.000	1.000	140.00	1.0
QCL4	0.54	0.03	14.000	1.000	34.00	0.5
QCX1	0.18	0.03	16.000	1.000	130.00	1.0
QCX2	0.13	0.01	41.000	1.000	210.00	1.0
QCX2	0.12	0.01	45.000	1.000	220.00	1.0
QCX3	0.11	0.01	310.000	1.000	360.00	1.0
QCX3	0.11	0.01	370.000	1.000	310.00	1.0
QCX2	0.12	0.01	56.000	1.000	42.00	0.5
QCX2	0.14	0.01	49.000	1.000	43.00	0.5
QCX1	0.18	0.01	19.000	1.000	37.00	0.5
QCX1	0.18	0.01	16.000	1.000	130.00	1.0
QCX2	0.13	0.01	41.000	1.000	210.00	1.0
QCX2	0.12	0.01	45.000	1.000	220.00	1.0
QCX3	0.11	0.01	310.000	1.000	360.00	1.0
QCX3	0.11	0.01	370.000	1.000	310.00	1.0
QCX2	0.12	0.01	56.000	1.000	42.00	0.5
QCM1	0.16	0.01	57.000	1.000	50.00	0.5
QCM2	0.23	0.03	24.000	1.000	45.00	0.5
QCM3	5.60	0.10	10.000	1.000	75.00	0.5
QCM4	98.00	1.00	5.600	0.300	37.00	0.5
QCM4	108.00	1.00	6.300	0.300	29.00	0.5
QCM3	7.30	0.10	0.850	0.050	2.00	0.2
QCM2	0.22	0.03	0.120	0.025	0.42	0.05
QCM1	0.15	0.01	0.210	0.030	0.43	0.05

Table 9-8. *The vibration and strength stability sensitivities and tolerances of all the quadrupoles in the collimation system. The sensitivities are calculated as if the one quadrupole was assigned the entire budget of the section. There are about 100 quadrupoles, so in general the tolerance is 1/10th of the sensitivity. Continued ...*

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
QCX2	0.12	0.01	0.170	0.030	0.36	0.05
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCX2	0.14	0.01	0.180	0.030	0.38	0.05
QCX1	0.18	0.01	0.097	0.025	0.36	0.05
QCX1	0.18	0.01	0.097	0.025	0.36	0.05
QCX2	0.13	0.01	0.180	0.030	0.38	0.05
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX3	0.11	0.01	0.140	0.025	1.20	0.10
QCX2	0.12	0.01	0.180	0.030	0.36	0.05
QCM1	0.15	0.01	0.210	0.030	0.45	0.05
QCM2	0.23	0.03	0.130	0.025	0.44	0.05
QCM3	5.60	0.10	1.100	0.100	1.80	0.20
QCM4	98.00	1.00	5.800	0.300	35.00	0.50
QCN4	98.00	1.00	4.900	0.300	51.00	0.50
QCN3	11.00	1.00	8.800	0.300	43.00	0.50
QCN2	0.72	0.03	11.000	1.000	44.00	0.50
QCN1	0.47	0.03	24.000	1.000	52.00	0.50
QCX2	0.38	0.03	24.000	1.000	46.00	0.50
QCX3	0.33	0.03	100.000	1.000	1300.00	1.00
QCX3	0.34	0.03	110.000	1.000	710.00	1.00
QCX2	0.38	0.03	35.000	1.000	35.00	0.50
QCX2	0.42	0.03	30.000	1.000	36.00	0.50
QCX1	0.57	0.03	11.000	1.000	29.00	0.50
QCX1	0.57	0.03	8.700	0.300	36.00	0.50
QCX2	0.42	0.03	22.000	1.000	46.00	0.50
QCX2	0.38	0.03	24.000	1.000	46.00	0.50
QCX3	0.33	0.03	100.000	1.000	1300.00	1.00
QCX3	0.34	0.03	110.000	1.000	710.00	1.00
QCX2	0.38	0.03	35.000	1.000	35.00	0.50
QCP1	0.48	0.03	34.000	1.000	40.00	0.50
QCP2	0.75	0.03	14.000	1.000	36.00	0.50
QCP3	9.40	0.10	8.800	0.300	61.00	0.50
QCP4	73.00	1.00	5.800	0.300	47.00	0.50
QCP4	78.00	1.00	6.200	0.300	38.00	0.50
QCP3	11.00	1.00	1.900	0.100	3.10	0.20
QCP2	0.72	0.03	0.220	0.030	0.78	0.10
QCP1	0.47	0.03	0.370	0.030	0.76	0.10

Table 9-8. *Cont.*

Quad name	Strength Sensitivity (10^{-3})	Sensitivity Tolerance (10^{-3})	Vertical Vibration Sensitivity (10^{-7} m)	Vertical Vibration Tolerance (10^{-7} m)	Horizontal Vibration Sensitivity (10^{-7} m)	Horizontal Vibration Tolerance (10^{-7} m)
QCX2	0.38	0.03	0.320	0.010	0.64	0.10
QCX3	0.33	0.03	0.240	0.030	2.10	0.20
QCX3	0.34	0.03	0.240	0.030	2.10	0.20
QCX2	0.38	0.03	0.320	0.030	0.64	0.10
QCX2	0.43	0.03	0.320	0.030	0.68	0.10
QCX1	0.57	0.03	0.170	0.030	0.64	0.10
QCX1	0.57	0.03	0.170	0.030	0.63	0.10
QCX2	0.41	0.03	0.330	0.030	0.67	0.10
QCX2	0.38	0.03	0.320	0.010	0.65	0.10
QCX3	0.33	0.03	0.240	0.030	2.10	0.20
QCX3	0.34	0.03	0.240	0.030	2.10	0.20
QCX2	0.38	0.03	0.320	0.030	0.65	0.10
QCX2	0.43	0.03	0.330	0.030	0.68	0.10
QCX1	0.57	0.03	0.170	0.030	0.65	0.10
QMD1	3.00	0.10	0.490	0.050	1.70	0.20
QMD1	2.20	0.10	0.520	0.050	1.40	0.20
QMD2	0.98	0.10	0.570	0.050	0.90	0.10
QMD2	0.99	0.10	0.560	0.050	0.90	0.10
QMD3	3.60	0.10	0.740	0.050	6.80	0.20
QMD3	3.80	0.10	0.760	0.050	7.50	0.20
QMD4	25.00	1.00	1.100	0.100	5.90	0.20
QMD4	26.00	1.00	1.100	0.100	5.70	0.20

Table 9-8. *Cont.*

- Horizontal Sextupole Position—The calculation of the preceding section can be used to estimate the horizontal position stability of the sextupole since any sextupole motion adds a quadrupole field to the beam line of magnitude $k_S dx$. For a 0.5% contribution to spot size increase the displacement would be limited by the relationship (valid for the FD-phase sextupoles only)

$$k_s dx \leq \frac{1}{10}\,\frac{1}{\beta_y} \tag{9.83}$$

depending on whether one is looking at the vertical waist or horizontal waist.

For the sextupoles of the stage 1 of the FD phase at large $\beta_x = 1.7\ 10^4$ where $k_S = 1.16\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 5\,\mu$m. For the sextupoles of the stage 2 of the FD phase at large $\beta_x = 5.3\ 10^3$ where $k_S = 2.6\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 7.5\,\mu$m. These four pairs sextupoles contribute to the horizontal waist aberration. Since the quadrupole stability tolerances are mandated by vertical waist considerations, the entire 1/2% horizontal waist aberration budget may be allocated to these four sextupole pairs. If we require the same tolerance for all pairs, the result is $dx < 3\,\mu$m.

For the sextupoles of stage 1 of the FD phase at large, $\beta_y = 4\ 10^4$ where $k_S = 1.5\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 1.6\,\mu$m. For the sextupoles of stage 2 of the FD phase at large $\beta_y = 1.3\ 10^4$ where $k_S = 2.8\,\mathrm{m}^{-2}$, Eq. 9.83 implies $dx \leq 2.7\,\mu$m. These sensitivities are based on the vertical waist aberration, for which we have budgets 0.5% for both pairs. If the tolerance of both pairs is taken to be identical, the result is $dx \leq 1.4\,\mu$m.

Beam-line Location	Sextupole Name	Horizontal Sensitivity (μm)	Horizontal Tolerance (μm)	Vertical Sensitivity (μm)	Vertical Tolerance (μm)
V_FD_1	SCY1	1.6	1.4	0.5	0.3
H_FD_1 (2)	SCX1&2	5.0	3.0	1.2	0.4
V_FD_2	SCY4	2.7	1.4	0.8	0.4
H_FD_2 (2)	SCX7&8	7.5	3.0	1.6	0.5

Table 9-9. *Stability sensitivities and tolerances for collimation system sextupole positions.*

The exact meaning of these sextupole-position tolerances must be stated with some care. A change in the position of any quadrupole between large-β_y paired sextupoles that would cause a displacement of the beam at the second sextupole by 1.4 μm would give rise to the vertical waist aberration. Thus one sees that it is not the absolute position of the sextupole that is of concern, but its position with respect to the sextupole with which it is paired, and the position of this sextupole pair with respect to the beam line as defined by the quadrupoles. For details on stabilization strategies see Section 9.2.6 below.

- Vertical Sextupole Position—A change in the vertical sextupole position gives rise to a skew quad aberration of strength $k_S dy$. For a 0.5% spot size increase we have, for FD-phase sextupoles, the condition

$$k_s dy \leq \frac{1}{10}\sqrt{\frac{\varepsilon_y}{\varepsilon_x}}\,\frac{1}{\sqrt{\beta_x \beta_y}} \tag{9.84}$$

At the $\beta_{y,\max}$ in stage 1 we have $\sqrt{(\beta_x\beta_y)} = 1.710^4$. The resulting 0.5% sensitivity, for $k_s = 1.5\,\mathrm{m}^{-2}$, is $dy \leq 0.5\,\mu$m. At the $\beta_{x,\max}$ in stage 1 we have $\sqrt{(\beta_x\beta_y)} = 0.95\ 10^4$ and $k_s = 1.1m^{-2}$, with a sensitivity $dy \leq 1.2\,\mu$m.

If the 1/2 percent budget is divided with 1/6% going to each of the $\beta_{x,\max}$ sextupoles and 1/3% going to the $\beta_{y,\max}$, we obtain tolerances of $dy \leq 0.7\,\mu$m and $dy \leq 0.4\,\mu$m, respectively.

At the $\beta_{y,\max}$ in stage 2 we have $\sqrt{(\beta_x\beta_y)} = 5.3\ 10^3$. The resulting sensitivity, for $k_s = 2.8\,\mathrm{m}^{-2}$, is $dy \leq 0.8\,\mu$m. At the $\beta_{x,\max}$ in stage 2 we have $\sqrt{(\beta_x\beta_y)} = 3.1\ 10^3$ and $k_s = 2.6\,\mathrm{m}^{-2}$, with a sensitivity $dy \leq 1.6\,\mu$m.

If the 0.5% budget of the collimation system for the principal skew aberration is divided amongst these six sextupole pairs, one possible assignment are the tolerances listed in Table 9-9. We note that the vertical sextupole position tolerance has the same meaning as the horizontal tolerances discussed in the previous section. See Section 9.2.6 for stabilization strategies.

Energy Jitter Considerations. Variations in bunch intensity give rise to variations in energy through beam-loading effects in the linac, hence the energy can vary bunch-to-bunch within a bunch train, and also bunch-train to bunch-train. This variation can be no larger than the band width of the final-focus system, about $\pm 0.4\%$, but we consider here whether such variations give rise to wakes in the collimation system. This jitter is about 1/10th of 4% and hence is 1/10th of the collimation aperture.

By designing the collimation system with a symmetric dispersion function and in such a way that the transverse wake is the same magnitude at the horizontal spoiler and/or absorber at each end of the $-I$s, the energy jitter will produce the same $\Delta x'$ at the second collimator as the first. However because of the $-I$ transform between these locations, the kicks cancel. Thus a properly designed collimation system can be insensitive to energy jitter.

Quadrupole Position. Several quadrupoles between sextupoles have $k_Q R_{12}$ multipliers to the sextupole of 2.2. Thus the tolerances in Table 9-9 must be divided by 2.2 to get the stability tolerance for such quadrupoles.

Jitter of a quadrupole in the FD phase will cause beam jitter at the IP. Using 1/2% of the jitter budget for the collimation system, and assuming that the 8 quadrupoles located at $\beta_{\max}$ points are moving independently, the tolerance on the quadrupole jitter is given by

$$\Delta y < \frac{1}{5}\,\frac{1}{\sqrt{8}}\,\frac{\sigma_{y'}}{k_Q} = \frac{1}{5}\,\frac{1}{\sqrt{8}}\,\frac{\sigma_y}{k_Q \beta_y} \tag{9.85}$$

Since $\sigma_y = 40\,\mu$m, and $k\beta_y = 1,200$, the tolerance on jitter is $\Delta y < 2\,$nm. A similar consideration of the quadrupoles at the β_x maximum leads to a horizontal jitter tolerance of $\Delta x < 2\,$nm. The $k\beta$ is smaller, but there are twice as many quadrupoles and there is no help from disruption. See Table 9-8 for a list of vibration tolerance for all quadrupoles in the collimation system. These vibration tolerances are chosen so that the contribution from all quadrupoles in the collimation system adds up to at most a 0.5% spot size increase.

Note that these vibration tolerances do not apply to seismic ground motion. They can be taken to be the tolerance of the quadrupole vibration with respect to ground beneath it, or to high frequency cultural noise. In the FFTB the vibration between the quadrupole and the ground beneath it is the order of 1 nm, hence the stands and vibration from cooling water in the FFTB meet the requirements of the vibration tolerances of Table 9-8.

The seismic motion, although larger than the tolerances calculated above, is very correlated at low frequencies (smaller than 5 Hz) and the spectrum is small for frequencies above 5 Hz in tunnels below the ground surface. See Appendix C for details on ground motion. Because of the correlation, one must calculate a lattice response function to see the net effect on the IP beam positions. Figure 9-35 shows this lattice response function for the collimator section for the vertical plane. The response function is also discussed in Appendix C. and its application to the final-focus lattice is discussed in Section 11.5.5.

The lattice response function gives the square of the displacement at the IP for a unit amplitude wave of wave number k in the collimation section. At $k > 0.1$ (corresponding to $\lambda < 60\,$m, or with a velocity of 600 m/s at frequencies $f > 10\,$Hz) the response function average is about 0.04. This means if we adopt the IP position requirement as 0.5 nm, the tolerance in the collimation system is 2.5 nm. This number is larger than quiet ground conditions in this frequency region, but care must be taken not to introduce noise of this magnitude. At lower k the tolerance drops off abruptly. At $k = 0.01$ (corresponding to $\lambda < 600\,$m, or with a velocity of 600 m/s a frequency of $f > 1\,$Hz) the response function $G(k)$ is 10^{-5}, corresponding to a tolerance of 160 nm. The integrated ground motion to this frequency is only a few nm, so is well within this limit.

One can fold the quiet spectrum ground motion together with the lattice response function to calculate the relative beam motion at the IP due to the entire spectrum of seismic noise. The result is a beam displacement at the IP that is the order of 0.1 nm for seismic motion in the collimation system.

Dipole Strength Stability. The tightest tolerance on dipoles occurs for dipoles between the vertical spoiler and vertical absorber of stage 1 of the FD-phase collimation. Since the dispersion function at these sextupoles is $\eta = .026\,$m, we may deduce that $\Sigma\theta_i R_{12,i} = 2\eta = 0.052\,$m, where the sum is over the bends between the sextupoles. The tolerance condition may be written as $\Sigma\Delta\theta_i R_{12,i} \leq 1.4\,\mu$m. Since all θ_i are equal, these two relations can be combined to conclude that $\Delta\theta/\theta \leq 2.7\,10^{-5}$. Using the relation for the two sextupole pairs at the horizontal spoilers and absorbers we have $\Sigma\theta_i R_{12,i} = 0.08\,$m and $\Sigma\Delta\theta_i R_{12,i} \leq 3\,\mu$m, we conclude $\Delta\theta/\theta \leq 3.8\,10^{-5}$. Since these are different aberrations, the smallest of the two estimates prevails.

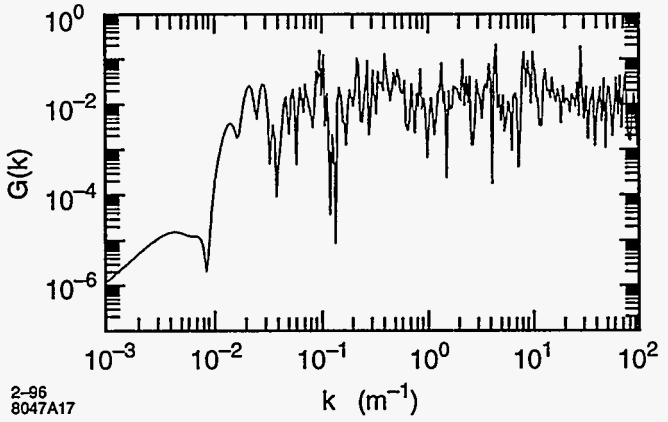


Figure 9-35. *The vertical lattice response function for the 1-TeV-c.m. collimation system.*

Beam Line Location	Dipole Name	$\Sigma\theta_i R_{12,i}$ (cm)	0.02 σ (mm)	Jitter Sens. (10^{-5})	Jitter Toler. (10^{-5})	Stab. Toler. (10^{-5})
Angle Adjust	B2M (2)					N/A
Match In	BCA1 (2)	4	6	15	5	N/A
IP_1	BCA1 (4)	8	6	8	5	N/A
IP→FD	BCM1&2	4/4	6/6	15/15	5	N/A
FD_1	BCA2 (4)	8	6	8	2.7	2.7
FD→IP	BCN1&2	4/2	6/4	15/20	5	N/A
IP_2	BCB1 (4)	4	4	10	5	N/A
IP→FD	BCP1&2	2/2	4/4	20/20	5	N/A
FD_2	BCB2 (4)	4	4	10	4.8	4.8
Match Out	BCB2 (2)	2	4	20	5	N/A

Table 9-10. *Tolerances for the dipoles in the collimation system.*

For comparison, the tolerance for the dipoles of the stage 2 FD phase are $\Delta\theta/\theta \leq 4.8\ 10^{-5}$, the vertical pair ($\Sigma\theta_i R_{12,i} = 0.029\,\text{m}, \Sigma\Delta\theta_i R_{12,i} \leq 1.4\,\text{mm}$), and $\Delta\theta/\theta \leq 6.7\ 10^{-5}$ for the horizontal pairs ($\Sigma\theta_i R_{12,i} = 0.045\,\text{m}, \Sigma\Delta\theta_i R_{12,i} \leq 3\,\text{mm}$). The smallest of these will prevail.

The tolerances calculated in the previous two paragraphs are stability tolerances. Beam spray and change of centroid due to beam loss at collimators, would make it very difficult to use beam-based techniques in FD_1. Since these tolerances relate to conditions at the IP, and do not affect collimation system function, they may be monitored and compensated in the skew correction system. Also, see Section 9.2.6.

There are no similar stability implications for the remaining dipoles, and tolerances can be considered as jitter tolerances. The dipoles of the match in and match out would give jitter in the FD phase. The transition-region dipoles between IP and FD can cause jitter in both planes. The dipoles within a stage, can cause jitter at the phase of that stage. The change in position due to a change in angle can be found from the dispersion function change created by the bends. These are listed in column 3 of Table 9-10.

To calculate the tolerances for the remaining dipoles we use the condition that the dipoles should not give rise to detectable beam jitter. Quantitatively we will require that the contribution to jitter by all collimation system dipoles should be less than $0.02\,\sigma$ (this should be chosen to be negligible compared to expected jitter values). This number, at the location where the dispersion was measured, is given in column 4 of Table 9-10. The sensitivities are the ratios of the dispersion function to the $0.02\,\sigma$ numbers. These are given in column 5 of Table 9-10.

The sensitivities are equally divided between the two phases. Assuming that the bends in each location listed in Table 9-10 are on the same power supply, and that the errors in the bends at different locations are independent, we arrive at the tolerances listed in column 5.

Collimator Jaw Stability. In this section we investigate the effect of the wake from the two horizontal collimators between the vertical sextupoles on the offset of the beam at the downstream sextupole. Pulse-to-pulse displacement at the downstream sextupole from the wakefield kick must be smaller than the sextupole position tolerance of $dx = 0.3\,\mu$m.

The horizontal collimators are almost π apart in x-phase, hence the beam jitter at the horizontal collimators will give rise to a wake kick of opposite sign. But since the R_{12}s to the vertical sextupoles are the same sign, the net displacement at the vertical sextupole will be zero. In this regard it is important that the two collimators have the same wakefield kick strength.

If one of the jaws moves there will be a net kick of course. Since the wake kick from a 1-σ jitter is less than $0.7\,\sigma'$ for the sum of all spoilers or absorbers at a given phase, the wakefield from a single collimator for a 1-σ motion will give a kick of about $0.7/4\,\sigma' \approx 0.2\sigma'$. Since 1σ at the horizontal collimators is about $180\,\mu$m, the kick will be less than 0.2 (17 nr)=3.4 nr. The displacement at the downstream "vertical" sextupole will be about $0.17\,\mu$m.

This consideration also yields the stability tolerance on the collimator jaw of about $320\,\mu$m. This is very loose, and corresponds to the absolute alignment tolerance for the jaws which will be calculated next.

If these considerations had been carried out for the vertical collimators the tolerance would be about 1σ, which for the vertical collimation is about $40\,\mu$m.

Collimator Jaw Alignment. If the beam passes off-center through a collimator the centroid receives a kick, and also the beam emittance grows slightly. These effects were calculated in Section 9.2.4. For an optimized taper it was shown that the luminosity loss due to emittance growth was 2.3 times smaller than the luminosity loss due to the centroid kicks. When kicks arise because of beam jitter, they all have the same sign, and the luminosity loss goes as the number of collimators squared. When the collimator jaws are misaligned, the kicks will have a random sign, and thus the luminosity loss will grow linearly with the number of collimators. So we can calculate the luminosity loss due to each collimator and then add the effects.

Suppose we assume a 1-σ collimator jaw misalignment. Then the t-ratio for that collimator gives the kick magnitude (in numbers of σ'). The luminosity loss, from emittance growth, due to that collimator will be $0.055(t/t')^2$ (see Eq. 9.40). So the luminosity loss for all collimators at a given phase will be $0.055\Sigma(t/t')_i^2$. For the vertical final-doublet phase (FD_V) this sum is 0.1. Hence a random displacement of the vertical collimators by 1σ would result in a luminosity loss of 1/2%. (Note that the emittance growth which occurs in the IP phase does not effect the luminosity. The growth occurring in this phase effects the angular distribution at the IP.) The horizontal sum is 0.05 so the luminosity loss from a random 1σ horizontal collimator displacement would be 1/8%.

Thus the tolerances on the collimator jaw alignment can be taken to be about 1σ. This would result in a total luminosity loss of 3/4%. See Section 9.2.6 for a discussion of how to achieve this alignment. Also see Section 11.5.5 in Chapter 11.

Particle Tracking and Mapping Studies

Tracking Between Collimator Apertures. In order to insure the passive protection of the collimation system absorbers and the final-focus system collimators, it is essential that they lie in the shadow of the initial collimation system spoilers. Our concern is that the presence of uncompensated sextupoles and chromaticity within the system could lead to unexpected beam offsets at the absorbers and collimators. To check this we have traced particles through the system.

The clearest method we have found to verify this functionality and diagnose any failures is to initialize particles uniformly in an x-x' or y-y' (or both simultaneously) phase space for a fixed energy offset δ. In practice one need consider only negative values of δ, since large positive values of δ can not occur. Figure 9-36 shows the result of such an exercise for the horizontal phase space for an energy off-set of $\delta = -0.02$. The profile is shown at four important horizontal absorbers in the collimation system, and for the first collimator in the final-focus system. Two horizontal absorbers not shown have very large apertures and two are in the stage-2 FD-phase section which we have argued is unnecessary.

We first note that the distributions of Figure 9-36 look like straight lines since the x' range arising from even small α functions is much larger than the internal spread of x' within the bunch. Secondly we note that except for the profile at the final-focus collimator, the distributions fall within the collimated aperture, or in the few exceptions, the apertures can be enlarged to accommodate. The apertures of the absorbers in IP_1 and FD_1 can be enlarged to accommodate the slight beam offsets. They must be set back somewhat to avoid beam that could pass near the spoilers and be deflected into the absorbers. The apertures of FD_2 are not crucial, and can be enlarged, since this phase of collimation could be omitted altogether.

The large offset of the profile at the collimator in the final-focus system was a surprise to us because the profiles tracked in an earlier lattice, that were shown above in Figure 9-6, were not a problem. The analysis of this situation is described in the following paragraph.

Mapping between Collimator System and IP. The results of tracking described in the previous section show that there is a problem with the beam displacement at the first collimator in the final-focus section. The changes between this lattice and an earlier lattice that did not have this problem are: i) a modified big bend without chromatic correction, ii) the addition of a 400-m skew correction section, iii) a new angle adjustment module at the entry to the final-focus system, and iv) an enlarged beta match to the final focus which includes a phase adjustment capability.

To pinpoint the source of the position offset, we created maps [Yan] of each section of the transport line from the collimation system to the final-focus system. Analysis of these maps showed that linear chromaticity and δ^2 dispersion must be corrected. When these terms are corrected, the image falls within the aperture. Figure 9-37 shows the maps for each section and the complete beam line before the chromatic corrections are introduced, and Figure 9-38 shows the situation afterward. The pictures are drawn in a normalized phase space with the IP phase horizontal and the FD phase vertical.

As pointed out, the stage-2 FD-phase collimation is redundant, and can be removed. Then the IP-phase stage-2 collimation will be located between the last FD-phase collimation and the final-focus system and could be used for chromatic correction.

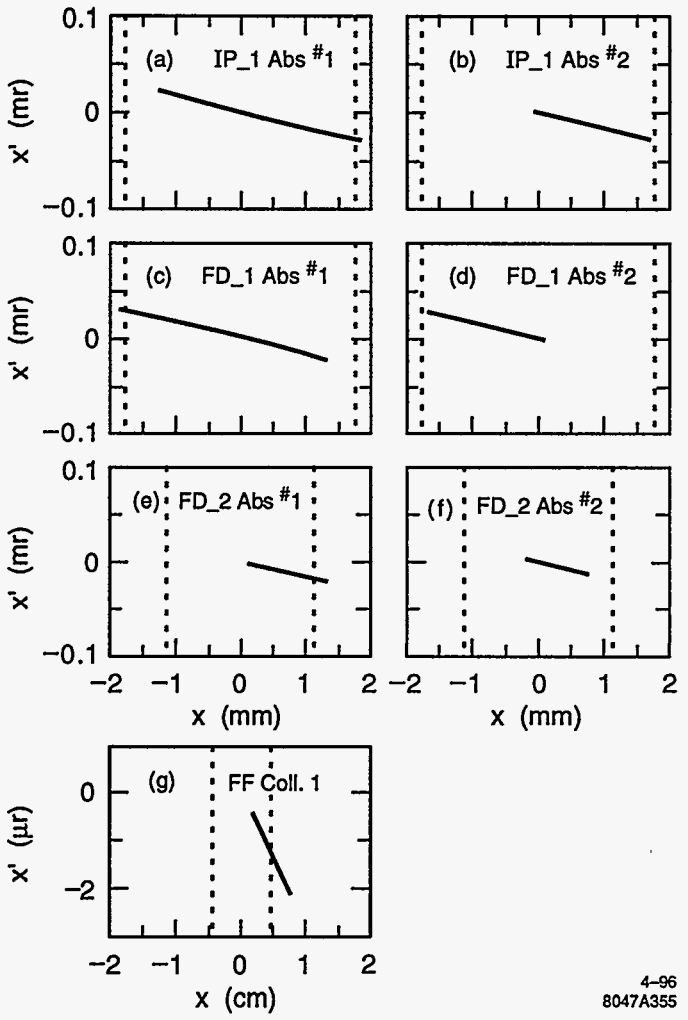


Figure 9-36. *This figure shows the downstream image of particles released at the entrance to the collimation system at $\delta = -0.02$ and on a 6σ square in $x - x'$ phase space. The four pictures are the image at all absorbers in the collimation system and at the collimators in the final-focus system.*

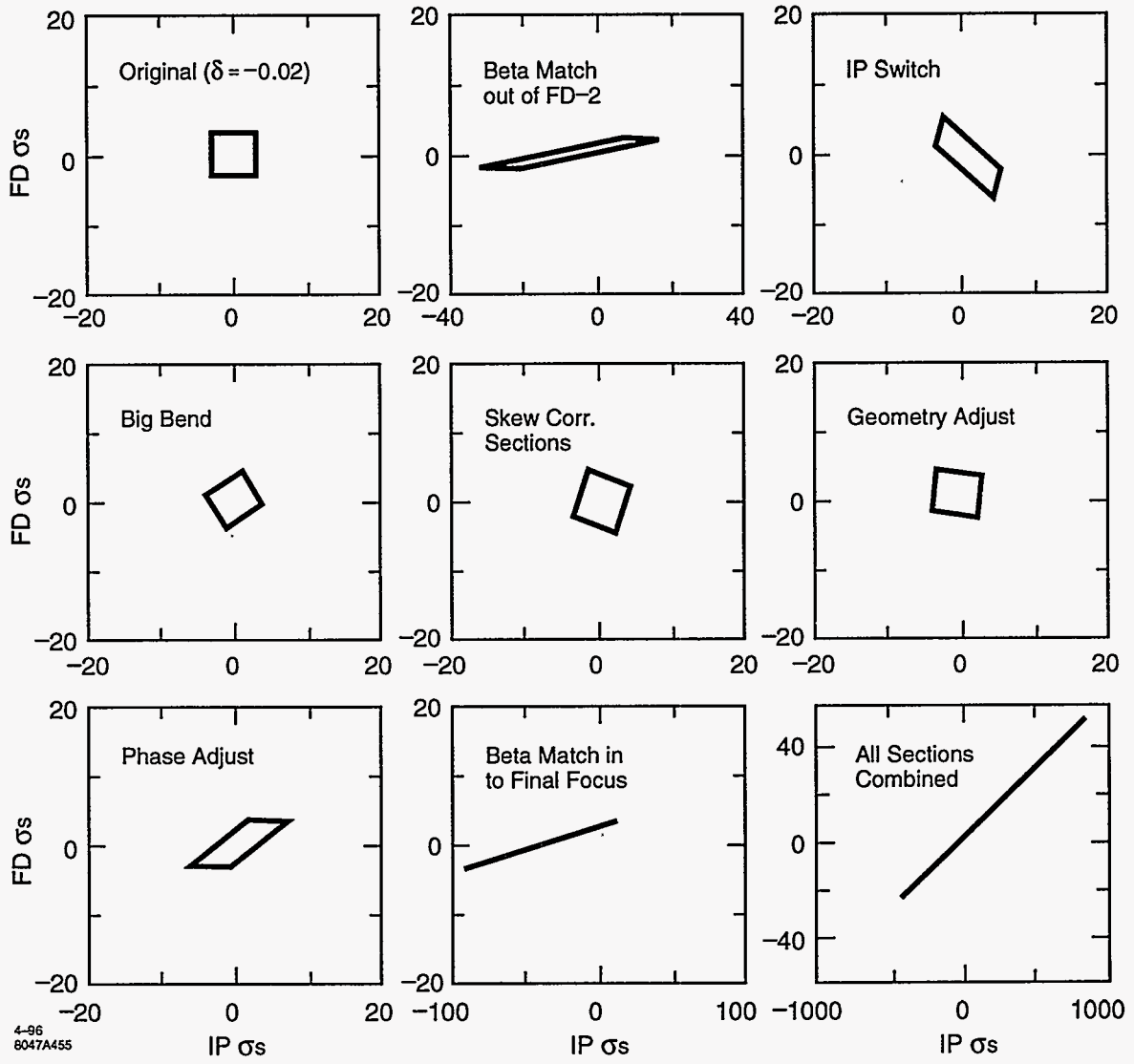


Figure 9-37. *This figure shows a sequence of maps for sections of the beam delivery system. Each picture corresponds to the image of a $3\sigma \times 3\sigma'$ piece of phase space with $\delta = -0.02$. The map is the composition of a linear map from IP to section entrance, then through the section, then a linear map back to the IP. The last picture is a map through all sections.*

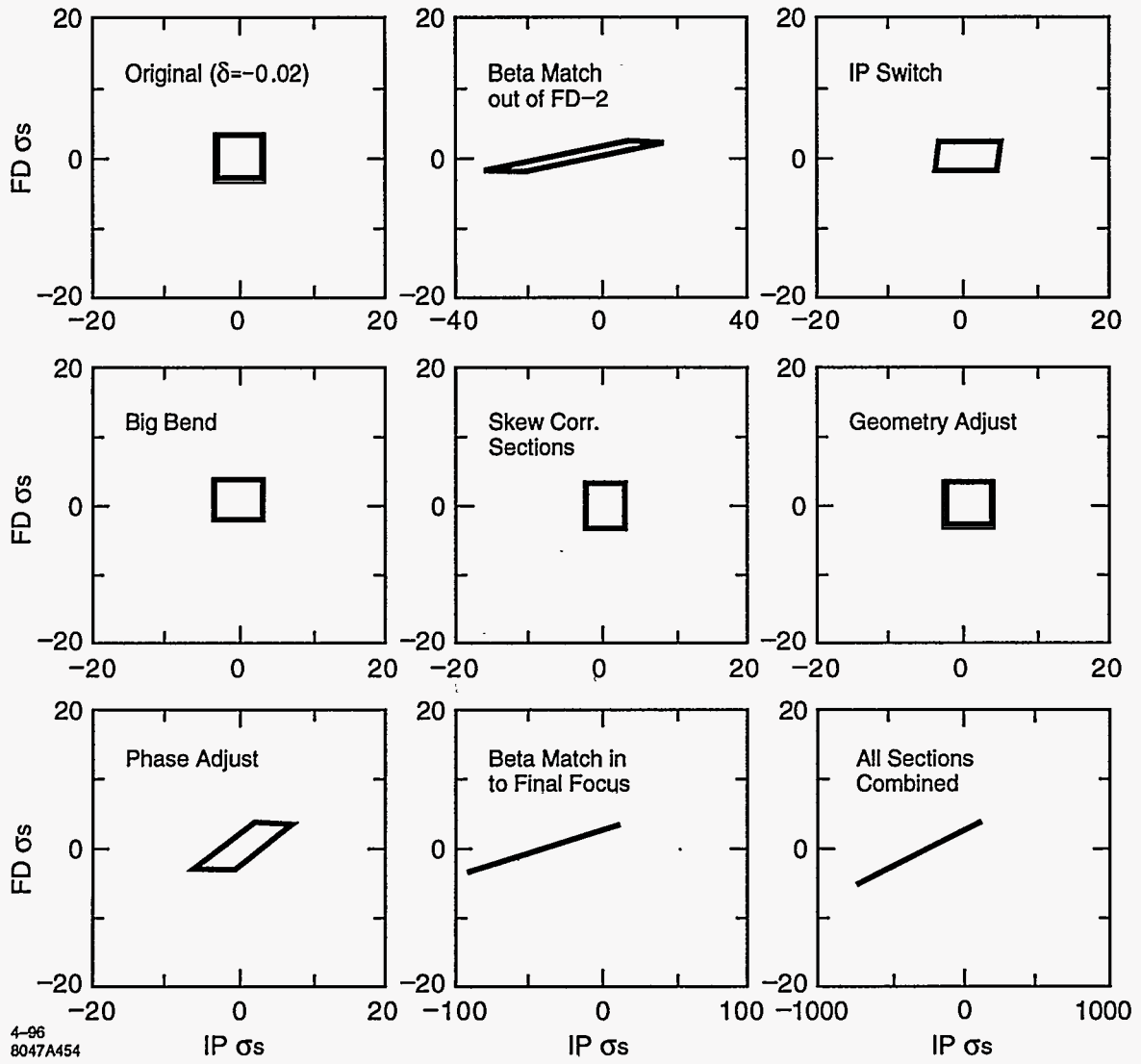


Figure 9-38. *This set of maps is the same as Figure 9-37 with the additional feature that the linear chromaticity and 2nd-order dispersion have been locally corrected. Note the dramatic change in the final map through all sections. (The scale has changed.) The FD phase is now less than 5σ and will pass through the collimator.*

Next-generation lattice. Following up on the mapping studies described in the previous paragraph, we have developed a next-generation lattice that incorporates the following features:

- Removes the FD stage-2 collimation section,
- Inserts sextupoles into the big bend to chromatically fix the big bend,
- Uses the big-bend sextupoles to also compensate the chromaticity of the skew-correction system and the final-focus angle-adjustment module,
- Uses the IP stage 2 sextupoles to compensate the chromaticity of the IP switch,
- Fixes the phase of modules with second-order dispersion so that the dispersion is always in the IP phase,
- Compensate the effects of interleaved sextupoles in the collimation-system modules with octupoles.

This lattice has been tracked and mapped, and has achieved the behavior described in Figure 9-38.

Power Deposition from Spoilers. Figure 9-39 shows the result of tracking particles initialized at the edge of the last spoiler in stage 1 with the angular and energy distributions of Section 9.2.2 [Cai]. We have carried out similar studies for all of the spoilers, and both edges of each. The results of this study show that about 1% of the particles survive to the first section of stage 2, and none hit the beam line after that [Drozhdin 1996].

A study of an earlier but similar beam line is shown in Figure 9-40 [Keller]. This figure shows a section of the beam line corresponding to the first half of one phase of one of the $-I$s in the collimation lattice indicating the placement of absorbers and spoilers. Note the horizontal and vertical scale. The beam pipe aperture is a small slit along the bottom of the frame. The horizontal and vertical spoiler are indicated by h/e and v, respectively, and the corresponding absorbers by H/E and V. Following the first h/e spoiler is a drift length $L = 37\,\mathrm{m}$, then a quadrupole pair with total $k_Q = 0.53$, then a second drift of length 37 m before the absorber H/E. Using the formula of Section 9.2.2 yields $\sqrt{(R_{12}R_{34})} = 55\,\mathrm{m}$. Following the first vertical spoiler (v) is a drift length of 37 m, then a quadrupole pair with $k_Q = .037$, followed by a drift of length 11.6 m, for which $\sqrt{(R_{12}R_{34})} = 43\,\mathrm{m}$. A photon dump will be required in the region of the four bends, indicated in Figure 9-42 by the letter BCXR, and an absorber will be located before QX1.2 as well as QX1.1.

The incident beam was 50 mm from the edge of the spoiler slit h/e_1. The quadrupoles are simulated by 8-cm-radius copper cylinders with a 0.25-cm-radius bore. The four 10-m dipoles were simulated by 5-cm radius copper cylinders with a 0.5-cm radius. The percent energy absorbed is indicated in Table 9-11.

For a 500-GeV beam energy, at 120 Hz, with $0.9\,10^{12}$ particles per bunch train, the average power in the beam is 8.4 MW. Assuming 1% is incident on a spoiler, downstream absorbers will need to accept a continuous 84 kW of power. The main absorber placed at the end of the second straight section following the spoiler will receive about one-half of this energy, hence it will need to disperse 40 kW of energy. About 20% of the energy exiting the spoiler is in the form of photons. A large number of electrons will pass through the bend and impact absorbers in the second half of the collimation section, or in the $3\pi/2$ section between collimation sections.

Power Deposition from "Virtual" Spoilers. There have been suggestions that it would suffice to have just one stage of collimation, omitting the second IP-stage section. This seems on the face of it improbable

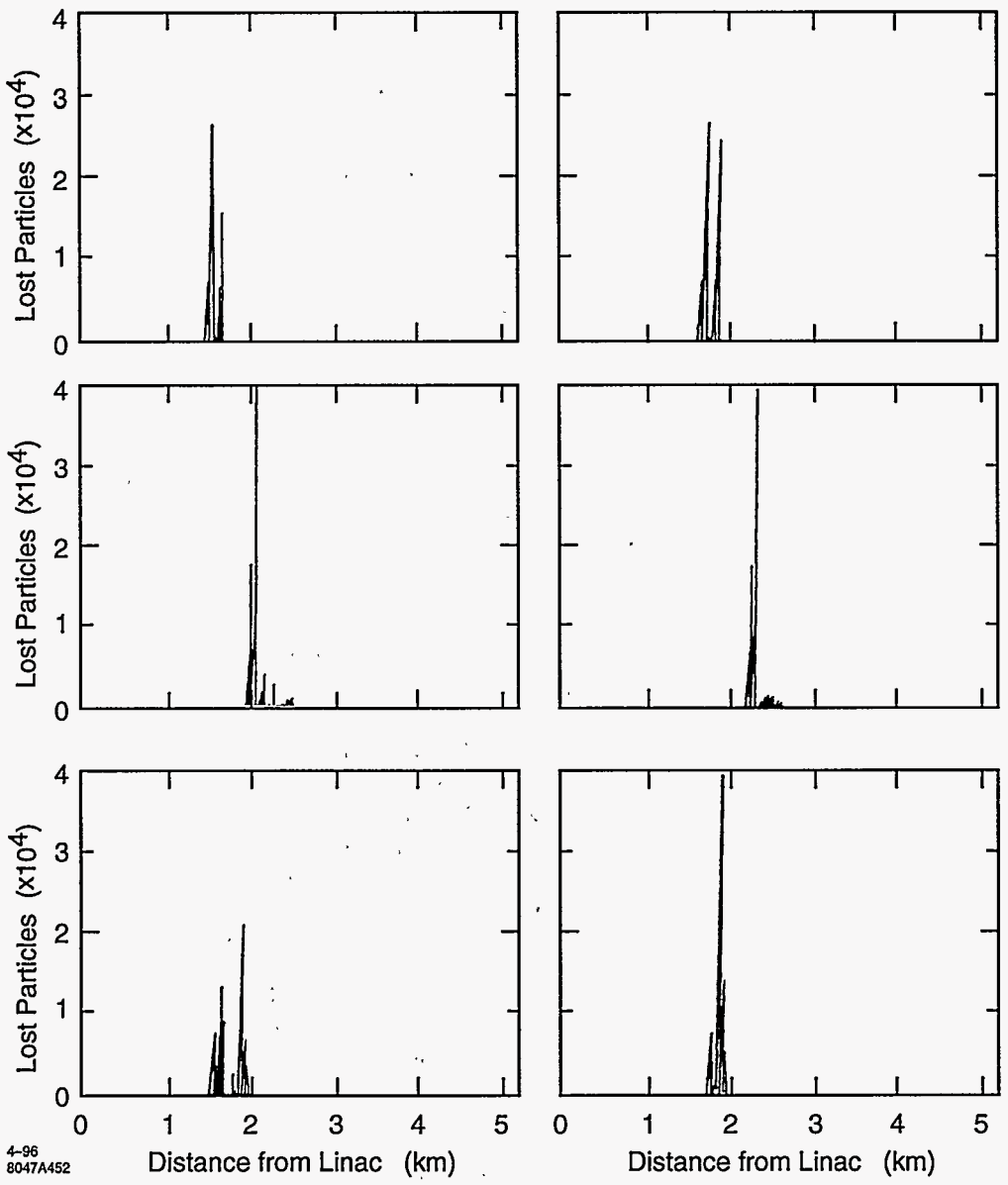


Figure 9-39. *These six pictures show the energy deposition from particles impacting the spoilers in the collimation system. The initial angular and energy distribution correspond to those shown in Figure 9-13–9-14. The six pictures are for the left or top edge of the six spoilers.*

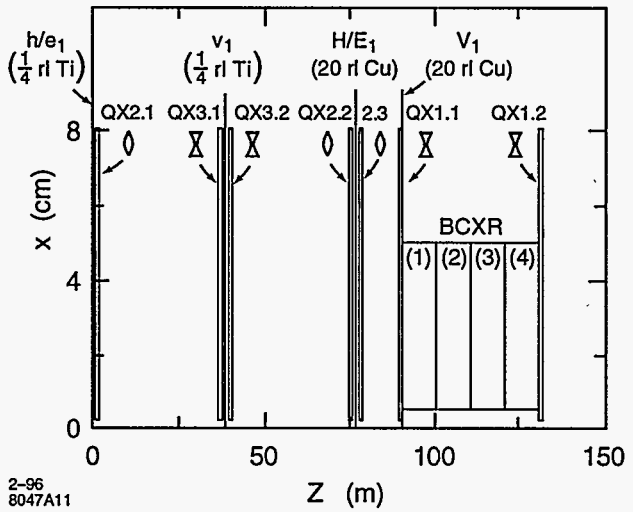


Figure 9-40. *A close-up view of a 150-m section of lattice, starting at the first horizontal spoiler. This physical arrangement was used for an EGS study of energy deposition in the 1-TeV-c.m. collimation system. The results of this study are presented in Table 9-11.*

Element	% Energy Absorbed
QX2.1	0.04
QX3.1	4.0
V1 (0.25 rl Ti)	0.1
QX3.2	1.0
QX2.2	30.1
HE1 (20 rl Cu)	36.0
QX2.3	0.4
QX1.1	0.2
V1 (20 rl Cu)	11.6
Thru V1	16.4

Table 9-11. *Results of an EGS run showing energy deposition downstream from the first spoiler.*

because, according to our assumptions, there are potentially 10^{10} particles impacting the last spoiler. Of these about 2/3 are expected to impact absorbers placed immediately downstream from the spoiler. Indeed tracking shows about 79% hit absorbers in this section. Another 19% are lost in the bend that terminates the dispersion and the beta-match out of the collimation system. That leaves 4 10^8 particles traveling toward the IP with large IP-phase amplitudes. If a particle is off-energy one would expect it to be lost in the bends. But according to the energy distribution of Section 9.2.2, 35% of the particles have an energy loss less than 4%. So many particles have energies near the beam energy, and the majority of the ones that get lost because of low energy are lost at the end of the collimation system. Only 0.5% are lost between the end of the collimation system and the beginning of the final-focus system. 1% are lost in the final-focus system and the remaining 2.5% pass through the IP. Since 1% represents 10^8 particles, this is two orders of magnitude larger than the number which can be collimated in the final-focus system. The only caveat to this statement would be a detector design that could veto 100 muons arriving parallel to the beam line.

Figure 9-41 shows the loss distribution along the beam line from "virtual" spoilers placed at locations in stage 2 where there would be spoilers if this was the only collimation stage [Cai].

9.2.6 Operational Issues

Orbit Stabilization

Changes in waist and skew aberrations that arise from orbit changes within the collimation system can be monitored non-invasively in the skew correction system (SCS) downstream, and corrections introduced. This strategy will relieve the tolerances listed in Table 9-9 in the sense that the stability tolerance can be considered to be for the time between skew and waist checks in the SCS rather than the time between skew and waist scans at the IP. Also the results are intrinsically better since the two beams are not intertwined here as at the IP.

However, it will still be important to occasionally monitor the orbit within the collimation system. This requires BPMs located in these sections. BPM performance will be adversely affected by the presence of intense backgrounds originating primarily from spoilers. Hence the BPMs must be carefully protected from the spray originating at the spoilers, or from any secondary particles that might be generated by this spray.

One can take advantage of the fact that the number of bunches in the beam-train and the emittance of a single bunch can be changed rapidly (a criteria that is important for the machine protection systems). If the collimators are designed so that they can be withdrawn rapidly, then the beam could be reduced to a single small-emittance bunch and the beam trajectory can be observed without the presence of spray and offsets due to beam loss. We note that it takes an acceleration of 2 10^{-3} g, where g is the acceleration of gravity, to move an object 5 mm in 1 second, with an original and final velocity equal to zero. The concern is that the collimator alignment accuracy not suffer in this process. Collimator alignment tolerances are typically 100 μm horizontally and 40 μm vertically, so this may not be a problem.

BPM Protection

Even with a single bunch it is important to protect and shield the BPMs. Because of their small length, and high resolution, and with an inner diameter that is twice the size of the quadrupole bore, the rf BPM [Shintake 1995] tested in the December 95 FFTB run should be easy to protect. One should choose the recess of the BPM and the shape of the protection collimator in such a way that particles could not reach or

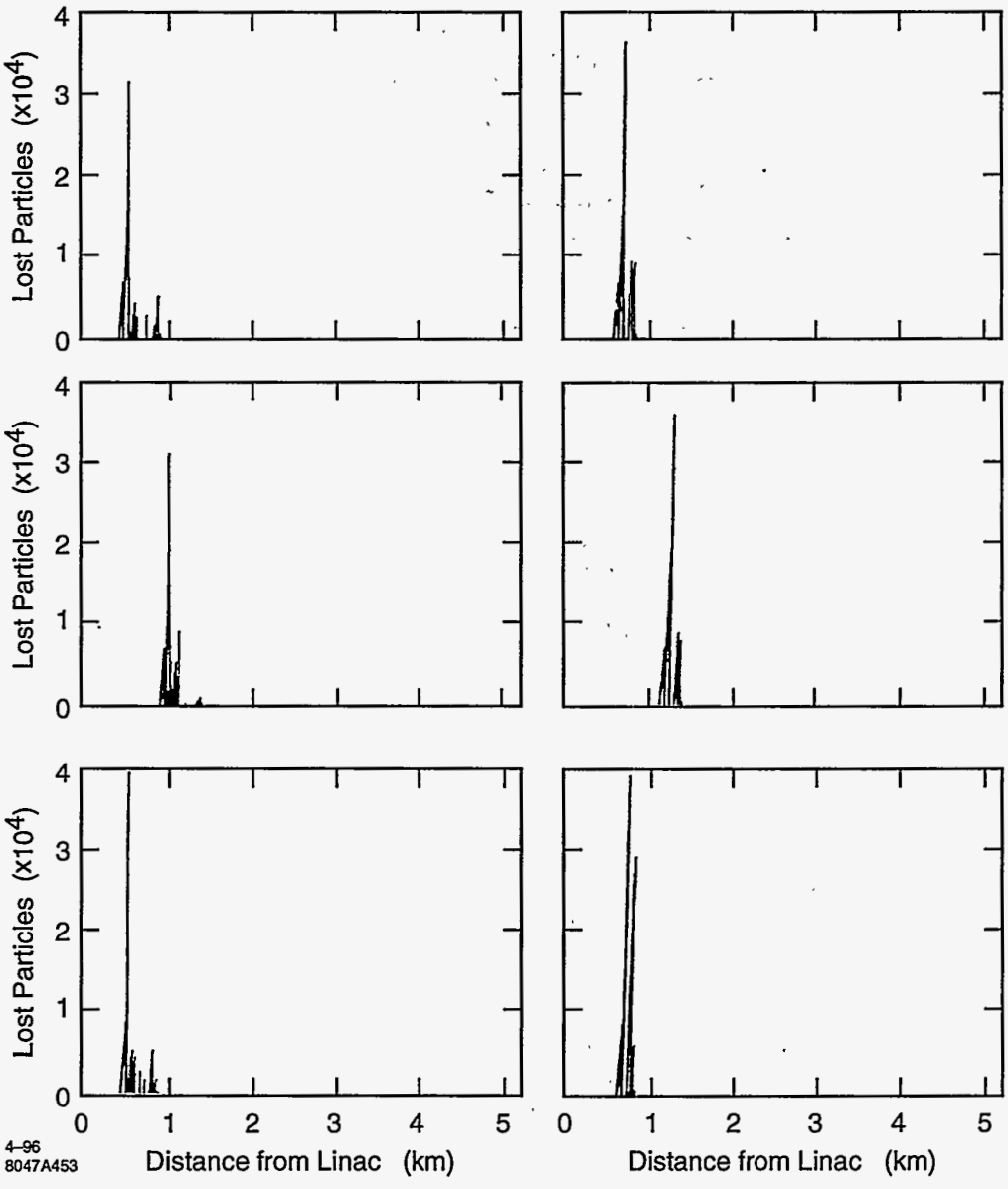


Figure 9-41. *These six pictures show the energy deposition from particles impacting "virtual" spoilers in stage 2 of the collimation system. The intent is to explore the consequences of having only 1 stage of collimation.*

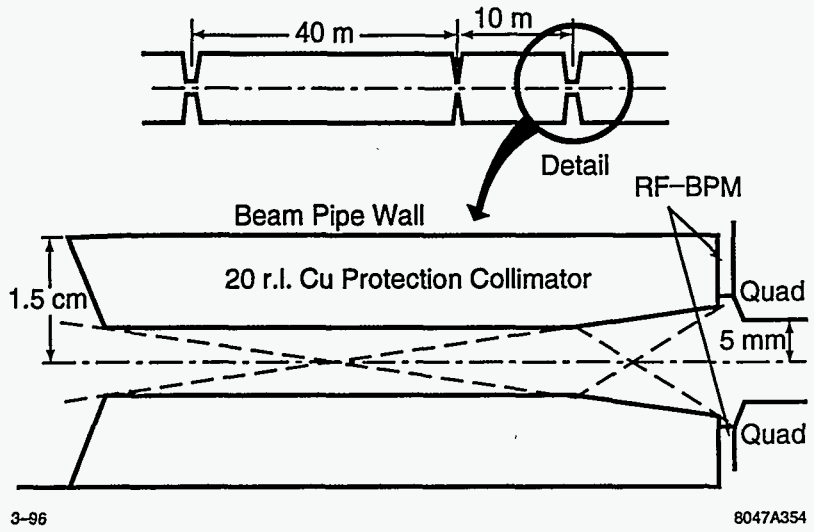


Figure 9-42. *A conceptual sketch of a possible protection collimator for BPMs of the collimation system.*

influence the BPM with a two-surface process (that is the beam would have to scatter or create a secondary electron at two distinct surface before reaching the BPM). The rf BPMs tested in the FFTB had an inner radius of 10 mm, and a resolution of 40 nm. Figure 9-42 shows a conceptual sketch of a protection collimator for a recessed rf BPM with an inner radius of 8 mm. This sketch is only intended to be suggestive, and does not represent a design. Research with rf BPMs has been very promising. It will be important to do studies of spray protection geometries to determine their effectiveness.

System Alignment

Magnet Alignment. It is presumed that all magnetic elements will be located on movers similar to the type used in the FFTB. These have step sizes of $1/3\,\mu$m with a total range of 1 mm. If the magnets are originally aligned to say $250\,\mu$m it is a straightforward task to align the magnets to tens of micrometers using beam-based techniques that rely on changing the current in each magnet. Such techniques were employed successfully in the FFTB.

Collimator Jaw Alignment. Collimator jaws can be aligned by changing the separation of the jaws while keeping the mid point between jaws fixed, and observing any centroid disturbance of the beam. As clarified in Section 9.2.5, the tolerance on collimator position is about 1 σ. For most collimators a 1-σ beam displacement produces about a 0.1-σ kick. In the first stages of the collimation system, with the large R_{12}s and R_{34}s to downstream stages, it is easy to detect these 0.1-σ kicks. Vertically 0.1 σ can be $4\,\mu$m, horizontally it is $30\,\mu$m. This will be detectable with the contemplated BPMs.

In the last stage of the collimation system this is more difficult. The typical beta functions in the big bend and skew correction system are 40 m, hence 0.1 σ is be $1.4\,\mu$m horizontally and $0.14\,\mu$m vertically. The former could be detected with an FFTB type strip line BPM. The latter could be detected with the rf BPM that has been tested. In principle there is no problem detecting these orbit changes. One has the additional advantage that one can observe the trajectory through the entire remaining beam delivery system, so there is a lot of redundancy.

The alignment of the final-focus system collimator is more sensitive yet. Here the maximum R_{34} to a downstream BPM will be much smaller, say 200 m. A 0.1-σ kick at the CCY collimators has a magnitude of 0.05 nr, and the downstream displacement will be 10 nm. It is expected that this tolerance can be achieved with an rf BPM since the first efforts resulted in a 40-nm resolution. This BPM would have to be placed at an optimum location in the final telescope.

We note that if the beam has a slight banana shape as a result of upstream wakes, this collimation alignment technique will be weighted to the head of the bunch, since the head is causing the wake on the tail. Any result coming from the banana part of the bunch would tend to move the collimator in the direction of the bunch. This will have the effect of causing a kick to the tail in the direction opposite to the banana tail. In a phase space diagram this kick is perpendicular to the tail in phase space.

Tuning and Stabilization

Since the IP-phase quadrupoles are in phase with the IP they have little effect on the IP beam size, hence the tolerances of the two stages of IP-phase collimation are much looser than the tolerances on the FD-phase collimation, and so we shall concentrate on the FD phases. Furthermore, stage 2 of the FD-phase collimation is optional, since the beam has been already well collimated by previous stages, and since the β functions are smaller, the tolerances are looser than stage 1. Thus we can focus our attention on stage 1 of the FD-phase collimation. We will assume that the BPMs have the required resolution and have been protected from spray.

There are five quadrupoles in each of two sections of this stage. Each section is separated from the other by π phase advance. The first section has one horizontal spoiler and one vertical spoiler; the second section has one horizontal spoiler. If we assume that 1% of the beam is indeed intercepted at the spoilers, then since their gaps are about 2 mm, a change in the beam centroid of 20 μm can occur. Since the tolerance on the centroid position at the sextupole located next to the vertical spoiler is about 0.4 μm, we must determine the vertical centroid to a factor of 50 better than the change due to beam scraping. Thus stabilization to maintain the tolerances of Section 9.2.4 is not possible when the beams are being collimated to this extent.

As mentioned above, either the aberrations caused by a change in the collimation system orbit can be measured out and compensated in the skew correction section, or the collimators can be withdrawn after reducing the train to a single bunch. In the latter case methods similar to those described in Section 11.5.6.

Another strategy is to insure that the elements (strengths and positions) of the beam line are stable without any intervention. This is always a good course of action if possible. This strategy is greatly enhanced by the ability to non-invasively monitor the waist and skew aberration in the skew correction system (SCS) which follows the big bend. In this case the required stability times is a fraction of an hour rather than an hour. The stability tolerances on the bends are all a few times 10^{-5} which is achievable and should not cause a problem. The smallest vertical beam position tolerance at the sextupole is 0.3 μm, which could result from the motion of a quadrupole by abut 0.1 μm. Recent FFTB measurements show a beam-line element motion of about 0.1 μm per hour, which would suggest a drift of less than 10 nm in a six-minute stability time [Assman]. The FFTB measurements require further analysis and a model development to extrapolate to longer beam-line lengths. Since the FFTB line is located in a concrete block structure in the research yard at SLAC, not 20 m underground in an environment that should be thermally more stable, it appears we can achieve the required stability without intervention. (Recall that the tolerances were based on 1/2% luminosity loss for the skew and waist aberrations from the entire collimation system.)

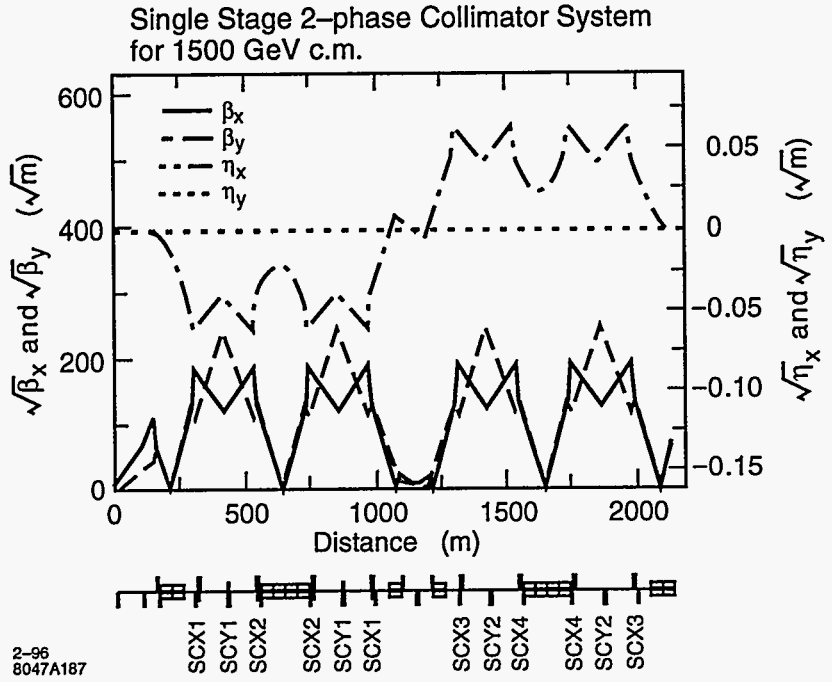


Figure 9-43. *The lattice functions for a 1.5-TeV-c.m. collimation system.*

9.2.7 Energy Scaling Laws

As we indicated in Section 9.2.4 the $\overline{n}$ are independent of energy, and since the equation for the gap is independent of energy, so is g_0. It follows that the spot sizes at the collimators are independent of energy but that β_0 must increase linearly with energy. Since energy deposition at the surface of a collimator is independent of energy, the requirements of beam-size to insure spoiler survival remain the same.

According to this scaling the beam divergence at the spoilers will scale as $\sqrt{(\varepsilon/\beta)} \sim 1/\gamma$. But the rms angle in the multiple scattering formula also varies as $1/E$, so the scattering angle to beam divergence ratios remain unchanged. Thus many of the considerations of particle counts we have presented in Section 9.2.3 above remain unchanged.

A Collimation System for 1.5-TeV c.m.

A lattice for a 1.5-TeV-c.m. system is shown in Figure 9-43. This lattice was designed so that the horizontal and vertical planes could be collimated at the same location, where the vertical β-function is a maximum. At this energy, collimating each phase one time requires a beam-line length of 2 km. As described in the text, the optimum system will have 2 IP phases and 1 FD phase of collimation. The second IP phases can have a smaller β-function. If more beam-line length is required we can extend the collimation system in to the linac tunnel.

Figure 9-44 shows the band pass and emittance growth for this lattice. Both are satisfactory.

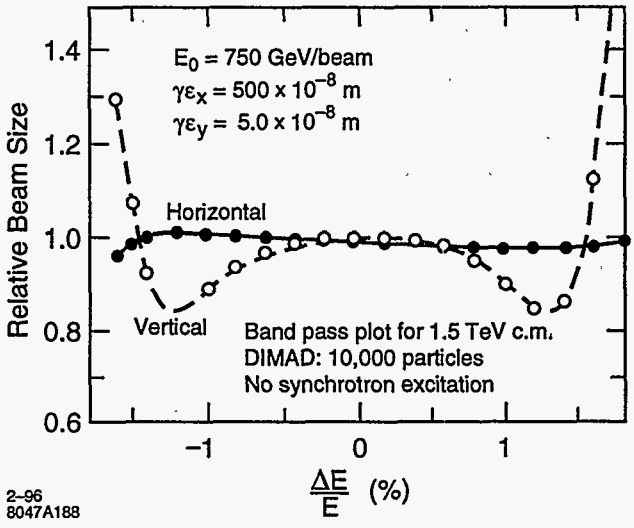


Figure 9-44. *The band pass plot for the 1.5-TeV-c.m. collimation system.*

9.2.8 Nonlinear Collimation Systems

Nonlinear collimation systems were proposed by Merminga *et al.*, [Merminga 1992] and described in several papers. The collimation system proposed for the JLC was also nonlinear and described in the JLC report [Oide]. It had a new interesting feature that the horizontal and vertical planes were collimated at the same time. Before beginning the design and analysis of the linear collimation system described here we assessed the results of these papers. We were impressed that the lengths of the nonlinear systems were not obviously shorter, and either the sextupoles were impossibly strong, or the beta functions were as large as linear system beta functions. This led us the conclusion that it would be best to begin with the more conservative linear system, pursue an optimization of that design, calculate its tolerances and assess its operational behavior, and then return to the nonlinear system for a comparison.

Now that the linear design has reached some maturity, it is appropriate to turn to the nonlinear systems, or to a combination of linear and nonlinear designs, and carry out the above intentions. Unfortunately that has not been completed.

9.2.9 Summary and Conclusions

We have described the specifications, discussed the relevant properties of materials, presented spoiler and edge-scattering distributions, clarified the relevant wake-field of tapered collimators and defined the optimum choice of collimator shape, quantified the impact of near-wall wakes for on-axis and mis-steered beam, designed lattices that implement the required functionality, calculated the position and strength tolerances of the magnetic elements, clarified the impact of ground motion, tracked the lattices to confirm their functionality, tracked spoiler distributions in the collimation system to determine power deposition, traced edge-scattered distribution in the final-focus system to determine the probability of particle impacts on the final doublet, determined the extent of tail repopulation due to gas scattering, discussed the operations problems, and addressed machine protection issues. To our knowledge all issues have been addressed with

the satisfactory conclusion that it is possible to build and operate a collimation system for the proposed beam parameters from beam energies of 175 GeV to 750 GeV, that will collimate the beam at apertures that are required by the final-focus system and final-doublet apertures. We are confident that we have an existence proof. However the systems are long and they are sensitive.

Because of the system length and its delicate nature, we feel that further research is warranted in four directions:

1. As the machine protection system (MPS) is better determined and clarified, it may evolve that the passive protection requirement we have assumed is not warranted by the magnitude of additional risk avoided. The second level of the MPS already requires careful monitoring of all system changes and limiting the range of change that can occur in fast systems. At the present moment the only fast change, which to our knowledge can occur upstream of the collimation system, is the sudden dropping of klystron families, resulting in a sudden large energy change. This is a special kind of change that must already be addressed in the linac beam diagnostic sections (chicanes). If the requirement of passive collimator protection were dropped, the collimation system would become much shorter and less delicate.

   It would be also worthwhile to better determine the worst-case tail population expected from the linac. Is it larger or smaller than the number we have assumed? This does not alter the lattice design, but does impact the power that must be handled by the absorbers.

2. Experiments should be carried out to verify the wakes of tapered collimators.

3. Nonlinear systems and combinations of linear with nonlinear systems should be studied. It is time to search for the "optimum" collimation system.

4. Improvements in the present system should be pursued, including a redesign to meet recent beam parameter changes.

5. Finally, non-conventional collimation ideas should be pursued. Though they have been discussed, it would be a worthwhile study to clarify the impracticality of all wild ideas, with the hope that one will actually survive.

9.3 Pre-Linac Collimation

The pre-linac collimation system is intended to remove the low energy, 10-GeV, contribution to the beam halo produced from the electron/positron source, damping ring, pre- accelerator and bunch compressor. The pre-linac collimation system should be able to collimate continuously 1% of the beam, 10^{10} particles, at the IP and FD phase in both planes and must withstand one full mis-steered bunch-train. Using simple energy scaling laws the post-linac collimation-system design can be scaled to produce the pre-linac design. There are some differences based on requirements for the dispersion function, and with synchrotron radiation in dipoles.

9.3.1 Pre-Linac Collimation Function:

Halo particles before the main linac are present upon extraction from the damping ring and generated in the bunch compressor. If not removed this low energy halo will be injected into the main linac and accelerated

up to full energy. By placing a collimation system at the entrance to the main linac the amount of this low-energy halo accelerated to full energy will be minimized easing the load on the final collimation system, and reducing potential background in the IR region. The strategy used to achieve the needed passive protection is "spoiler followed by absorber." The spoilers will be constructed of titanium. The spoilers function is to increase the angular divergence of an accidentally mis-steered beam so that it can be intercepted by the downstream copper absorbers.

9.3.2 Pre-Linac Collimation System Requirements:

To match the post-linac specifications, the beta functions at the collimators must collimate at a depth of 6 σ_x, and 35 σ_y. The collimator jaw gap is determined by two conditions. The first is the spoiler survivability which requires that $\sqrt{(\sigma_x\sigma_y)} \leq 100\,\mu$m so that the spoiler does not exceed its tensile strength and break due to thermal expansion. The second condition on the jaw gap is determined from minimizing the wakefields induced by the collimator jaws. The equation for the optimal collimator jaw gap is independent of energy so the optimum pre-linac collimation jaw gap will be the same as the post-linac gap, approximately 1 mm.

In order for the spoilers to survive it is required that the thermal stress created by the temperature rise remain below the tensile limit of the spoiler for one mis-steered bunch train. A titanium alloy has been found to offer the best performance and it can be coated with pure titanium or titanium nitride to reduce resistive-wall wakefield effects. Following the arguments for the temperature rise and tensile strength requirements for the post-linac collimation system one can scale the results for the pre-linac collimation system. The energy deposition at the surface of a material is independent of energy above 1 GeV, hence the instantaneous heat load at 10 GeV is the same as 500 GeV. The heat load in the absorber however will simply decrease from 84 kW to 1.7 kW due to the fact that the beam has less energy to deposit. Hence the spoiler material and thickness remain the same as the 500-GeV design and the absorber design can be relaxed.

The beam divergence varies as $1/\gamma$ and the rms multiple scattering angles also vary as $1/\gamma$, thus the ratio of scattering angles to beam divergence remain constant. The energy scaling laws show that much of the previous work done on the post-linac collimation system can simply be carried over for the pre-linac collimation system. A detailed discussion of heat loads, multiple scattering, absorber properties, spoiler transmission, wakefield tolerances, etc., can be found in the post-linac Section 9.2 Energy scaling is specifically discussed in Section 9.2.7.

9.3.3 Pre-Linac Collimation Optical Design:

The optical design follows the post linac design utilizing an interleaving of the horizontal and vertical scraping to minimize the collimation section length. The beam is collimated at the IP and FD phase in both planes one time. Using the constraint $\sqrt{(\sigma_x\sigma_y)} \leq 100$ mm for spoiler survival the beta functions at the design emittances must satisfy the condition $\sqrt{(\beta_x\beta_y)} \leq 80$ m at 10 GeV. Figure 9-45 shows a 10-GeV lattice which satisfies these conditions. The spoiler and absorber gaps are about 1 mm. There are two subsections each having a set of horizontal and vertical spoilers and absorbers each separated by $-I$ transformations. The two subsections are separated by a $5\pi/2$ transform.

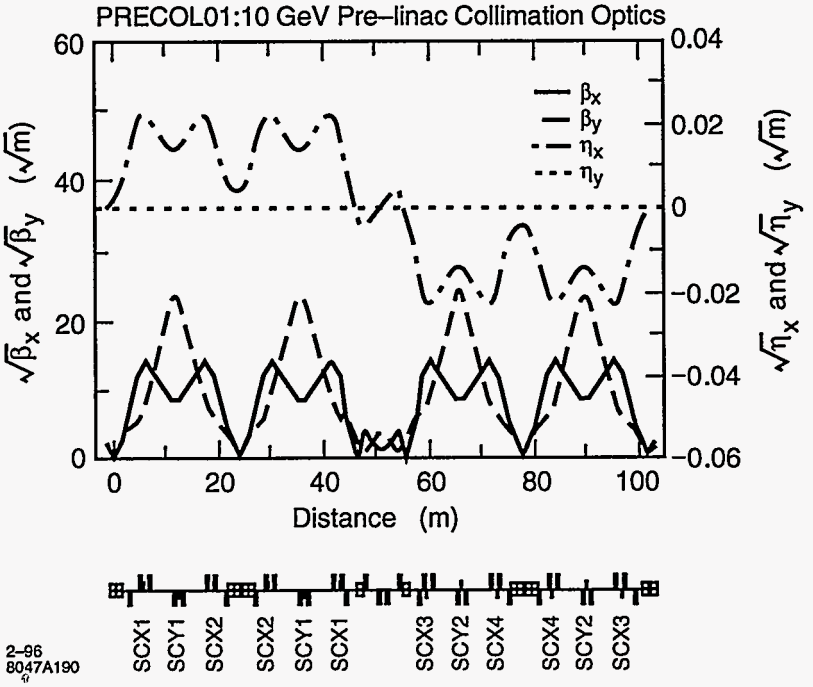


Figure 9-45. *The pre-linac collimation system lattice.*

9.3.4 Pre-Linac Collimation System Bandwidth

The pre-linac collimation system bandwidth is shown in Figure 9-46. Though it is larger than the post-linac bandwidth, the energy spread at the beginning of the linac is quite large because of the second bunch compressor. The bandwidth should be three times this large. This problem must be addressed, either by attempting to design a collimation system with 4.5% bandwidth or locating the collimation system downstream in the linac after the energy spread is reduced. The latter option, though seemingly attractive, has the problem that the first 600 m (to 30 GeV) of the linac is used to induce a coherent energy spread to facilitate BNS damping in the linac. Hence at 30 GeV the energy spread is still 1.5%, and it would be necessary to place the collimation system at 90 GeV. This is nine times the energy of the existing design, so the system would have to be redesigned for that energy. Since pre-linac collimation is optional we have not investigated this subject further.

9.3.5 Summary

We haved presented a design for an optional 10-GeV pre-linac collimation system. This system is a scaled-down version of the 500-GeV post-linac collimation system. The beta functions were scaled so that the spoiler survival criteria of $\sqrt{(\sigma_x \sigma_y)} \leq 100\,\mu$m, a 1% beam halo collimation, and a minimal wakefield condition are satisfied. The system uses the same quadrupole magnet design as the post-linac system and its length is 100 m. The design presented has a bandwidth of 1.5% which is too small to be located immediately after the bunch compressor. If pre-linac collimation is deemed necessary this system will need to be improved upon.

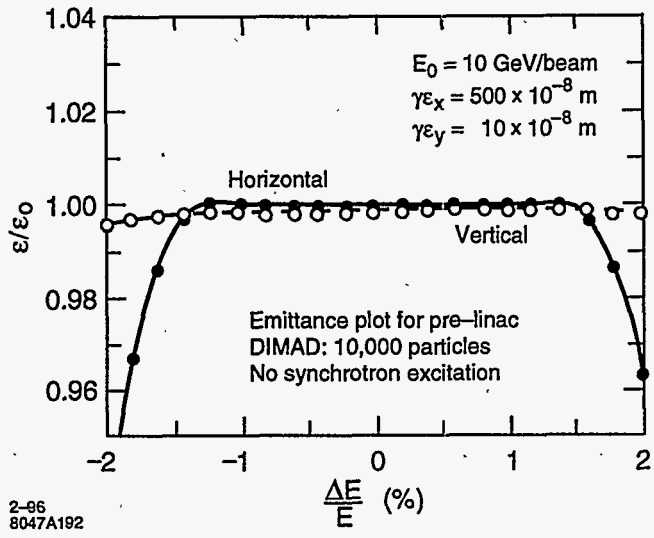


Figure 9-46. *The pre-linac emittance growth.*

9.4 Bunch Length Collimation

9.4.1 Introduction

In order to achieve high luminosity, the NLC requires very short bunches. The three primary considerations that determine the bunch length are: the bunch length should not be much larger than β_y^*, the emittance growth due to the transverse wakefields increases with the bunch length, and the energy spread due to the longitudinal wakefield decreases with increasing bunch length. For the NLC parameters, the optimal bunch length is between 100 and 200 μm.

When generating these bunches in the bunch compressors, one typically populates long bunch length tails. These tails will be deflected to large transverse amplitudes by the transverse wakefields. In addition, the long longitudinal tails will be converted into energy spread as they go through the main linac. In the post-linac collimation system and the final-focus system, the energy tails will generate transverse halo due to chromatic effects. Both sources of transverse tails will create unacceptable backgrounds in the detector if not collimated. Therefore, the longitudinal tails have to be collimated before entering the final focus. It is better, and presumably easier, to collimate them before the main linac.

For these reasons we decided to design a bunch length collimation system for the NLC. The system should be located before or at the beginning of the NLC main linac, and remove particles beyond $3\sigma_z$. The usual method to cut longitudinal tails is to convert them to transverse tails and remove them with a collimator. The ideal case is when the bunch length has a monotonic energy distribution. Then the energy spread can be transferred into horizontal displacement with dispersion.

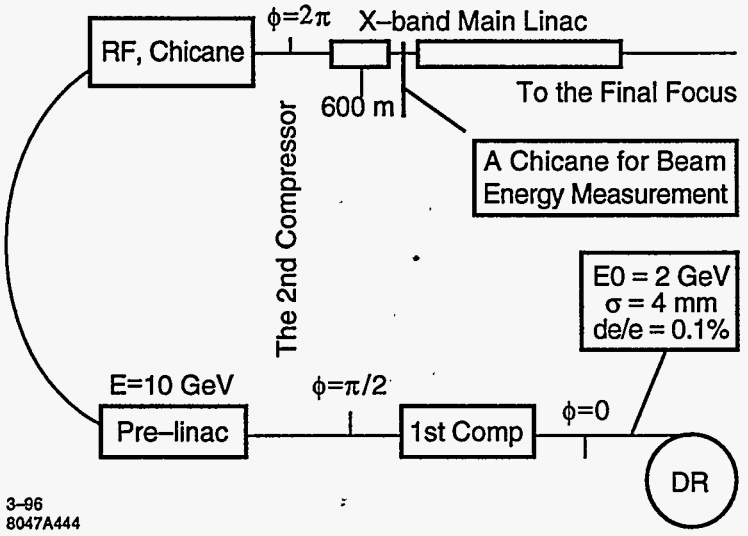


Figure 9-47. *The layout of the NLC.*

9.4.2 Collimator location

Figure 9-47 is the layout of the damping ring to the final focus of the NLC, showing the damping ring, two steps of bunch compression and the X-band main linac.

The longitudinal phase $\Phi = n\pi$ is the necessary phase to place the collimator. Hence we may choose the first collimation at the exit of the damping ring. Here, the energy spread is small (0.1%), and for $3\sigma_z$ bunch length collimation, it will need to cut beyond $+/-0.3\%\Delta E/E$. Assuming the collimator half gap equals 1.5mm and $\sigma_e/E = 0.1\%$, then a dispersion $\eta = 0.5$m is needed. Since there is already a spin rotation system at this location, the 0.5-m dispersion is available.

A second bunch-length collimation could be put after the second bunch compressor. Due to the space limitation, we choose to put this collimator after the first 600m of acceleration in the X-band main linac, where there is a chicane for beam energy measurement.

9.4.3 Particles Loss at Bunch Length Collimators

We put the two bunch length collimation systems into the present two-step compression system and the main-linac lattice design. Table 9-12 presents the expected particle loss at the collimators with a cut at $2\sigma_e$ [Raubenheimer 1994; Zimmermann].

The following figures show the bunch shape at the end of the main linac as calculated by a modification of the LITRACK program [Bane].

Collimation location	Beam loss %
1st collimation only	4%
2nd collimation only	2%
1st and 2nd together	5%

Table 9-12. *Particles loss percentage at the collimation systems.*

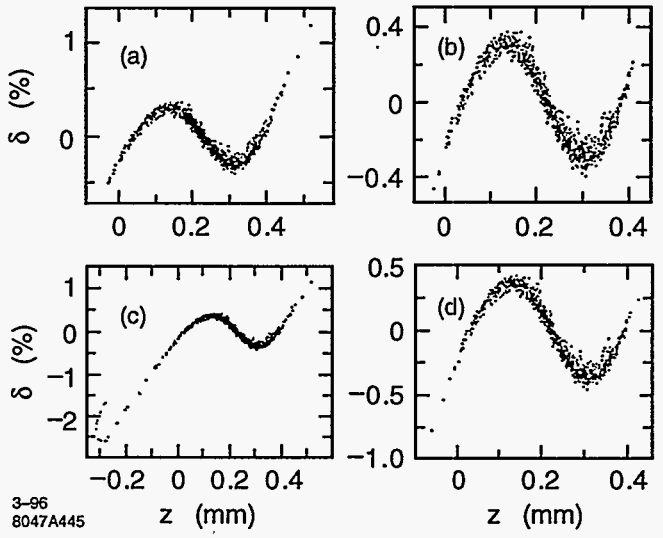


Figure 9-48. *(a) The bunch shape without collimation. (b) The bunch shape with the first collimation only. (c) The bunch shape with the second collimation only. (d) The bunch shape with both bunch length collimation systems.*

9.4.4 Collimation in the Linac

It is important to have the beam energy be longitudinally monotonic, so that the bunch length collimation can be made by energy tail cuts. To meet the above requirement, we can put the beam behind the rf crest, so that only 10% of the particles are beyond $4\sigma_z$ and have energy deviation greater than $4\sigma_e$. For a collimator half gap equal 750 μm, and $\sigma_e/E = 1\%$, a dispersion of $\eta = 2.5$ cm implements a $+/-3\sigma_e$ energy tail cut.

At the beginning of the main linac, the beam bunch length is 100 μm. The energy distribution may be calculated with the formula:

$$E(z) = E_0(z) + V_{\rm rf}\cos(\Phi_0 + kz) - \int W_L(z'-z)\rho(z')dz' \quad . \tag{9.86}$$

Here $E_0(z)$, is the initial energy distribution at the beginning of the main linac, with $< E_0(z) >= 10$ GeV. $V_{\rm rf}$ is the main linac rf peak voltage. Φ_0 is the center particles' accelerating phase, and k is the wave number of the rf wave form. $W_L(z'-z)$ is the main X-band linac longitudinal wakefield, which has been calculated by K. Bane. $\rho(z')$ is longitudinal distribution, which we will assume to be Gaussian. The minus sign indicates beam energy loss.

For the three NLC designs, when $\Phi = 16°$ behind the crest of the rf phase is chosen, the beam energy spreads are about 1%. If we set the collimator to cut $+/-2\sigma_e$ (*i.e.*, $+/-2\%$), less than 10% particles will be cut.

	NLC-I	NLC-II	NLC-III
Rf phase	16°	16°	16°
Energy gain	33 MV/m	55 MV/m	70 MV/m
In 600 m	20 GeV	33 GeV	42 GeV
# of particles	$7\ 10^9$	$11\ 10^9$	$14\ 10^9$
Energy spread	1.0%	1.2%	1.28%
# of particles	$1.54\ 10^8$	$3.30\ 10^8$	$7.84\ 10^8$
(%) Cut	(2.2%)	(3.0%)	(5.6%)

Table 9-13. *Some parameters for the three NLC design energies.*

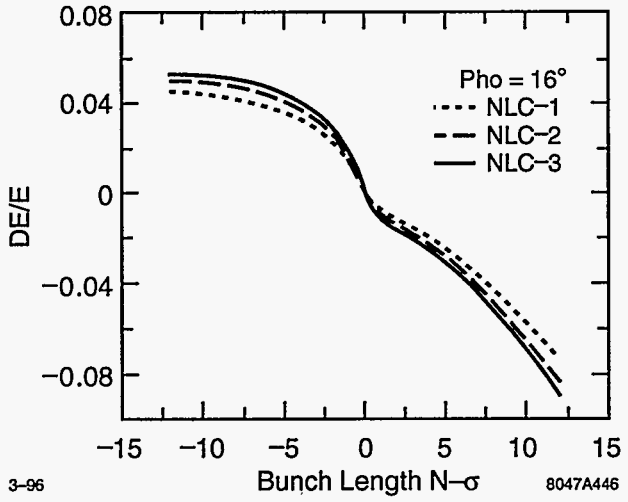


Figure 9-49. *Energy distributions for the three NLC design phases.*

Table 9-13 lists some parameters and percentages of particles that are lost at the collimators. Figure 9-49 shows the energy distribution along the bunch.

9.4.5 Conclusion and Discussion

It is possible to have two bunch-length collimation systems in the NLC design lattice. The bunch has a nearly monotonic energy distribution when we choose the rf phase to be 16° behind the rf crest for all three NLC design phases. The energy spreads are about 1%. If we decide to collimate $+/-2\sigma_e$ energy tails, there will be less than 10% of particles cut.

Before determining the utility of the collimators, a number of additional questions remain to be answered. In particular, we need to study the longitudinal and transverse wakefields induced by the collimators and related tolerances.

Our results show that the collimator at the exit of the damping ring is relatively straightforward and effective. Because of a non-linear $\delta - z$ relationship, and therefore a nonlinear $x - z$ dependence at the collimator, the second collimation is more difficult and less beneficial.

References

[Assman] R. Assman and C. Montag have studied quadrupole motion in the FFTB using the DESY wire system. Data analysis is in process.

[Bane] K. Bane wrote the program LITRACK (1994). Undocumented.

[Brown 1979] K. Brown and R. Servranckx, "Chromatic Corrections for Large Storage Rings", SLAC–PUB–2270, in *IEEE Trans. Nucl. Sci.* **26**, 3598 (1979).

[Cai] Y. Cai performed the tracking studies for this chapter using a C++ mapping and tracking code he has written called LEGO.

[Chao 1992] A. Chao, "Coherent instabilities of relativistic bunched beam", SLAC–PUB–2946 (1982).

[DeStaebler-Walz] Spoilers followed by absorbers have been in use for many years at the Stanford linac.

[DeStaebler 1994] H. De Staebler first pointed this out to us. Private communication.

[Drozhdin 1996] A. Drozhdin and N. Mokhov, "NLC-1000 Beam Collimation System Simulation", Jan. 18, 1996. The authors have prepared distributions similar to those presented in the text generated by Y. Cai. We thank them for supplying us with a preprint of this study they did of the NLC collimation system.

[FFTB BPMs] The FFTB strip line BPMs have a resolution of 1 mm. The stability of the zero point has not yet been determined. According to S. Smith the one-hour stability could conceivably be a factor of ten better than the resolution. A proposal has been submitted to study the zero point drift of these BPMs.

[Helm] All of the collimation lattices have been prepared by R. Helm. Lattice angle control was studied by Y. Nosochkov.

[Irwin 1991] J. Irwin, N. Merminga, R. Helm, and R. Ruth, "Optimizing a Nonlinear Collimation System for Future Linear Colliders", SLAC–PUB–5507, in *Proc. 1991 Part. Acc. Conf.* (1991).

[Keller] L. Keller provided this tracking study.

[Merminga 1992] N. Merminga, R. Helm, J. Irwin, and R. Ruth, *Particle Acclerators* (1992).

[Nelson 1993] R. Nelson has provided the EGS simulations described in this chapter. See The EGS Code System, SLAC–PUB–265 (1985).

[Oide] JLC collimation system, Section 4.8.6, "JLC-I", KEK Report 92–16 (1992).

[Piwinski] A. Piwinski, "Wake Fields and Ohmic Losses in Flat Vacuum Chambers", DESY–HERA–92–04 (1992).

[Rossi 1952] B. Rossi, *High-Energy Particles*, (Prentice-Hall 1952).

[Raubenheimer 1995] T. Raubenheimer has carried out these first estimates of expected tail populations. Further detailed studies are recommended. Private communication (1995).

[Raubenheimer 1994] T.O. Raubenheimer, "Bunch Compressor Parameters", NLC–Note–2 (1994).

[SLC] Recent experiments with collimator wakefield kicks are indicating that the wake kick is about three times larger than predicted by theoretical formulas. It appears to have the gap behavior appropriate to a resistive-wall wake, but a geometric source is not ruled out.

[Shintake 1995] Rf BPMs with a resolution of 40 nm were demonstrated in the Dec. 95 FFTB run.

[Smith 1995] Private communication. Steve Smith designed and built the BPMs for the FFTB.

[Stupakov 1995] "Geometric Wake of a Smooth Taper", SLAC–PUB–7086 (Dec. 1995).

[Stupakov 1996] This limit was computed by G. Stupakov. Private communication.

[von Holtey] G. von Holtey suggested the use of curved collimators for this reason. G. von Holtey, "LEP Main Ring Collimators", CERN/LEP–BI/87–03.

[Walz 1973] D. Walz, D. Busick, T. Constant, K. Crook, D. Fryberger, G. Gilbert, J. Jasberg, L. Keller, J. Murray, E. Seppi, and R. Vetterlein, "Tests and Description of Beam Containment Devices and Instrumentation–A New Dimension in Safety Problems", SLAC–PUB–1223 (1973).

[Walz 1991] D. Walz has provided the information on properties of materials used in this chapter.

[Walz 1992] "NLC Final-focus Collimation and Dumping", Working Group Talk, Final Focus and Interaction Region Workshop, (1992).

[Warnock] According to Warnock the resistive-wall wake is not given by this integral in perturbation theory (for small disturbances of the beam wall). Private communication. However, we know of no better approximation to use for these deep tapers.

[Yan] Y. Yan has created and analyzed the maps used to analyze the chromatic behavior of the beam delivery beam line.

[Yokoya 1988] "Impedance of Slowly Tapered Structures", CERN SL/90-88 (AP).

[Yokoya 1991] K. Yokoya and V. Telnov prepared estimates of this at the LC91 workshop in Protvino, USSR. The argument here follows a transparency prepared by K. Yokoya.

[Yokoya 1995] K. Yokoya has pointed out the importance of the on-axis wakes for parallel plate collimators. We thank K. Bane for communicating these results.

[Zimmermann] F. Zimmermann, "Longitudinal Single-bunch Dynamics and Synchrotron Radiation Effects in the Bunch Compressor", NLC–Note 3 (1994).

Contributors

- K. Bane
- Y. Cai
- A. Drozhdin
- R. Helm
- L. Keller
- J. Irwin
- R. Messner
- R. Nelson
- T. Raubenheimer
- G. Stupakov
- F. Tian
- D. Walz
- M. Woodley
- Y. Yan
- F. Zimmermann

10

IP switch and big bend

Contents

10.1 Introduction 644
10.2 The IP Switch 644
10.2.1 Optics Design 644
10.2.2 Chromatic Emittance Dilution 644
10.2.3 Synchrotron Radiation 646
10.2.4 Tolerances 646
10.2.5 Diagnostics and Correctors 648
10.2.6 Beam Correction Issues 648
10.2.7 Other Issues 648
10.3 The Big Bend 648
10.3.1 Optical Design 649
10.3.2 Chromatic Emittance Dilution 651
10.3.3 Synchrotron Radiation Effects 651
10.3.4 Tuning, Tolerances, and Corrections 652
10.3.5 Spin Transport and Depolarization 654
10.3.6 Vacuum System 655

10.1 Introduction

The IP switch follows the main linac and collimation section and allows slow switching between multiple IPs. The big bend provides muon protection and IP separation. It also generates the IP crossing angle which facilitates extraction of the spent beams. Figure 10-1 shows a schematic layout (drawn to scale) of the IP switch, the two big bend sections, and the skew correction and diagnostic sections (Chapter 11). These sections follow the collimation section (Chapter 9) and a 100-m-long emittance diagnostic section. An NLC design with two IPs will require two IP switches and four big bends.

10.2 The IP Switch

The purpose of the IP switch is to provide capability for switching the beam between two alternate final focus beam lines. The IP switch should provide enough separation so that most of the major transport elements are not shared by the two beam lines. Rapid switching of the beam between alternate transport lines is not necessary. It is probably possible to make the switch in a period of less than one hour. Emittance growth from aberrations and synchrotron excitation should be negligible (*i.e.*, a few percent).

10.2.1 Optics Design

The IP switch bends the beam a total of 1.5 mr. Figure 10-2 shows the IP switch optics. The QS quadrupole is horizontally movable in order to switch between IPs. It is displaced by ± 3.25 cm (± 2.6 cm) for the 500-GeV/beam (750-GeV/beam) configuration.

The upgrade to 750 GeV/beam is accomplished by adding ten 3-m-long dipoles inboard of the existing dipoles (Figure 10-3a and 10-3b). Some quadrupole strength changes are also necessary. Figure 10-3 shows the beam line offsets for the 500-GeV/beam (Figure 10-3a) and the 750-GeV/beam (Figure 10-3b) cases. The plots start from the end of the collimation section and continue through the big bend matching section. The BS dipoles should be thin C-magnet types to fit in the 11.7-cm (7.8-cm) center-to-center separation at the face of the first dipole just downstream of the QS quadrupole. The first five BS dipoles (first ten for the 750 GeV/beam configuration) provide the separation, while the QS quadrupole and the next five (ten) BS dipoles make the system achromatic. Beam line elements through QS are common to both beam lines. Three configuration modes are possible. The first is at 500 GeV/beam with 10 dipole magnets installed (Figure 10-3a), the second is at 500 GeV/beam with 20 dipole magnets installed (Figure 10-3b), and the third is 750 GeV/beam with 20 dipole magnets installed (also Figure 10-3b). Tables 10-1 and 10-2 list the magnets for both energies. The match into the IP switch and also into the big bend has been accomplished with two sets of four quadrupoles (QM1,...8) of modest design. The center-to-center separation at the face of QM5 is 15.7 cm, requiring a special thin quadrupole design at least for QM5 and QM6.

10.2.2 Chromatic Emittance Dilution

Tracking studies using the tracking code TURTLE [Carey 1982] have been made for beams with Gaussian energy distributions having rms of 0.2, 0.4, and 0.6%. The results for the entire beam line (from end of

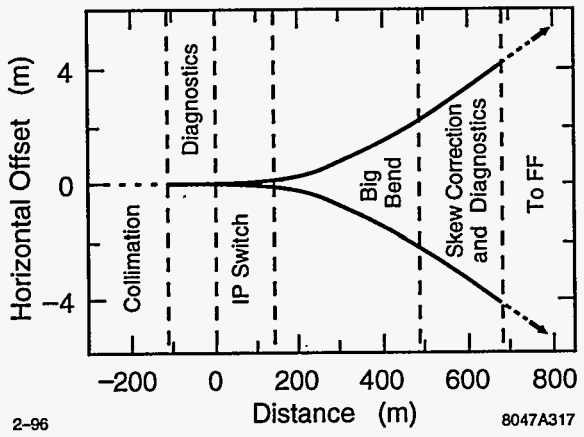


Figure 10-1. *IP switch/big bend layout (to scale).*

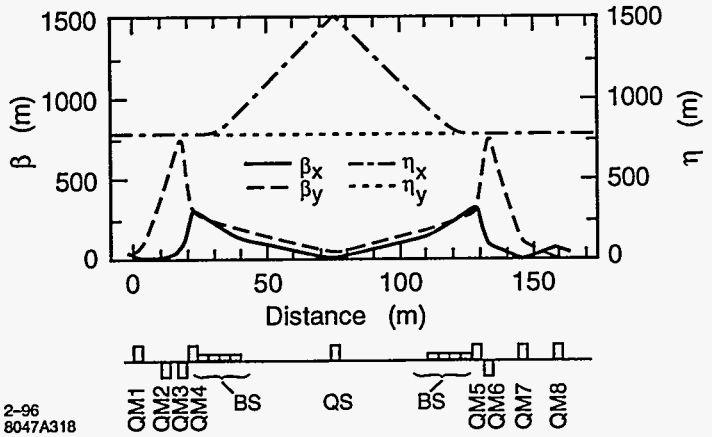


Figure 10-2. *IP Switch Optics (500 GeV/beam).*

Name	Number	Length (m)	Radius (mm)	Pole-Tip Field (kg)
QM1	1	2.5	6	+5.85 (+8.78)
QM2	1	2.5	6	−0.33 (−0.38)
QM3	1	2.5	6	−6.57 (−10.0)
QM4	1	2.5	6	+6.62 (+9.97)
QS	1	2.0	6	+2.35 (+4.43)
QM5	1	2.5	6	+6.08 (+9.15)
QM6	1	2.5	6	−6.60 (−9.97)
QM7	1	2.5	6	+3.01 (+4.62)
QM8	1	2.5	6	+6.66 (+10.0)

Table 10-1. *IP-switch quadrupole magnets for 500 GeV/beam (750 GeV/beam). Fields for the 20-dipole, 500-GeV/beam mode (not shown) are simply scaled from the 750-GeVcase.*

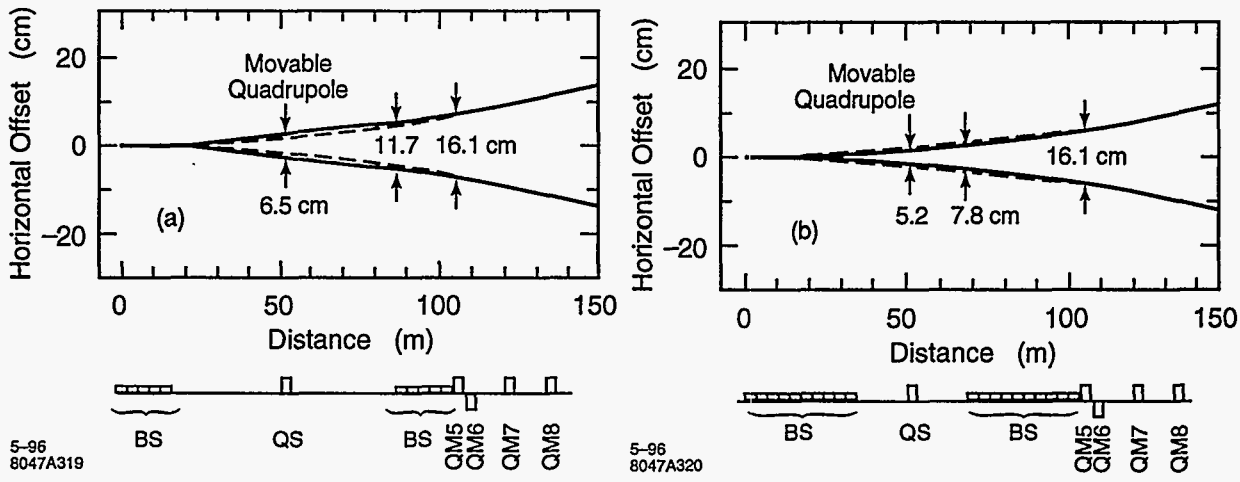


Figure 10-3. *(a) IP switch beam line (500 GeV/beam). Dashed line is nominal beam line at 750 GeV. (b) IP switch beam line at 750 GeV/beam after adding 10 more 3-m dipoles inboard of existing dipoles. Dashed line is nominal beam line at 500 GeV/beam.*

Beam Energy (GeV)	Name	Number	Length (m)	Half Gap (mm)	Field (kGauss)
500	BS	10	3.0	6	0.834
500	BS	20	3.0	6	0.417
750	BS	20	3.0	6	0.625

Table 10-2. *IP switch dipole magnets for different energy modes.*

collimation section to beginning of final focus) are tabulated in Table 10-9. The chromatic contribution to emittance increase for the IP switch alone at 0.3% rms energy spread is <1% in each plane.

10.2.3 Synchrotron Radiation

The horizontal emittance dilution, energy spread increase, and energy loss due to synchrotron radiation (SR) through the IP switch are summarized in Table 10-3. The 7.3% emittance increase quoted in Table 10-3 is with respect to the main damping ring extracted emittance. For a more realistic end-of-linac emittance of $\gamma\epsilon_{x0} = 5 \times 10^{-6}$ m the increase is 4.4%.

10.2.4 Tolerances

The single-element tolerances for the IP switch magnets are listed in Tables 10-4 (dipoles) and 10-5 (quadrupoles). Each tolerance represents a 2% luminosity loss for that single element's effect on one beam. The effects of these errors generally increase the IP beam size except in the case of dipole field regulation and quadrupole transverse vibration which continuously steer the beams out of collision. In this case, since the

Beam Energy (GeV)	Number of Dipoles	$\Delta\epsilon_{x\mathrm{SR}}/\epsilon_{x0}$ (%)	$\sigma_{\delta\mathrm{SR}}$ (10^{-4})	Energy loss (MeV)
500	10	3.3	7	66
500	20	0.7	3	33
750	20	7.3	10	167

Table 10-3. *Horizontal emittance dilution, energy spread increase and energy loss due to synchrotron radiation through the IP switch ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m).*

NAME	Quantity	roll (mr)	$\Delta B/B_0$ (%)	b_1/b_0 (%)	b_2/b_0 (%)
BS	10	10	0.033	3.4	300

Table 10-4. *IP switch dipole magnet single element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 4 mm. The sextupole field harmonics for the dipoles are extremely loose ($\sim 300\%$ at $r = 4$ mm).*

NAME	Quantity	roll (mr)	Δx offset (μm)	Δy offset (μm)	Δx_{rms} vibrate (μm)	Δy_{rms} vibrate (μm)	$\Delta B/B_0$ (%)	b_2/b_1 (%)
QM1	1	4.1	360	21	1.1	0.063	4.5	1650
QM2	1	27	6800	130	20	0.400	11	4000
QM3	1	0.47	150	5.2	0.45	0.016	0.33	55
QM4	1	0.40	83	7.9	0.25	0.024	0.56	52
QS	1	1.2	1680	76	5.0	0.230	5.1	150
QM5	1	0.43	90	8.7	0.27	0.026	0.61	58
QM6	1	0.46	150	5.2	0.44	0.016	0.33	55
QM7	1	12	1500	30	4.5	0.090	4.9	3750
QM8	1	7.2	190	61	0.58	0.180	4.5	1000

Table 10-5. *IP switch quadrupole magnet single element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3\times10^{-6}$ m, $\gamma\epsilon_{y0} = 3\times10^{-8}$ m, $\sigma_\delta = 0.3\%$). Sextupole field harmonics (b_2/b_1) are evaluated at a radius of 4 mm and are very loose ($> 50\%$).*

exact betatron phase to the IP is not calculated, phase averaging is applied. The tolerances given in the tables have not yet been distributed out into a weighted tolerance budget; the numbers are for reference. In fact, given multiple errors over multiple elements, these tolerances are much too loose. However, since tuning considerations have not been folded in, most static, non-steering errors may also be corrected over some reasonable range.

10.2.5 Diagnostics and Correctors

Beam position monitors (BPM) will be required, probably one horizontal and one vertical BPM per $\pi/2$ of betatron phase. A minimal number of dipole orbit correctors should be used; the optimal locations for these devices have yet to be established. The diagnostic section at the end of the big bend will be used to determine the quality of the match into the big bend.

The horizontal dispersion of 32 mm (26 mm at 750 GeV/beam) at QS provides an excellent location for the measurement of beam energy and energy spread, using a BPM and a profile measurement device such as a wire scanner. At the entrance to QS, the dispersive horizontal spot size for a beam with 0.3% rms energy spread is 96 μm (78 μm), while the betatron spot size ($\gamma\epsilon_x = 3 \times 10^{-6}$ m) at this location is only 5.2 μm (4.2 μm), allowing a good energy spread measurement. In addition, placement of a 1-μm-resolution BPM at this location will provide a relative energy measurement resolution of $\sim 3 \times 10^{-5}$.

10.2.6 Beam Correction Issues

Beam-based techniques will be used to verify transverse alignment during commissioning of the beam line, but will probably be infrequently necessary thereafter. Orbit correction algorithms remain to be studied, but a simple point-to-point scheme will probably be sufficient. Feedback stabilization of beam position at the entrance to the big bend will probably be desirable, depending on the stability of the incoming beam.

10.2.7 Other Issues

Some issues remain to be considered, including:

- The need for machine protection collimators and their locations.
- Vacuum and pumping requirements.

10.3 The Big Bend

The big bend is a low-angle arc after the main linac which provides detector muon protection [Keller 1993], an IP crossing angle to facilitate extraction of the spent beams, and allows switching between multiple IPs. The total bend angle (including 1.5-mr IP switch angle) is 10 mr (20-mr IP crossing angle) which provides ~40-m spatial separation between the two IPs (~700-m transport to an ~1600-m-long final focus). At 500–750 GeV/beam, the horizontal emittance growth due to SR sets lower limits on the system design length. The following Section describes an optimized optical design of this big bend section for 500 GeV/beam and 750-GeV/beam electrons (or positrons).

10.3.1 Optical Design

For electrons, the emittance growth due to SR is calculated using [Helm 1973, Raubenheimer 1993]:

$$\Delta\gamma\epsilon_x \approx \left(4\times10^{-8}\,\mathrm{m^2\,GeV^{-6}}\right)\cdot E^6\sum_i \frac{L_i\langle H\rangle_i}{|\rho_i|^3} \quad , \tag{10.1}$$

where the summation is over bending magnets, L is the magnet length, ρ is the bend radius of the magnet, E is the beam energy, and $< H >$ is the mean of the usual "curly-H" function.

$$\langle H\rangle = \frac{1}{L}\int_0^L \frac{\eta^2+(\eta'\beta+\eta\alpha)^2}{\beta}dz \tag{10.2}$$

This integral has been solved [Helm 1973] for a magnet with bending and focusing. The mathematical result is lengthy and is not reproduced here. To find the optimal parameters for a string of FODO cells, this result is used with Equation (10.1) in a convenient computer program to calculate the SR emittance growth with maximum quadrupole pole-tip fields of 10 kg at 750 GeV/beam with a 6-mm radius. Preliminary resistive wall calculations [Bane 1995] indicate that this radius might be decreased to 3 mm, shortening the system length by ~100 m. However, it is thought that this possibility may not provide an adequate safety factor.

The number of FODO cells and the phase advance per cell are varied to find the minimum total length for a ~2% horizontal SR emittance growth ($\gamma\epsilon_{x0} = 3\times10^{-6}$ m). Both separated function and combined function magnet systems were explored. The parameters reached represent a compromise between theoretical optimal values and realistic constraints on magnet lengths and reasonable phase advance per cell.

The phase advances per cell chosen ($\Psi_x = 108°, \Psi_y = 90°$) do not represent the precise optimum ($\Psi_x = 135°, \Psi_y < 72°$). The values are biased towards a more reasonable design considering beam position monitor sampling, chromaticity, a potential sextupole resonance and magnet alignment tolerances. The effect on the total length of this slight bias is small (<10%). Splitting the horizontal and vertical tunes in the separated function lattice improves the bending magnet density because it allows the D-quadrupoles to be shorter than the F-quadrupoles for a constant pole-tip field. However, in a combined function design, this split does not improve the SR emittance dilution. The tune split may also desensitize the beam to a potential ion instability. The choice between combined function (CF) or separated function (SF) lattice reduces to a few points listed below.

Combined Function: The CF lattice is more space-efficient and only one type of magnet needs to be built. The main disadvantage is that the focusing strength is not tunable without changing the bending strength. The magnets are long (~6 m) since they bend and focus, but the net length is still shorter than the SF design (~30%).

Separated Function: In its favor, the SF lattice may be more easily tuned since the focusing strength will be independent of the bend strength. For example, the phase advance per cell may be changed to provide a trombone tuner between the collimation phase and the IP. Also beam-based alignment techniques can be applied. However, there are three types of magnets to build in this scheme and the overall length needs to be longer than the CF design (~30%).

Due to its tunability and beam-based alignment potential, we have chosen the SF design. The dispersion and beta functions of the 500-GeV/beam SF design are shown in Figure 10-4. The final design uses 15 FODO cells with four dipoles per cell. Table 10-6 lists the FODO parameters for the big bend SF design.

A 76-cm quadrupole-to-dipole space has been maintained for BPM and ion pump placement and a 50-cm dipole-to-dipole space is held so that the dipole and quadrupole magnet lengths are no more than 3 m and

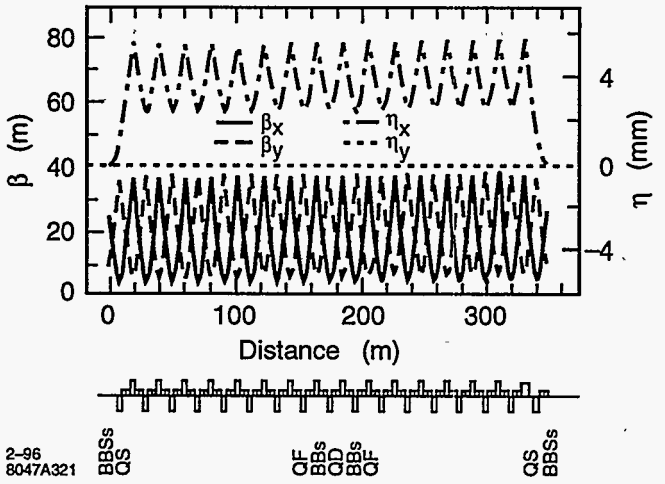


Figure 10-4. *Big bend optics for separated function lattice.*

Total Length (FODO cells)	(m)	312
Total bend angle	(mr)	8.5
Number of FODO cells [a]		15
Maximum $\beta_{x,y}$	(m)	36
Maximum η_x	(mm)	5.6
x-phase advance/cell	(°)	108
y-phase advance/cell	(°)	90
Spin phase advance/cell	(°)	37 (55)
Dipole magnet length	(m)	3.0
Bend radius (per dipole)	(km)	22.6
F-quad length	(m)	2.50
D-quad length	(m)	2.26

[a] There are four dipole magnets per cell.

Table 10-6. *Optimized big bend parameters at 500 GeV/beam (750 GeV/beam) for 6.7-kGauss (10-kGauss) quadrupole pole-tip fields at 6 mm radius and tolerable SR emittance growth.*

2.5 m, respectively. A "missing-magnet" dispersion matcher/suppressor is used at the entrance and exit of the FODO string. The matcher/suppressor magnets are identical in size and strength to those in the FODO section. Only their longitudinal position has been adjusted to obtain the periodic dispersion function in the FODO section. In this way, all 64 dipoles may be powered in series with one power supply, the 17 defocusing quadrupoles powered with a second supply, and the 16 focusing quadrupoles on a third. This allows tuning of the phase advance per cell independently in each plane. Table 10-7 lists the dipole magnet parameters and Table 10-8 lists the quadrupole magnet parameters for the big bend as shown in Figure 10-4. The 3-m-long dipole magnet design may also be used in the IP switch as long as it is a C-magnet design with dimensions which meet the requirements described in Section 10.2.1. The 2.5-m-long QF design may be used for the QM1-8 matching quadrupoles in the IP switch. The center-to-center beam line separation between the two big bends at the face of the first QD magnet is 28.8 cm (for either 10 or 20 IP switch dipoles) which sets an upper limit on the outer horizontal dimension of the big bend quadrupole magnets of <28 cm full width.

Name	Number	Length (m)	Half Gap (mm)	Field (kGauss)
BB	64	3.0	6	0.738 (1.108)

Table 10-7. *Big bend dipoles for 500 GeV/beam (750 GeV/beam).*

Name	Number	Length (m)	Radius (mm)	Pole Tip Field (kGauss)
QD	17	2.258	6	-6.70 (-10.05)
QF	16	2.500	6	+6.70 (+10.05)

Table 10-8. *Big bend quadrupoles for 500 GeV/beam (750 GeV/beam).*

Rms energy spread (%)	$\Delta\epsilon_x/\epsilon_{x0}$ (%)	$\Delta\epsilon_y/\epsilon_{y0}$ (%)
0.2	1.2	0.2
0.4	3.7	1.6
0.6	7.7	4.2

Table 10-9. *Chromatic emittance increase at 500 GeV/beam for various rms Gaussian energy spreads for beam line including initial diagnostic section, IP switch, all matching sections, big bend, and the skew correction and diagnostic section (synchrotron radiation effects not included).*

10.3.2 Chromatic Emittance Dilution

Tracking studies using TURTLE [Carey 1982] have been made for the entire beam line described in this chapter (nearly 800 meters of beam line from the end of the collimation section to the beginning of the final focus). A Gaussian energy distribution with rms of 0.2%, 0.4%, and 0.6% energy spread was used. For all cases, the emittances used were $\gamma\epsilon_{x0} = 3\times10^{-6}$ m and $\gamma\epsilon_{y0} = 3\times10^{-8}$ m. The chromatic emittance dilution at 500 GeV/beam for each case is listed in Table 10-9 (SR effects not included). Sextupole compensation is not necessary.

10.3.3 Synchrotron Radiation Effects

Optics

Table 10-10 lists the SR parameters for the 500-GeV/beam big-bend design of Figure 10-4. The 750 GeV/beam parameters are also given. The fractional emittance growth referred to is the main damping ring extracted emittance of $\gamma\epsilon_x = 3\times10^{-6}$ m. The energy loss across the length of the big bend is 0.07% at 500 GeV (0.22% at 750 GeV). Given the large chromatic bandpass of the big bend (less than 2% emittance increase at 1% rms Gaussian energy spread), it is not necessary to taper the fields through the system.

Beam Energy	(GeV)	500	750
Critical Energy (u_c)	(MeV)	12	42
SR-generated rms Energy Spread	(%)	0.015	0.040
Energy Loss	(GeV)	0.331	1.675
Horizontal SR Emittance Growth	(%)	0.3	3.3
Number of photons/electron		90	456

Table 10-10. *Synchrotron radiation parameters for big bend at 500 GeV and 750 GeV/beam for* $\gamma\epsilon_x = 3 \times 10^{-6}$ *m.*

Detector Backgrounds

The energy distribution generated by SR has a long tail which falls off rapidly for energies, u, well above the critical energy, u_c ($u/u_c \equiv \xi >> 1$) [Sands 1970].

$$n(\xi) \propto \xi^{-1/2} e^{-\xi} \tag{10.3}$$

To estimate the number of electrons in the tail which achieve an oscillation amplitude comparable to one rms horizontal beam size (in the interest of staying clear of the final doublet face), the necessary energy deviation, u_1, is written in terms of the rms SR energy spread, σ_δ, the relative horizontal SR emittance increase, $\Delta\epsilon_x/\epsilon_{x0}$, and the beam energy, E_0.

$$\xi_1 \equiv \frac{u_1}{u_c} = \frac{\sigma_\delta}{u_c/E_0} \cdot \frac{1}{\sqrt{\Delta\epsilon_x/\epsilon_{x0}}} \sim 40 \text{ (at } E_0 = 750 \text{ GeV)} \tag{10.4}$$

For the worst case (750 GeV/beam), a particle which is 40 critical energies lower than nominal will oscillate at one sigma. The number of electrons per bunch at or beyond this energy, N_{e^-}, is calculated in Equation (10.5) where N_γ is the total number of photons/e^-. Even for 10^{10} electrons per bunch at 750 GeV/beam, this energy tail is insignificant and will not generate a background.

$$N_{e^-} = 10^{10} \cdot \frac{9\sqrt{2}}{15} N_\gamma \int_{\sqrt{2\xi_1}}^{\infty} e^{-x^2/2} dx \sim 10^{-4} \tag{10.5}$$

10.3.4 Tuning, Tolerances, and Corrections

The single-element tolerances for the big bend magnets are listed in Tables 10-11 (dipoles) and 10-12 (quadrupoles). Each tolerance represents a 2% luminosity loss for that single element's effect on one beam. The effects of these errors generally increase the IP beam size except in the case of dipole field regulation and quadrupole transverse vibration which continuously steer the beams out of collision. In this case, since the exact betatron phase to the IP is not calculated, phase averaging is applied. The tolerances given in the tables have not yet been distributed out into a weighted tolerance budget; the numbers are for reference. In fact, given multiple errors over multiple elements, these tolerances are much too loose. However, since tuning considerations have not been folded in, most static, non-steering errors may also be corrected over some reasonable range.

At this time, detailed tolerance and tuning studies have not been performed. However, the big bend design must include dispersion tuning elements for both planes and betatron phases to correct any residual dispersion

NAME	Quantity	roll (mr)	$\Delta B/B_0$ (%)	b_1/b_0 (%)	b_2/b_0 (%)
BB	64	45	0.014	35.0	7500

Table 10-11. *Big bend dipole magnet single-element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Quadrupole and sextupole field harmonics (b_1/b_0 and b_2/b_0) are evaluated at a radius of 4 mm. The sextupole component tolerances for the dipoles at $r = 4$ mm are extremely loose.*

NAME	Quantity	roll Δx (mr)	Δy offset (μm)	$\Delta x_{\rm rms}$ offset (μm)	$\Delta y_{\rm rms}$ vibrate (μm)	$\Delta B/B_0$ vibrate (μm)	b_2/b_1 (%)	(%)
QD	17	4.4	680	26	2.1	0.077	6.9	2400
QF	16	4.3	230	57	0.70	0.170	3.5	500

Table 10-12. *Big bend quadrupole magnet single-element tolerances at 500 GeV/beam for 2% luminosity loss each ($\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m, $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m, $\sigma_\delta = 0.3\%$). Sextupole field component tolerances (b_2/b_1) are evaluated at a radius of 4 mm and are very loose ($\geq 500\%$).*

due to magnet misalignments and gradient errors. Coupling and matching corrections as well as diagnostics exist just after the big bend (Chapter 11). Vertical dispersion correction can be provided by adding four small skew quadrupoles (of zero nominal field)—one per cell in the last four cells. This scheme takes advantage of the 90° vertical phase advance per cell by pairing skew quadrupoles at $-I$ transfer matrix (2 cell) separation so that, for equal and opposite skew quadrupole settings, no betatron cross-plane coupling is generated. The second pair of skew quadrupoles then handles the other betatron phase. The range of correction for one pair of 50-cm-long, ±5-kGauss pole-tip field, 6-mm pole-tip radius skew quadrupoles located 10 cm downstream of each QD at $\beta_y = 32$ m, $\eta_x = 2.8$ mm is $\epsilon_y/\epsilon_{y0} = 13.4$ at 0.3% rms energy spread for $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m at 500 GeV/beam (the vertical dispersion induced at the center of a QD magnet is as much as 4.6 mm). No significant coupling or horizontal beta function perturbation is generated over this range. However, for very large corrections some second order dispersion may be induced which will limit the correction range or require similar skew sextupole tuners. This level has not yet been studied.

The horizontal dispersion may be controlled similarly by adding two pairs of small normal quadrupoles (of zero nominal field). Due to the 108° horizontal phase advance per cell these quad pairs must be spaced by 5 cells to provide a $-I$ separation. If the 9th and 10th as well as the 14th and 15th cell include a 50-cm-long, ±8-kg, 6-mm-radius quadrupole which is 10-cm upstream of each QF (at $\beta_x = 31$ m, $\eta_x = 5.2$ mm) the emittance correction range per pair will be $\epsilon_x/\epsilon_{x0} = 5.1$ (the additional horizontal dispersion induced at the center of a QF magnet is as much as 14 mm). However, since the vertical transfer matrix between paired normal quadrupoles (5 cells) is not equal to $-I$, there will be a small perturbation to the vertical beta function which amounts to a 10% beta beat amplitude at full horizontal dispersion correction (±8 kGauss quadrupole fields). This small effect is correctable with the matching quadrupoles just upstream of the pre-final focus emittance diagnostic section (Chapter 11). The dispersion correction specifications are summarized in Table 10-13.

Given the dispersion correction available, the big bend quadrupoles will probably not require movers. However, beam-based alignment techniques will greatly benefit in speed and convergence if movers (at least in the vertical plane) are available. The movers should control the vertical position to ~5-μm resolution over a range of approximately ±500 μm. Roll control is not required given the fairly loose roll tolerances as well as

Quad Type	Quantity	Length (m)	pole radius (mm)	max. field (kGauss)	rms reg. tolerance (%)	$\Delta\eta_{x,y}^{max}$ at $\beta_{x,y}^{max}$ (mm)	max. ϵ_y/ϵ_{y0}
skew	4	0.5	6	±5	0.1	4.6	13
normal	4	0.5	6	±8	0.5	14	5.1

Table 10-13. *Big bend dispersion tuning magnet specifications at 500 GeV/beam for 0.3% rms energy spread, $\gamma\epsilon_{x0} = 3 \times 10^{-6}$ m and $\gamma\epsilon_{y0} = 3 \times 10^{-8}$ m. There is one skew quadrupole in each of the last four FODO cells 10 cm downstream of the QD and one normal quadrupole in each of cells 9, 10, 14 and 15 at 10 cm upstream of the QF.*

the skew (Chapter 11) and vertical dispersion corrections available. Beam-based alignment of the big bend quadrupoles has not yet been studied in detail. However, an independent partial current shunting switch across each big bend quadrupole will probably be a significant advantage for any beam-based alignment algorithm.

Horizontal and vertical dipole correctors at each QF and QD, respectively, will be required to initially steer the beam line and to use in fast feedback applications. Correctors with ±1.0-kg fields and 25-cm length will be adequate to displace the beam nearly ±500 μm at the next similar quadrupole at 500 GeV/beam. The horizontal correctors will then need to regulate at $\sim 1 \times 10^{-3}$ over the 100-ms (10 pulse) range while similar vertical correctors will need $\sim 1 \times 10^{-4}$ regulation (~0.3% luminosity loss due to all correctors in both big bends for both planes).

The tolerance on the beta match into the big bend is quite loose. It can be shown that the SR emittance increase approximately scales with the amplitude of the incoming beta mismatch.

$$\Delta\varepsilon_{SR} \approx B_{\rm mag}\Delta\epsilon_{\rm SR-nom.} \tag{10.6}$$

Here $B_{\rm mag}(\geq 1)$ is the beta mismatch amplitude in the horizontal plane and $\Delta\varepsilon_{\rm SR-nom.}(\ll \varepsilon_{x0})$ is the nominal SR emittance increase for a matched incoming beam. A very large mismatch of $B_{\rm mag} = 2$ ($\beta_x \approx 4\beta_{x0}, \alpha_x = \alpha_{x0} = 0$) will amplify a nominal 0.3% SR emittance increase to 0.6%. The vertical match has no such constraint.

10.3.5 Spin Transport and Depolarization

The spin phase advance per cell (spin tune) has also been tabulated in Table 10-6. A spin tune-betatron tune resonance is to be avoided, or small vertical oscillations will precess the electron spin into the vertical plane [Limberg 1993]. However, even without a resonance there may be significant vertical alignment error induced spin rotation due to the large gradient magnets and the extremely high energy. A 100-μm vertical beam offset at 500 GeV/beam in a single QD magnet will rotate a longitudinally oriented spin vector 1° into the vertical plane. If the errors are static this may be compensated by properly orienting the incoming spin vector using the 2-GeV solenoid rotator system (Chapter 5).

The depolarization for a bend through θ, at an energy γ, and an incoming Gaussian rms energy spread of σ_δ is

$$\overline{P}/P_0 = \frac{1}{\sqrt{2\pi}\sigma_\delta}\int_{-\infty}^{\infty} e^{-\delta^2/2\sigma_\delta{}^2}\cos(a\gamma\delta\theta)d\delta = e^{-(a\gamma\theta\sigma_\delta)^2/2} \ . \tag{10.7}$$

where $a = (g-2)/2$ is the anamolous magnetic moment. For a 10-mr bend at 500 GeV/beam with a 0.3%-rms energy spread the relative depolarization is 0.06% (0.13% at 750 GeV).

10.3.6 Vacuum System

The pressure requirements for the big bend section are set by tolerable detector background levels [Irwin 1993]. At present, it is desirable to achieve an average pressure of $\sim 5 \times 10^{-8}$ Torr in the big bend. If the chamber is cylindrical and made of aluminum with specific outgassing rate $q = 5 \times 10^{-10}$ T-l/s-m (similar to mature SLC arcs), a specific conductance for a 6-mm-radius of $c = 0.18$ m-l/ and ion pumps of speed $S > 5$ l/s placed three per FODO cell ($L \sim 7$ m for a total of 45 ion pumps), the system is conductance-limited with average pressure [Ziemann 1992].

$$\overline{P} \cong qL^2/3c \approx 5 \times 10^{-8}\text{Torr} \ . \tag{10.8}$$

An order-of-magnitude-lower pressure is probably achievable by using a baked stainless steel chamber with a much lower specific outgassing rate.

References

[Keller 1993] L.P. Keller, "Muon Background in a 1.0-TeV Linear Collider", SLAC–PUB–6385 (October 1993).

[Helm 1973] R.H. Helm, M.J. Lee, P.L. Morton, M. Sands, "Evaluation of Synchrotron Radiation Integrals", SLAC–PUB–1193 (March 1973).

[Raubenheimer 1993] T.O. Raubenheimer, P. Emma, S. Kheifets, "Chicane and Wiggler Based Bunch Compressors for Future Linear Colliders", SLAC–PUB–6119 (May 1993).

[Bane 1995] K. Bane, private communication (1995).

[Sands 1970] M. Sands, "The Physics of Electron Storage Rings", SLAC–121 (November 1970).

[Limberg 1993] T. Limberg, P. Emma, R. Rossmanith, "The North Arc of the SLC as a Spin Rotator", *Proc. of the 1993 Part. Acc. Conf*, Washington, DC (1993).

[Irwin 1993] J. Irwin, R. Helm, W.R. Nelson, D. Walz, "Conventional Collimation and Linac Protection", SLAC–PUB–6198 (May 1993).

[Brown 1977] K.L. Brown *et al.*, "TRANSPORT", SLAC–91 (May 1977).

[Carey 1982] D.C. Carey *et al.*, "Decay TURTLE", SLAC–246 (March 1982).

[Ziemann 1992] V. Ziemann, "Vacuum Tracking", SLAC–PUB–5962 (October 1992).

[Emma 1994] P. Emma, T. Limberg, R. Rossmanith, "Depolarization in the SLC Collider Arcs", *Proc. of the 1994 European Part. Accel. Conf.*, London, England (July 1994).

Contributors

- Paul Emma
- Dick Helm
- Tor Raubenheimer
- Mark Woodley

11

Final Focus

Contents

11.1 Introduction 661
11.2 Parameters and Specifications 662
11.2.1 Goals and System Boundaries 662
11.2.2 Parameter List 663
11.2.3 Energy Flexibility 663
11.2.4 Overview 664
11.3 Skew Correction and Diagnostic Section 665
11.3.1 The Skew Correction Section 665
11.3.2 The Diagnostic Sections 666
11.3.3 Tuning Simulations 668
11.4 Beta-Matching Section 669
11.4.1 Optical Design 670
11.4.2 Tuning Elements 672
11.4.3 Tolerances 675
11.5 Chromatic Correction and Final Transformer 676
11.5.1 Introduction, Parameters and Dilutions 676
11.5.2 Layout and Optics 679
11.5.3 Performance 687
11.5.4 Tuning 696
11.5.5 Tolerances 698
11.5.6 Feedback and Stability 731
11.5.7 Operations and Controls 736
11.5.8 Components 737
11.5.9 Summary 743
11.6 The Final Doublet 743
11.6.1 Doublet Parameters 744
11.6.2 An Analytical Model of the Doublet 745
11.6.3 Final Doublet Wake Effects 751
11.6.4 Synchrotron Radiation Effects 753
11.6.5 Nominal Final Doublet Designs 755
11.6.6 Tolerances 758
11.6.7 Steering Jitter Budget 759
11.6.8 Sources of First-order Dispersion in the Doublet 761
11.6.9 Superconducting Quadrupole Q1.5 764
11.6.10 Summary 765
11.7 Crossing Angle, Crab Cavity, and Solenoid 767
11.7.1 Determination of Crossing Angle 767

11.7.2 Crab Cavity 768
11.7.3 Solenoid Effects 771
11.7.4 Summary 774
11.8 The Beam Extraction and Diagnostic System (The Dump Line) 774
11.8.1 Comparison with the SLC 775
11.8.2 Basic Design Procedures, Constraints, and Assumptions 776
11.8.3 The Beam-Beam Calculations 778
11.8.4 Beam Line Optics 784
11.8.5 Beam Control Hardware 789
11.8.6 Beam Monitors and Diagnostics 791
11.8.7 Beam Dump 798
11.8.8 Secondary Beams 798
11.8.9 Energy Recovery and Its Applications 799
11.8.10 Other Questions and Problems 800
11.9 Conclusions and Comments 800
11.A The Beam Dumps 802
11.A.1 The Beam Dump Vessel 804
11.A.2 The Window 806
11.A.3 Isotope Production 808
11.A.4 Radiolysis and Hydrogen Evolution 808
11.A.5 Summary 809

11.1 Introduction

The final-focus system comprises the region between the big bend and the main beam dump. Its function is to demagnify the transverse sizes of electron or positron beams by a factor 80 horizontally and 300 vertically, down to a value of about 250 nm × 4 nm at the interaction point (IP), where the beams collide. After the collision, an extraction line guides the beam remnants onto the beam dump.

The final-focus designs for 500 GeV and for 1 TeV which are described in this chapter, have a comfortable momentum bandwidth and can operate in the full ZDR parameter plane. A dedicated geometry-adjustment section at the entrance to the final focus facilitates the adiabatic upgrade from 350-GeV to 1.5-TeV-c.m. energy, with a constant IP position and only minor transverse displacements.

Important design issues of the final-focus system are linear and nonlinear optics; momentum bandwidth; effects of synchrotron radiation; tuning schemes; sensitivity to varying beam conditions, such as incoming orbit, emittance, and energy; tolerances on alignment, vibrations and field changes; response to ground motion; wakefield effects; stability of the final doublet; design and tolerances of the crab cavity; compensation of the solenoidal detector field; beam removal from the IP; maintenance, tuning and stabilization systems.

When compared to previous colliders, the NLC beam parameters and tolerances enter a distinctly new regime. The spot size at the IP is 125 times smaller than that at the SLC, which is already as small as 500 nm, and is 15 times smaller than the 70-nm spot size achieved at the FFTB. In addition to delineating the stringent tolerances which are a consequence of the small spot size, we will attempt to clarify how we propose to achieve these tolerances. The story is complex since there are several significant timescales, several distinct aberrations to consider at each timescale, many elements and element parameters which can influence each aberration, and many sources of change for the parameters of each element. Each timescale is associated with a tuning system and a maintenance system which stabilizes the final focus between tunings. We use the word tuning when the diagnostic involves determination of IP beam and collision conditions. Hence missteering at the IP is one of the aberrations which is tuned.

In the case of alignment, because everything is moving, it is often difficult to clearly convey what elements need to stay aligned, by how much, and with respect to what other elements or coordinate system. An important result of ground-motion studies, which are described in Appendix C, is that the ground (bedrock) can be used as a reference, in the sense that if all elements were moving as the bedrock below them, then the beams would remain in collision. Section 11.5.5 discusses these results as they apply to the final-focus system.

Aberrations are described in Section 11.5.3. A table in that Section (Table 11-9) lists all aberrations that will be tuned, how they will be tuned, frequency of tuning, what budget of luminosity is allotted to the final-focus system (since other upstream elements can contribute to these aberrations), and what systems maintain (stabilize) the system between tunings. Tuning and maintenance systems are described in more detail in Sections 11.3.1, 11.3.2, and 11.5.4. 11.5.4 and 11.5.8.

In Section 11.5.5, relevant timescales, each of which corresponds to a tuning frequency, are listed and named, and all tolerances described are specified relative to these timescales. Tolerances can be achieved if:

- The diagnostic systems used for tuning have the required resolution.
- Tuning and adjustment knobs have the required sensitivity.
- Maintenance systems stabilize the final focus between tunings.

The performance of the maintenance systems in turn depends on achieving the required resolution in their diagnostic systems. Diagnostic systems are described in Section 11.5.8.

In the SLC, only the IP steering and incoming beam launch conditions have been automated. Though a dither feedback has been considered for the other IP aberrations, it would be speculative at this point to believe such a system would work at the NLC, and we do not make this assumption. Automated tuning (steering) and automated maintenance systems are described in Section 11.5.6.

Tables 11-13 through 11-18 list the sensitivities of the parameters of each element. Using these sensitivities a tolerance is assigned to each element (Table 11-19) so that the system budget of each aberration is met. Elements with similar tolerances have been grouped together. Table 11-21 gives the luminosity loss that results from these tolerance assignments, to be compared with the final-focus system budget of Table 11-9. Corresponding to each tolerance in Table 11-19, there is a brief statement on how this tolerance can be met (Table 11-20). Finally, in Section 11.5.8, we present a table that lists every diagnostic element and the resolution required.

It is worth noting that both the optical design and proposed tuning algorithms draw heavily from SLC and FFTB experience. In the limited time available, every small aspect of the design could not be studied in minute detail. However, we feel that all potentially critical or problematic topics have been addressed in depth. Outstanding work includes, for instance, a more detailed study of certain tuning and alignment procedures, and a more exhaustive description of commissioning, operation and machine protection. These items are not expected to be difficult.

One of the reviewers of the August 95 ZDR workshop (K. Oide) has created an alternative final-focus design for NLC parameters at 500-GeV-c.m. energy, which is based on the odd-dispersion scheme [Oide 1992]. The bandwidth of this system, without any additional 'Brinkmann'-sextupoles, is comparable to that of the present NLC design. An advantage of the odd-dispersion final focus is that it only uses about half the number of quadrupoles. A potential disadvantage is the nonexistence of an IP pre-image point, which may or may not be an operational aid. The tunability, upgradability, performance at higher energy, and background situation of the odd-dispersion design need to be evaluated, prior to a final decision on this alternative. In the present report, only the modular final-focus design by R. Helm is discussed, the performance of which appears to be entirely satisfactory.

11.2 Parameters and Specifications

11.2.1 Goals and System Boundaries

The purpose of the NLC final-focus system is to transport electron and positron beams of energy 180 GeV to 750 GeV from the end of the big bend to the IP, where the demagnified beams are collided, and to remove the beam remnants cleanly to facilitate crucial post-IP diagnostics. To accomplish its task, the final-focus system has to be stable and reliable, and it needs redundant diagnostics to detect and compensate all changes of beam parameters, magnet positions, or field strengths, which would otherwise reduce the luminosity.

For a c.m. energy of 500 GeV, the design spot size at the interaction point (IP) is about 4.2–6.5 nm vertically and 250–300 nm horizontally. The normalized emittances at the entrance of the final focus are assumed to be $\gamma\epsilon_x \approx 4 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 7$–$10.5 \times 10^{-8}$ m. At a c.m. energy of 1 TeV, the design spot size is

3.4–5.2 nm vertically and 200–250 horizontally, for normalized emittances of $\gamma\epsilon_x \approx 4 \times 10^{-6}$ m, and $\gamma\epsilon_y \approx 9$–$13.5 \times 10^{-8}$ m. The beam energy distribution in the final focus depends on bunch compressor and linac configurations. For most studies presented in this chapter, a Gaussian momentum distribution with a relative rms momentum spread of $\delta \approx 0.3\%$ is assumed. This distribution makes it easy to study the effect of a typical energy spread, but the real energy distribution will be non-Gaussian (compare Chapters 5 and 7).

The entrance of the final-focus system is formed by a skew correction section (SCS) and a diagnostics section (DS). These are followed by a geometry-adjustment section (GAS), beta- and phase-matching section (BMS), horizontal and vertical chromatic correction sections (CCX and CCY), which are separated by a beta-exchange module (BX), the final transformer (FT), the interaction region (IR), and the beam removal system (BRS). This chapter describes all these sections except for the IR which is discussed separately in Chapter 12.

The final-focus system is flexible enough to be operated in the entire c.m. energy range from 350 GeV to 1.5 TeV. Most of the data and figures in the following sections refer to the design for either 500-GeV or 1-TeV-c.m. energy.

11.2.2 Parameter List

Table 11-1 exemplifies beam parameters at the interaction point for c.m. energies of 500 GeV and 1 TeV. In this table, the values listed for luminosity and spot size do not include any dilutions, which may arise from high-order aberrations, synchrotron radiation residual uncorrected low-order aberrations, timing offsets or crab crossing errors. When dilutions are taken into account, the expected luminosity is reduced by about 20–30%, see Table 11-4 and the discussions in Section 11.5.

Most beam parameters are variable, as the NLC should function at any point inside a multi-dimensional operating plane [Burke 1995]. From the final-focus point of view, the largest IP divergences represent the worst case, since in this case the effect of nonlinear aberrations, the aperture requirements, and the Oide effect are most severe. The beam parameters listed in Table 11-1 refer to this case.

The minimum horizontal beta function at the IP is limited by nonlinear aberrations, by the maximum number of beamstrahlung photons that can be tolerated Chapter 12. and by the Oide effect (Sections 11.5.3 and 11.6.4). The value presently chosen was imposed by the beamstrahlung, whereas Oide effect and nonlinearities would allow for a 20% smaller beta function. It is, therefore, possible to compensate the increase of the horizontal spot size due to residual low-order aberrations by reducing β_x^*.

11.2.3 Energy Flexibility

The final focus should operate at least in the energy range from 350 GeV to 1.5 TeV. In the present design, this flexibility is accomplished with slightly different geometries for three overlapping energy ranges (Figures 11-9 and 11-10). The implication is that two minor reinstallations of magnets and supports are necessary during the energy upgrade, at around 500 GeV and 1.1 TeV. The IP position is held constant by means of a special bending section—the geometry-adjustment section—which is located at the entrance to the final focus.

At 500-GeV-c.m. energy, the final focus is operated with quadrupoles scaled down from the 1-TeV design. A complication arises in the final doublet, since the last quadrupole is a permanent magnet. Here, the

Comments	c.m. energy	500 GeV	1 TeV
Luminosity w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.52	1.16
Luminosity w. pinch w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.75	1.63
# particles per bunch	N_b (10^{10})	0.65	0.95
# bunches	n_b	90	90
# bunch trains per s	f	180	120
Enhancement factor w/o dil.	H_D	1.46	1.41
Hor. spot size w/o dil.	σ_x (nm)	253	200
Vert. spot size w/o dil.	σ_y (nm)	4.18	3.35
Hor. IP beta function	β_x^* (mm)	8	10
Vert. IP beta function	β_y^* (mm)	0.125	0.125
Norm. hor. emittance	$\gamma\epsilon_x$ (10^{-8} mr)	400	400
Norm. vert. emittance	$\gamma\epsilon_y$ (10^{-8} mr)	7	9
Hor. IP divergence	$\sigma_{x'}$ (μrad)	31.6	20.0
Vert. IP divergence	$\sigma_{y'}$ (μrad)	33.5	26.8
Bunch length	σ_z (μm)	100	125
Crossing angle	θ_c	20 mr	20 mr
Energy spread	δ_{rms}	$\geq 3 \times 10^{-3}$	$\geq 3 \times 10^{-3}$
Free length from IP	l^* (m)	2	2

Table 11-1. *Basic worst-case interaction-point beam parameters without dilutions (see also Table 11-4).*

energy scaling is accomplished by an adjacent superconducting quadrupole whose field changes sign during the energy raise from 500 GeV to 1 TeV. In order to keep the cost of a further upgrade low, the remaining final-focus magnets could be designed such that their strength can be increased to the 1.5-TeV values.

11.2.4 Overview

On the next pages, the skew correction section (SCS) is discussed, along with two options for the adjacent diagnostics section (DS). Section 11.4 is devoted to the beta- and phase-matching section (BMS), which will be used to adapt the IP beta functions and the waist position to varying incoming beam conditions, and to adjust the betatron phase advance between the collimator section and the IP. The geometry-adjustment section (GAS) located upstream of the BMS, the horizontal chromatic correction section (CCX), the beta-exchange module (BX), the vertical chromatic correction section (CCY) and the final transformer (FT), as well as a tolerance analysis for this region are described in Section 11.5, which also contains lists of magnets and diagnostics for the entire region between the SCS and the IP. Section 11.6 discusses the final doublet; Section 11.7 discusses effects of the solenoidal detector field and the crab cavity. The beam removal system (BRS) and beam dump are described in Section 11.8. Section 11.9 concludes the chapter with a short perspective on the present design and on outstanding questions.

11.3 Skew Correction and Diagnostic Section

This beam line section follows the big bend and is used to measure and correct any anomalous cross-plane coupling and to precisely match the beam into the main body of the final focus. In order to minimize the projected emittance, especially for a flat beam, it is necessary to remove all cross-plane correlations. SLC experience has shown this to be a very difficult problem when provisions are not included in the design to reliably measure and/or correct all four betatron correlation phases. Rather than placing skew quadrupoles and wire scanners as post-design space allows, a dedicated coupling correction and diagnostic section is highly desirable, especially for large emittance aspect ratios as in the NLC ($\varepsilon_{x0}/\varepsilon_{y0} \sim 100$).

11.3.1 The Skew Correction Section

The ideal Skew Correction Section (SCS) contains four skew quadrupoles separated by appropriate betatron phase advance in each plane such that the skew correctors are orthonormal (orthogonal and equally scaled). A simple realization of such a section is possible if the skew quadrupoles each correct one of the four beam correlations $< xy >$, $< x'y >$, $< xy' >$, $< x'y' >$ and if, in addition, the values of the product of horizontal and vertical beta functions are equal at each of the skew quadrupoles. The relative emittance dilution for a thin skew quadrupole of focal length f is

$$\frac{\varepsilon_y}{\varepsilon_{y0}} = \sqrt{1 + \frac{\varepsilon_{x0}}{\varepsilon_{y0}} \cdot \frac{\beta_x \beta_y}{f^2}} \quad , \qquad \frac{\varepsilon_x}{\varepsilon_{x0}} = \sqrt{1 + \frac{\varepsilon_{y0}}{\varepsilon_{x0}} \cdot \frac{\beta_x \beta_y}{f^2}} \quad . \tag{11.1}$$

For a flat beam ($\varepsilon_{x0}/\varepsilon_{y0} \gg 1$), and the vertical emittance is much more sensitive. Orthogonality of the skew quadrupoles is achieved by separating the first and second and also the third and fourth skew quadrupoles by FODO cells with betatron phase advances of $\Delta\psi_x = \Delta\psi_y = \pi/2$, and separating the second and third skew quadrupoles by $\Delta\psi_x = \pi, \Delta\psi_y = \pi/2$. Then if the first skew quadrupole controls the xy phase (by definition here) the second controls the $x'y'$ phase, the third is the $x'y$ phase and the fourth is at the xy' phase. This scheme allows total correction of any arbitrary linearly coupled beam with correction range limited only by available skew quadrupole strength. Initially (at zero strength) the skew quadrupoles are orthonormal. As correction strength is applied there is some deviation from this ideal situation due to the slight effect on the in-plane optics. For large corrections ($\varepsilon_y/\varepsilon_{y0} > 2$). applied by repeated minimization of the projected emittance, some iteration may be necessary (Section 11.3.3). Figure 11-1 shows beta functions and quadrupole locations for the SCS. The full length of the correction section, $L_{\rm SCS}$, is limited by achievable skew quadrupole pole-tip field, $|B_0|_{\rm max}$, its length, l, and its pole-tip radius, r, as well as the maximum correctable vertical emittance dilution, $(\varepsilon_y/\varepsilon_{y0})_{\rm max}$, the nominal emittance aspect ratio, $(\varepsilon_{x0}/\varepsilon_{y0})_{\rm nom}$, and the beam energy, expressed here as magnetic rigidity, $(B\rho)$,

$$L_{\rm SCS} \geq \frac{15}{2} \cdot \frac{r(B\rho)}{l|B_0|_{\rm max}} \sqrt{\frac{(\varepsilon_y/\varepsilon_{y0})^2_{\rm max} - 1}{(\varepsilon_{x0}/\varepsilon_{y0})_{\rm nom}}} \quad . \tag{11.2}$$

Note, Eq. 11.2 is a thin-lens approximation for both the skew and the FODO cell quadrupoles. It does not hold for skew or FODO cell quadrupoles with focal lengths comparable to their magnetic lengths. For a 500-GeV/beam system intended to correct up to a factor of three in emittance dilution, and for an emittance aspect ratio of 100 with 30-cm-long, 6-mm-radius, ±4-kGauss-skew quadrupoles, the system is ~170 m in length. While it would be possible to reduce this length by ~20%, the matching into the diagnostic section (Section 11.3.2) is simplified with the slightly longer design.

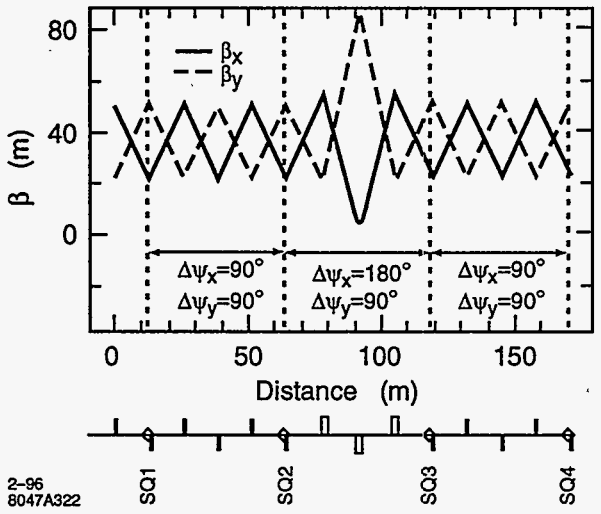


Figure 11-1. *Skew correction section (SCS) optics.*

11.3.2 The Diagnostic Sections

In order to tune the skew quadrupoles and other matching elements an emittance measurement must be made downstream of the correction. In what follows, only emittance measurements made with multiple wire scanners [Ross 1991] are considered. The emittance measurement may be either a single plane (2D) measurement where the coupling is only inferred by the dilution, or it may be a fully coupled (4D) measurement where all four coupling parameters as well as the two intrinsic transverse emittances are measured. The following presents a design for each scheme with ideal wire-to-wire phase advances and also constant matched spot sizes and aspect ratios. In the present NLC design, a 2D section will be used immediately downstream of the post-linac collimation section, while a 4D section will be used in conjunction with a skew-correction section at the input to the final focus.

The 2D Emittance Measurement Scheme

A space-saving, economical 2D emittance measurement system is probably adequate for regions in NLC where low levels of coupling are expected. In this case the optimal wire-to-wire phase advance per plane is π/N, where N is the number of wire scanners.[1] This conclusion is clear by viewing the normalized matched phase space as a circle with beam size measurements made at π/N phase intervals. With three single plane parameters to measure (ε, β, α) a four-scanner measurement provides some redundancy as well as better phase coverage for poorly matched beams. By separating each wire with a FODO cell of $\mu_{x,y} = 45°$, the phase coverage is optimum and, in addition, the matched beam will have a constant size per plane. Furthermore the aspect ratio, a, produced at (or near) a vertically focusing quadrupole is very reasonable and does not require precise wire-scanner roll alignment tolerances.

$$a \equiv \frac{\sigma_x}{\sigma_y} \cong \sqrt{\frac{\varepsilon_{x0}}{\varepsilon_{y0}} \cdot \frac{1-\sin\frac{\mu}{2}}{1+\sin\frac{\mu}{2}}} \approx \frac{2}{3}\sqrt{\varepsilon_{x0}/\varepsilon_{y0}} \tag{11.3}$$

[1] The arguments presented here are the same whether these are actual carbon-filament-wire scanners or multibunch-capable laser-wire scanners.

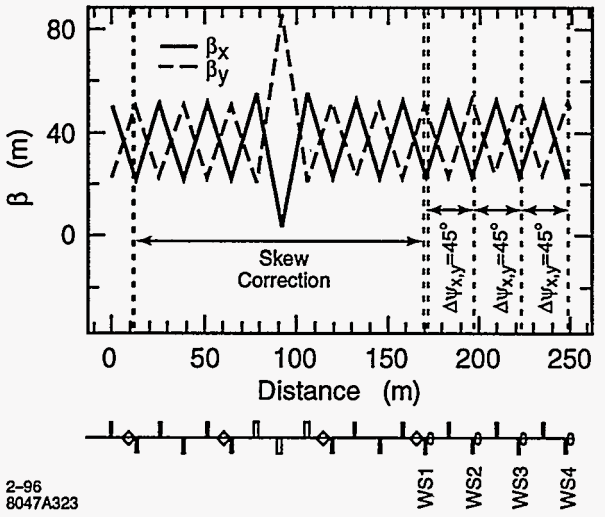


Figure 11-2. *SCS plus 2D emittance diagnostics section optics.*

In the case of the NLC $a \approx 6.7$ which requires an easily achieved roll tolerance[2] of $|\psi| \leq \sin^{-1}\left(\sqrt{2}/10\sqrt{3}a\right) \sim 1°$. Each wire scanner need only measure the x and y beam sizes (the correlation is not necessary in this 2D scheme). Figure 11-2 shows the lattice which includes the SCS followed by a matched 2D/four-scanner diagnostic section. The coupling is fully corrected by repeatedly minimizing the measured vertical emittance with each skew quadrupole in turn (some iteration may be necessary). Such a system should probably be preceded by four beta-matching quadrupoles to facilitate compensation of the slight in-plane focusing effect of energized skew quadrupoles. A clear advantage to this scheme is that the uncoupled, matched beam will be clearly evident as four precisely equal x and four precisely equal y beam sizes on the scanners. The lower limit on the length of the diagnostic section, L_{2D}, is set by the minimum measurable beam size, σ_y, the normalized emittance, ε_{yN}, and the beam energy, γ.

$$L_{2D} \geq \frac{3}{2}\frac{\sigma_y^2 \gamma}{\varepsilon_{yN}} \tag{11.4}$$

For a 500-GeV beam with ~1.5-μm vertical beam size and a normalized emittance of 5×10^{-8} m, the minimum length is ~70 m.

The 4D Emittance Measurement Scheme

If it is expected that coupling may be a significant problem or the application of a faster, calculated correction is desirable, the 4D measurement scheme may be preferable. In this case a precise measurement of the four linear coupling coefficients is made with a set of σ_x, σ_y, and σ_{xy} measurements made for each of six wire scanners. Each wire scanner must have three independent angle filaments—in the simplest case, a horizontal, a vertical and a 45° filament typically used in the SLC [Ross 1991]. The wire-to-wire phase advances are chosen in a similar way to the skew quadrupole placement in the SCS (Section 11.3.1). With the first wire measuring the xy correlation, a second wire at $\Delta\psi_x = \Delta\psi_y = \pi/2$ measures $x'y'$ and a third wire advanced again by $\Delta\psi_x = \pi, \Delta\psi_y = \pi/2$ measures $x'y$ and finally a fourth at another $\Delta\psi_x = \Delta\psi_y = \pi/2$ sees xy'.

[2]The following relation holds only for $a >\sim 3$ and achieves a systematic emittance measurement error of <2%.

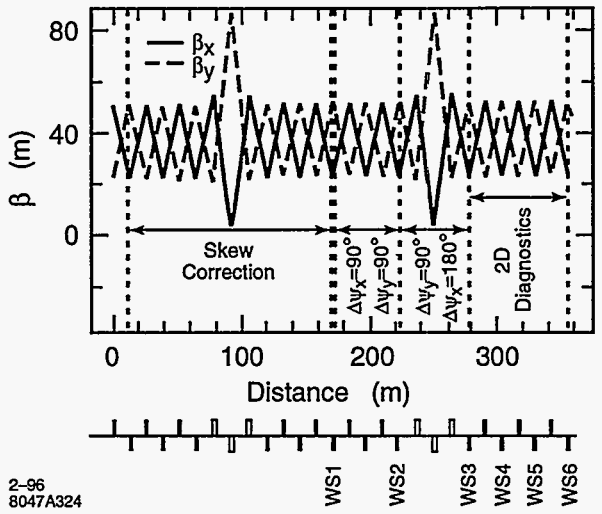


Figure 11-3. *SCS plus 4D emittance diagnostics section optics.*

The problem with this four-wire scheme is that the in-plane measurements are not determined. A fifth and sixth wire must be added to cover the missing single plane phases. Figure 11-3 shows the six-wire solution including the SCS, which is ideal for both the single plane and the coupling measurements. There are 10 parameters to measure ($\varepsilon_{x,y}$, $\beta_{x,y}$, $\alpha_{x,y}$, $< xy >$, $< x'y' >$, $< x'y >$ and $< xy' >$) and up to 18 profiles are scanned leaving 8° of freedom. It is also possible to use a subset of the 18 profiles in order to speed the measurement process. Note that this system can also be used as a 2D system by making single plane profile scans with wires WS3, WS4, WS5 and WS6 of Figure 11-3. The minimum length is approximately twice that of the 2D system.

$$L_{4D} \geq \frac{\pi \sigma_y^2 \gamma}{\varepsilon_{yN}} \tag{11.5}$$

11.3.3 Tuning Simulations

Tuning simulations were run using the Final Focus Flight Simulator [Woodley 1994] to test the convergence of the skew correction for a 500-GeV beam with an intrinsic emittance aspect ratio ($\varepsilon_{x0}/\varepsilon_{y0}$) of 100. Figure 11-4 shows the relative vertical emittance achieved with each skew quadrupole scan. In this case, with the input coupling diluting the vertical emittance by more than a factor of two, a correction to less than 1% dilution is achieved after the first pass through all four skew quadrupoles. This beam line setup operation could probably be executed in less than an hour of real machine time. Of course, using the 4D system here could conceivably produce a reasonably well-calculated correction within a period of minutes. The final choice of appropriate systems depends on the expected phase space stability and magnitude of the errors.

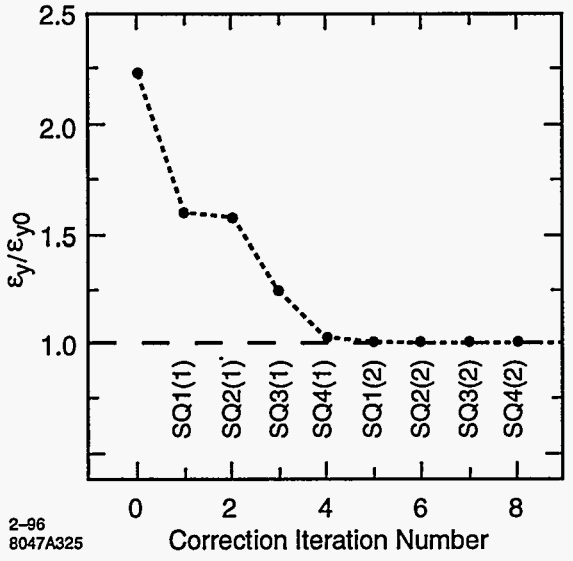


Figure 11-4. *Skew tuning simulation results.*

11.4 Beta-Matching Section

The purpose of the NLC Beam Delivery Beta-Matching section is to provide a set of magnets upstream of the horizontal chromatic correction section which can be used to respond to variations in the incoming beam and/or the desired beam parameters at the IP, and also can be used to set the desired phase advance between the collimators in the Post-Linac Collimation region (Section 9.4) and the Final Doublet. Experiences at the SLC Final Focus and the Final Focus Test Beam have demonstrated that it is useful to be able to reduce the angular divergences at the IP (and hence the beam sizes inside all the other magnets in the Final Focus), while maintaining the chromatic correction and final telescope optics which are used in collision. Reducing the IP divergence is also useful during the early stages of a run cycle, when the emittances in the beam delivery region may be too large to permit collisions at the design betatron functions. Finally, the beta-matching quadrupoles can be used to correct various mismatches between the design incoming beam and the beam which is actually delivered from the linac.

In the design of the NLC Beta Match, it has been assumed that coupling correction and measurement of the incoming beam phase space has been accomplished in the preceding Decoupling and Diagnostics section (Section 11.3), leaving only the problem of the uncoupled beam matching. The design constraints on the beta-matching region can be summarized as follows:

- The region should contain at least six quadrupoles, in order that the six independent first-order transfer matrix values to the IP can be independently adjusted, and that for a given input beam and set of Twiss parameters there is a solution which provides the desired output conditions and phase advances.
- The quadrupole specifications are consistent with warm-iron devices which can be easily manufactured and do not introduce emittance growth due to wakefields. Tolerances on field quality and mechanical stability should be no tighter than those upon quadrupoles in the chromatic correction sections. In

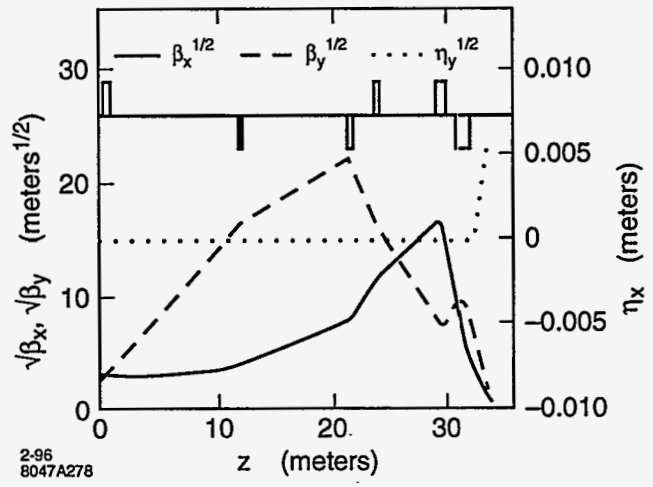


Figure 11-5. *Schematic representation of the SLC Final Focus Beta Matching region.*

general, the magnet specifications should be consistent with the technology and techniques of the rest of the beam delivery section.

- The contribution of the beta-matching region to overall chromaticity should be small enough to be easily cancelled out by reasonable increases in chromatic correction sextupole strengths.
- The range of accessible solutions for beam conditions at the IP should be as large as possible. Independent control of magnification and waist position control in x and y is essential.
- A dispersion-free pre-image point of the IP should be provided in each final focus, at which the magnification and waist position can be verified. This is necessary because the resolution of the incoming beam phase space is rarely adequate to match to collision conditions without such pre-IP verification. In addition, high-resolution (rf) BPMs that are placed close to the pre-image point will allow to correct the beam-beam deflection scans for pulse-to-pulse trajectory jitter in the IP betatron phase. Similar jitter-correction techniques have proven very successful at the SLC [Raimondi 1995] and are also being tested at the FFTB.
- The system should be as short, simple, and robust as possible.

11.4.1 Optical Design

At this time, there are two linear collider final-focus systems in existence, each of which has a unique beta-matching system. The SLC Final Focus uses a cluster of quadrupoles to focus on an image of the IP in the center of the first chromatic correction bend (Figure 11-5). A wire scanner in the center of the bend is used for match verification and tuning. In this case, the image is an existing symmetry of the design, and no additional quads are necessary to match from the image onto the CCS proper. The Final Focus Test Beam has no such naturally-occurring image point, and the five quadrupoles of its beta-matching section are used to focus the beam onto a pair of wire scanners in between the two chromatic correction sections (Figure 11-6). The SLC design is a more conventional mix of quadrupoles and modest drift regions, while the FFTB design precedes the first matching quadrupole with a 120-m drift space.

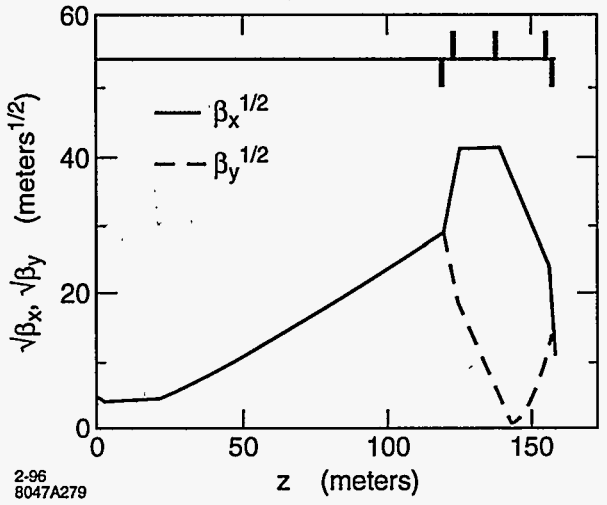


Figure 11-6. *Schematic representation of the Final Focus Test Beam Beta Matching region.*

The current design of the NLC Final Focus includes several images of the IP, including a dispersion-free image in the beta exchange region (between the horizontal and vertical chromatic correction sections). The NLC Beta Matching section has been designed to resemble the FFTB system, including the use of the beta exchange image for tuning the match. This is for several reasons. First, the optics of the beta exchanger guarantees the presence of horizontal and vertical IP images which can be exploited for match verification, while use of an upstream image constrains the design of the dispersion matching section. In addition, the energy spread of the NLC is sufficient that chromatic correction of the beam at the IP image will be necessary during tuning. This can only be done downstream of the first chromatic correction section, by retuning the sextupoles in that section. As it is, only one plane can be corrected at a time at the beta exchanger tune-up point; and the sextupole field required to correct the vertical chromaticity is opposite in sign to that used in normal running, necessitating a reversible power supply for the CCX sextupoles.

The optical functions of the NLC Beta Match are shown in Figure 11-7. The current design includes a total of seven quadrupoles, which satisfies the numerical requirement set forth above.

The total length of the system is 254 m. The maximal value of β_x within the section is 5,150 m, while the maximal value of β_y is 2,530 m. These values are far smaller than the typical values within the downstream region of the Final Focus, and the tolerances are expected to be correspondingly looser. Note that the beta-match region contains a drift of length 100-m upstream of the maximal betatron functions. Such a drift is an ideal location for a Single Beam Dumper (SBD), a kicker magnet capable of extracting individual bunch trains (*i.e.*, firing at 180 Hz). Extracting the beam after the Big Bend and Decoupling and Diagnostic sections would allow the beam to pass through the collimation section and the diagnostic wire scanner regions, while preventing it from entering the sensitive detector area. Thus, even beams which are not of sufficient quality to pass through the doublet into the detector can be maximally measured for diagnostic purposes if such an extraction magnet is placed at this location.

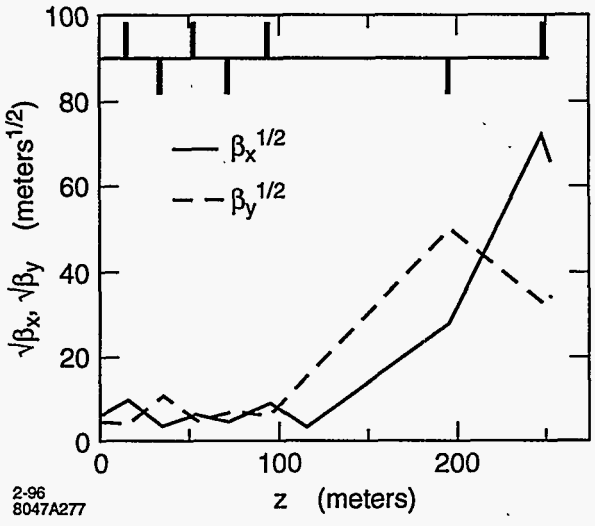


Figure 11-7. *Betatron functions of the beta-matching section.*

11.4.2 Tuning Elements

The tuning elements of the beta match include the seven normal quadrupoles in the region itself; the horizontal chromatic correction sextupoles (referred to as SX1 sextupoles); a beam size monitor at the IP image in the beta exchanger; and a divergence monitor just upstream of the vertical chromatic correction section.

Beta Matching Quadrupoles

Table 11-2 summarizes the design specifications of the seven quadrupole magnets in the beta-matching section. The magnets are physically identical with one another, with a length of 1 m and an aperture of 6 mm. The pole-tip fields for 1-TeV-c.m. go to a maximum of 6.5 kGauss for luminosity operations, and 7.3 kGauss for low-divergence operations ($\beta_x^* = 10\,\text{cm}, \beta_y^* = 1\,\text{cm}$); the same optics at the 1.5-TeV c.m. would require 10.9-kGauss pole-tip fields, and therefore these magnets are not usable at 1.5-TeV c.m. For all IP conditions in the operating plane, 6 mm is sufficient to maintain clearances in this region: worst-case rms beam sizes are 250 microns in the horizontal and 30 microns in the vertical.

The precise tunability of the beta-match design, in terms of range of IP betatron functions which can be delivered, range of waist knobs, and orthogonality of knobs, has not been studied for this design. A detailed study of the tunability should be carried out for final acceptance of this design.

The SX1 Sextupoles

Table 11-3 shows the beam size at the IP image location in the center of the beta exchanger with different sextupole configurations. In order to achieve the monochromatic beam size in the horizontal, the main CCX sextupoles (SX1) need to be set to a value roughly 33% of their design, while achieving the monochromatic vertical size requires a value 130% of design in the opposite polarity. This in turn indicates that the CCX

Quadrupole Name	Pole-tip Field (kGauss)	Length (m)	Aperture (mm)
QB1	5.9	1.0	6.0
QB2	-6.5	1.0	6.0
QB3	5.6	1.0	6.0
QB4	-3.2	1.0	6.0
QB5	6.4	1.0	6.0
QB6	-1.6	1.0	6.0
QB7	3.8	1.0	6.0

Table 11-2. *NLC beta-match quadrupole specifications.*

Conditions	σ_x at IP Image (μm)	σ_y at IP Image (μm)
Monochromatic Beam	1.85	0.20
Nominal Sexts, $\sigma_E/E = 0.003$	31.1	0.63
Sexts Off	14.8	0.41
SX1 = 33% design	2.00	0.48
SX1 = -130% design	74.1	0.20

Table 11-3. *Beam sizes at IP image location under varying conditions.*

sextupoles require reversible power supplies, with a capability to deliver significantly more current than is required for normal operations. The requirement of tuning the two planes with different sextupole strengths is also onerous in terms of time and difficulty level. While tuning the IP Image spot is the best tune-up procedure possible, next we will discuss a short-cut to this.

IP Image and Beam Size Monitor

The IP image at the center of the Beta Exchanger is a true image, with a pure demagnification in the horizontal and the vertical between it and the IP. The demagnifications are $(8.184)^{-1}$ in the horizontal, and $(57.65)^{-1}$ in the vertical. The NLC operating plane indicates that beam sizes from 3.5 nm to 7.0 nm may be required at the IP, which in turn requires vertical beam sizes of 200 to 400 nm at the image. The horizontal beam sizes range from 226 nm to 320 nm at the IP, indicating sizes at the image from 1.85 microns to 2.6 microns. The horizontal beam sizes achievable at the image are within the expected range of laser wire technologies, while the vertical sizes are pushing the limits of these same technologies, and entering the range of Laser-Interferometer beam size monitors.

While the measurement of the horizontal corrected size in Table 11-3 presents relatively few problems, the vertical corrected size can only be achieved with a tremendous dilution of the horizontal spot, leading to an aspect ratio of 371:1. Measurements of the vertical beam size will therefore be tremendously sensitive to installation roll of the monitor: a roll of 1 mr will result in a contribution of 74 nm added in quadrature with the 200 nm of the focused vertical spot. For this reason as well as those outlined above, it may be impractical to use the IP image monitor to measure the size of the vertical beam; rather, the optimal use might be to

use the monitor to constrain the vertical waist, and use the divergence monitor (see next Section) to verify the vertical betatron function at the IP image.

Angular Divergence Monitor

In addition to beam size measurements at an image of the IP, it is useful to be able to measure the horizontal and vertical angular divergences in this area. The divergences, along with the emittances measured upstream of the beta-match region, give the best estimate of the actual betatron function at the IP image monitor, as the divergence is insensitive to virtually all aberrations.

The ideal tactic for an angular divergence measurement is to measure the beam size at a point which is dominated by the divergence of the IP or IP image, and with a known transfer matrix to the IP or IP image. In this case, such a location is immediately upstream of the vertical chromatic correction section, at the end of the beta exchanger. The beam size at this location is 206 μm (x) by 140 μm (y) in the 1-TeV-c.m. design with 10 $\times$ 0.125-mm betatron functions at the IP and 50:1 emittance ratio; the beam size at this location is not drastically different for any optics in the operating plane, allowing a conventional metal-wire scanner such as the SLC linac scanners to be employed. However, the horizontal beam size listed here is significantly enlarged by dispersion present at this location. Therefore, an additional scanner with spectrum-measurement capability is needed to complete this measurement.

Because of the difficulties in measurement at each scanner, there are two conceivable schemes for tuning the incoming beam: one which "Caps the Ts and Dots the Is (CTDI)" and one which is "Close Enough for Government Work (CEGW)."

The CTDI tune-up is as follows:

- Measure incoming phase space with diagnostic wires (Section 11.3).
- Compute match optics for desired IP/IP Image conditions.
- Tune SX1 sextupoles to 33% design, tune x waist on image monitor.
- Measure beam on spectrum monitor and SY1 wire, compute monochromatic contribution at SY1 wire, compute x divergence.
- Use divergence computation to compute betatron function at IP image, correct with magnification knob, repeat waist for verification.
- Tune SX1 sextupoles to -130% design, tune waist on image monitor.
- Measure roll of IP image monitor spot (if possible), correct (if possible).
- Repeat waist scan on IP image monitor.
- Measure vertical divergence on SY1 monitor, compute IP image betatron function, correct if necessary with magnification knob.
- Repeat waist scan.
- Tune SX1 to colliding-beam strengths, go to IP.

The CEGW algorithm is the following:

- Measure incoming phase space with diagnostic wires (Section 11.3).
- Compute match optics for desired IP/IP Image conditions.
- Tune SX1 sextupoles to 33% design, tune x waist on image monitor; use waist scan to compute divergence.
- Use size of horizontal IP image spot and/or waist scan divergence value to tune magnification knob, iterate waist.
- Set SX1 to design strength.
- Tune y waist on IP image monitor to constrain waist location, ignore minimum spot size.
- Measure vertical beam size on SY1 monitor, compute vertical betatron function at IP Image monitor; tune magnification knob to correct SY1 size to predicted value.
- Iterate waist tuning if necessary, go to IP.

The CEGW algorithm makes maximal use of the diagnostic devices to bypass measurements which are difficult or time-consuming to make. A similar tuning algorithm in FFTB converges in a matter of hours; similar speed can be expected for tuning the beta match in NLC.

11.4.3 Tolerances

At this time, no detailed study of the beta-match section had been undertaken. However, based on the optics and some understanding of the functions of various magnets, it is possible to say a few things about operational tolerances of the system.

Field Strength Tolerances

The absolute accuracy tolerances on the beta-match quadrupole strengths are quite forgiving. The quadrupoles are set by matching the measured incoming beam to the desired IP conditions. Because the incoming phase space measurements are unlikely to have a greater precision than 1%, absolute quad strength accuracy of 0.1% will cause beam mismatch to be dominated by the errors in measurement, rather than errors in the magnets. The final settings of the magnets will be determined by tuning the IP image and/or the IP itself. This may require a granularity of the field strengths at the level of 0.01% of the maximum (assuming a maximum of 9.5-kGauss pole-tip fields). Stability of the fields with time will need to be approximately the same as the requisite granularity of the field.

Multipole Content Tolerances

Because the beam is much smaller in the beta-match quadrupoles than it is in other final-focus magnets, tolerances on higher multipole content will be correspondingly looser.

Vibration Tolerances

The beta-match quadrupoles are upstream of the first sextupole of the chromatic correction sections, and therefore these magnets are unable to induce steering between the sextupoles of a section. Tolerances for vibration will therefore be determined by dispersion and steering at the IP.

11.5 Chromatic Correction and Final Transformer

This beam line extends from the beta-matching section to the IP, and it comprises horizontal chromatic correction, beta-exchanger, vertical chromatic correction, and final transformer. Located at the end of the final transformer is the final doublet, which provides the last focusing before the beams collide (see Section 11.6 for a detailed discussion of the final doublet). The chromaticity of the final doublet is similar to that at the FFTB, and about five times as large as in the SLC. If the chromaticity were not corrected, the vertical IP spot size would increase by about a factor of 100 from the design value. The chromatic correction is accomplished in two separate beam-line sections, CCX and CCY, each of which accommodates a pair of sextupoles separated by an optical transform ($-I$) that cancels geometric aberrations and second-order dispersion. The chromatic correction is very similar to the FFTB, but it is different from the SLC in which the two sextupole pairs are interleaved. An interleaved placement of sextupoles is not acceptable for the NLC, because of the large higher-order aberrations this would generate. The bending section upstream of the beta-matching section is an innovation, unfamiliar from either SLC or FFTB. It allows the final focus to operate for any beam energy between 175 GeV and 750 GeV, without change of IP position, and hence it is known as "geometry-adjustment" section.

To obtain a proper estimate of the NLC luminosity, it is important to account for all possible sources of spot-size dilutions. To limit the total dilution, a budget was established for each effect contributing to the IP spot size, and, in particular, for each low-order aberration originating between the linac and the interaction point. Within this section, we present these dilution budgets, and discuss the implied tolerances on element strengths and positions, the frequency and accuracy of aberration tuning, and also the maintenance systems which stabilize the final focus between tunings. We believe that all the requirements can be met, and that there is even a potential for a further increase of the luminosity beyond that estimated here.

11.5.1 Introduction, Parameters and Dilutions

Final-focus and IP-beam parameters for different NLC scenarios are listed in Table 11-4. These parameters, in particular the cases Ia and IIa, form the platform on which the following discussion will be based. As indicated by the table, the final-focus designs for 500 GeV and 1 TeV deliver the desired horizontal and vertical spot sizes of about 250 nm and 4–8 nm, respectively. Both nominal luminosity and spot sizes quoted include various sources of dilution which are summarized in Table 11-5. The total and subtotals in Table 11-5 are calculated according to the formulae

$$1 - d_{subtot_j} = \Pi_{i \text{ in}_j} (1 - d_i) \tag{11.6}$$

$$1 - d_{tot} = \Pi_j \left(1 - d_{subtot_j}\right) \tag{11.7}$$

where d_i denotes the different relative luminosity dilutions.

To further understanding of Table 11-5: some supplementary comments on each table subsection are given below:

- **Emittance Growth.**
  The phase of the beam-line modules with respect to the IP must be very well-regulated and tuned for the beam line to function as designed. This can be easily monitored and arranged by studying betatron oscillations through the system. As a result of this phase regulation, elements within different beam-line modules affect the beam distribution in a way that is phase-related. What is of concern is the axis of the vertical (or horizontal) emittance in the IP phase. This we have denoted in Chapter 9 by the symbol $r_{y,IP}$. (If the beam distribution is looked at in a normalized phase space, where the beta functions are factored out, the distribution is normally assumed to be round, of radius $r_y = \sqrt{\epsilon_y}$. What we want to emphasize here is that the distribution will not remain round. The collimators in the FD phase, for example, give a kick in the IP phase, and enlarge the distribution along that axis only.) Emittance growth is not only generated by wakefields but also by synchrotron radiation, chromaticity, and higher-order optics. All of the entries in this subsection of the table refer to the enlargement of the IP axis of the emittance.

- **Collision.**
  The jitter is assigned according to the jitter budget proposed in Chapter 9 The calculation of the luminosity loss assumes that for the same luminosity loss the vertical missteering can be twice as large than for a rigid Gaussian bunch by virtue of the disruption. This is well-documented from simulations even for bunch charges that are a factor of 2 smaller than the design.

  Crab-cavity phase jitter (which is jitter of the positron beam cavity with respect to the electron-beam cavity) introduces an additional component to the horizontal jitter. So do field variations of the bending magnets, and quadrupole vibrations.

- **Tuned Low-order Aberrations.**
  In particular five aberrations—waist (x and y), dispersion (x and y) and skew coupling—must be scanned and corrected in regular intervals. (We do not assume a sublime dither technique.) Other aberrations will be more stable and can be scanned less frequently.

  Based on SLC experience, we expect to tune aberrations so that only a residual 0.5% luminosity loss per aberration remains. This number could conceivably be improved by developing a sensitive luminosity monitor for beam tuning purposes. In Table 11-5, the existence of such a monitor has been assumed neither for current nor for possible columns.

  We have assumed that the tuning of the five major aberrations is performed every 15 min, and that without tuning the design stability tolerances result in a 2% additional luminosity loss per aberration after 1 h. We expect that the aberrations increase from the minimum setting tolerance (0.5% per aberration) directly after tuning to the minimum plus the product of (time between tunings/1 h) and the budgeted spot-size increase for that aberration (as determined by system stability tolerances; typically 1–2%), added quadratically since they are independent.

  For the δ-dependent aberrations (chromaticity, dispersion, and chromatic skew), we presume that a smaller minimum (namely 0.25%) can be achieved by doubling the bunch energy spread for these scans.

  Dispersion has an effect on beam size due to energy spread within the bunch (an effect which is included under table subsection Tuned Aberrations) and also on beam jitter, since the individual bunches in a train and train-to-train will have different energies (an effect which is considered in the table subsection Collision). Insofar as the dispersion arises in the beam-delivery system, these two dispersions will be identical, and so the dispersion changes leading to increased beam size may be monitored by observing the correlation of position-jitter at the IP (which is being measured for each

bunch train) with the energy of the bunch train (as can be measured for each train by observing the orbit in the final-focus system). Thus the dispersion aberration can be tuned almost continuously and does not require allocation of a growth due to time between scannings. This potential improvement has not been included in Table 11-5.

It should be noted that dispersion can be present in the linac, and since inter-bunch energy spread and intra-bunch energy spread have different sources, and arise in different locations in the linac, it may occur that the beam size and the jitter due to dispersion cannot both be tuned together. In this case, one could try to tune by introducing orbit changes in the linac. However, this effect is considered an emittance dilution that originates in the linac, and not in the beam-delivery system, and thus we have not added it to our table.

Note that we equally have not added any aberration-induced dilutions originating in the collimation system. The reason is that such dilutions can be monitored non-invasively in the skew-correction section and can be corrected on a more frequent basis.

When an aberration affects both the vertical and the horizontal spot size, only the vertical effect needs to be considered, because, when the latter is tuned, the horizontal effect will be negligible as a result of the 100/1 spot-size ratio.

The crab-cavity adjustment refers to the voltage stability of the crab cavity. A 2% effect corresponds to a 6% voltage stability. We suppose that it can be tuned to 0.5% using a voltage scan, and then that it will be stable to 0.5% between tunings.

- **Other.**
  The entry for e^+e^- arrival time assumes an rms timing difference of the two beams equal to 170 fs (corresponding to 0.2° S-Band or 50 μm). This implies that the two bunches will be $0.2\beta_y^*$ offset when colliding at the IP.

  Under table subsection Other, we have also listed the time devoted to beam-based alignment and tuning. Losses from tuning scans are calculated under the assumption that the five major aberrations are scanned every 15 min, the control system is designed to minimize the control-time overhead, the magnet settle time is 200 ms per field change, an aberration scan consists of seven steps with different magnet settings, and that at each step a beam-beam deflection scan over 50 pulses is performed.

  The time-out for beam-based alignment presumes the alignment to be completed in 24 h and to be executed every three months.

Table 11-5 lists two budgets of luminosity dilutions. The first refers to the Current Lattice and is based on tracking; the second is called Possible and assumes:

1. A reduced β_x^* which compensates the increase of the horizontal spot size due to design and tuning aberrations. This is consistent with the fact that the horizontal design spot size was chosen to limit the number of beamstrahlung-photons per electron radiated during collision. We do not suppose that the collision effect can be offset in this way, since it is not related to the beam size.

2. A redesign of the final-focus system which diminishes the effect of synchrotron radiation from bending magnets and final quadrupoles, We have learned during the design process, and we have reason to believe, that we can improve the performance of these modules. In many cases we have actually demonstrated this improvement.

3. Better control and adjustment of crab-cavity phase and voltage.

4. An improvement in the chromatic properties of big bend and IP switch.

The dilutions for the nominal lattice amount to a 46.6% overall loss in luminosity, which can possibly be reduced to a total of only 25.9% by the aforementioned means.

In Table 11-4, we have considered the nominal lattice, but have attributed a 13% spot-size increase to residual uncorrected vertical aberrations which is about twice the value listed in Table 11-5 and have assumed that the increase of the horizontal spot size due to residual tuned aberrations and horizontal steering (though not that due to horizontal emittance growth!) is compensated by a reduced β_x^*. The expected increase of the spot-size product $\sigma_x \times \sigma_y$ then amounts to 50%, for 1-TeV-c.m. energy, and translates into a total luminosity loss of about 33%, which lies between the Current and the Possible value of Table 11-5.

The momentum bandwidth of the NLC final focus, defined by a 10% increase of either spot size for a mono-energetic beam, is at least ±0.6% for both c.m. energies (500 GeV and 1-TeV c.m.), even assuming the most difficult parameter set of the ZDR operating plane. The final focus is very forgiving in regard to increased emittances or to orbit variations: If β_y^* is held constant, the luminosity decreases roughly as the square root of the emittance, while the momentum bandwidth remains unaffected by the larger divergence. A 0.5-sigma variation of the incoming orbit causes an average spot-size increase by less than 1%.

The length of the final-focus system, from the geometry-adjustment section to the IP, is about 1800 m. This length is independent of energy, and is determined by the requirements for operation at 1.5-TeV-c.m. energy. Optimization of the final-focus system calls for a different dispersion and thus for a different bending angle at different c.m. energies. The different bending angle helps to balance nonlinear aberrations and synchrotron radiation effects at each energy. A dedicated geometry-adjustment section at the entrance to the final focus keeps the resulting geometry change at an acceptable level and the IP position constant. The horizontal magnet displacements required during an energy upgrade from 350 GeV to 1.5 TeV do not exceed 45 cm. The following discussion mainly refers to the two final-focus systems at 500 GeV and 1-TeV-c.m. energy with equal geometry, *i.e.*, Version II in Figure 11-10.

11.5.2 Layout and Optics

General Description

The basic layout of the proposed NLC final-focus system is very similar to that of the Final Focus Test Beam. In the region between the diagnostics section (DS) and the IP, the NLC final focus is constructed from six functional modules. These are, in the order of their location: geometry-adjustment section (GAS), beta- and phase-matching section (BMS), horizontal chromatic correction section (CCX), beta-exchanger (BX), vertical chromatic correction section (CCY), and final transformer (FT). Two conventional magnets, one superconducting and one permanent quadrupole at the end of the FT (the final doublet, or, perhaps more appropriately, the final quartet) provide the last focusing before the two beams collide.

The total distance from the entrance of the GAS to the IP is about 1820 m. A schematic of the magnet configuration is depicted in Figure 11-8, which also shows the beta functions corresponding to the 500-GeV parameter set listed in Table 11-1 (Case Ia in Table 11-4). In this case, the vertical beta function has a maximum value of 190 km at the main sextupoles in the CCY and peaks at about 75 km in the final doublet. The maximum value of the horizontal beta function is 75 km, at the CCX-sextupoles. A large beta function at the sextupoles is advantageous for chromatic correction, considering the effects of synchrotron radiation, chromo-geometric aberrations, and orbit-stability tolerances [Zimmermann 1995]. The length of the CCX is about two-thirds that of the CCY, which reflects the larger horizontal beta function at the IP. Note that, per a recent design modification, the beta-matching section has been augmented by two additional quadrupoles,

	c.m. energy	0.5 TeV			1 TeV		
		Ia	Ib	Ic	IIa	IIb	IIc
Luminosity w. dilution	L (10^{34} cm^{-2} s^{-1})	0.41	0.40	0.41	0.77	0.82	0.73
Lum. w. dil. & pinch	L (10^{34} cm^{-2} s^{-1})	0.58	0.55	0.60	1.02	1.10	1.06
Luminosity w/o dilution	L (10^{34} cm^{-2} s^{-1})	0.52	0.49	0.51	1.16	1.20	1.02
Lum. w. pinch w/o dil.	L (10^{34} cm^{-2} s^{-1})	0.75	0.71	0.78	1.63	1.72	1.56
# particles per bunch	N_b (10^{10})	0.65	0.75	0.85	0.95	1.1	1.25
# bunches	n_b		90			90	
# bunch trains per s	f		180			120	
Enhancement factor	H_D	1.41	1.40	1.48	1.33	1.35	1.46
Hor. spot size w. dil.	σ_x (nm)	264	<294	<294	231	<250	<284
Vert. spot size w. dil.	σ_y (nm)	5.07	<6.26	<7.76	4.35	<5.08	<6.52
Hor. spot size w/o dil.	σ_x (nm)	253	283	283	200	219	253
Vert. spot size w/o dil.	σ_y (nm)	4.18	5.20	6.48	3.35	3.97	5.20
Hor. IP beta function	β_x^* (mm)	8	10	10	10	12	16
Vert. IP beta function	β_y^* (mm)	0.125	0.15	0.2	0.125	0.15	0.2
Norm. hor. emittance	$\gamma\epsilon_x$ (10^{-8} mr)		400			400	
Norm. vert. emittance	$\gamma\epsilon_y$ (10^{-8} mr)	7	9	10.5	9	10.5	13.5
Hor. emittance growth	$\Delta(\gamma\epsilon_x)_{SR}$ (10^{-8} mr)		~ 23			~ 38	
Hor. IP divergence	$\sigma_{x'}$ (μrad)	31.6	28.3	28.3	20.0	18.3	15.8
Vert. IP divergence	$\sigma_{y'}$ (μrad)	33.5	34.6	32.4	26.8	26.5	26.0
Bunch length	σ_z (μm)	100	125	150	125	150	150
Crossing angle	θ_c (mr)			20			
Rms energy spread	$\delta_{\rm rms}$ (%)			≥ 0.3			
Total energy bandwidth	$\Delta E/E$ (%)		≥ 1.2			≥ 1.2	
Free length from IP	l^* (m)			2			

Table 11-4. *Interaction-point beam parameters for different NLC scenarios. Spot-size increases are scaled from NLC-Ia and NLC-IIa. Luminosity enhancement is estimated from P. Chen's scaling laws. Nominal luminosity and nominal spot sizes include dilutions due to synchrotron radiation (0–16%), high-order aberrations (7%), residual uncorrected low-order aberrations (here assumed as 13%), e^+e^- timing offset (2%), crab crossing error (2%) and vertical orbit jitter (3%); see also Tables 11-5 and 11-9. Effect of residual low-order aberrations and horizontal orbit jitter on the horizontal spot size is assumed to be compensated by a reduced β_x^*. The numbers quoted for luminosity and spot size "without dilution" (w/o) are calculated according to linear optics and do not include any of the additional effects.*

Category	Description	Current Lattice		Possible	
		Horizontal	Vertical	Horizontal	Vertical
Emittance Growth (*actually* r_{IP})		**20.6%**	**11.8%**	—	**8.1%**
	Collimation System	3.0%	2.0%	—	2.0%
	IP Switch, Big Bend, SCS	3.0%	1.6%	—	1.0%
	Final Focus	15.4%	7.3%	—	4.0%
	Uncorrected Aberrations	6.0%	1.5%	—	1.5%
	Synchrotron Radiation	10.0%	5.9%	—	2.5%
	In Dipoles	10.0%	2.1%	—	0.1%
	In Quadrupoles	0.0%	3.5%	—	2.0%
	In Solenoid	0.0%	0.4%	—	0.4%
	Wakes (res.-wall & geom.)	0.2%	1.34%	—	1.34%
	Protection Collimators	0.1%	0.5%	—	0.5%
	Main Collimators	0.1%	0.65%	—	0.65%
	Misalignment	0.0%	0.25%	—	0.25%
	Jitter Wake	0.1%	0.3%	—	0.3%
	Quadrupole Wake	0.0%	0.1%	—	0.1%
	IR Beampipe	0.0%	0.2%	—	0.2%
Collision		**5.1%**	**3.5%**	**4.1%**	**3.5%**
	Incoming (to BDS) Jitter	0.5%	0.8%	0.5%	0.8%
	Jitter Amplification of Coll.	0.6%	0.7%	0.6%	0.7%
	Quad Motion	0.8%	1.5%	0.8%	1.5%
	Bend Power Supply Ripple	0.7%	0.0%	0.7%	0.0%
	Crab Cavity Phase-diff. Jitter	2.0%	0.0%	1.0%	0.0%
	Dispersion (inter-bunch & centroid)	0.5%	0.5%	0.5%	0.5%
Tuned Aberrations		**3.0%**	**5.5%**	—	**5.5%**
	Linear	2.5%	4.0%	—	4.0%
	Waists	1.0%	1.0%	—	1.0%
	Skew Quad (2)	0.0%	1.5%	—	1.5%
	Dispersion (intra-bunch δ)	0.5%	0.5%	—	0.5%
	Crab-angle Adjustment	0.5%	0.0%	—	0.0%
	δ-dependent 'linear'	0.5%	1.0%	—	1.0%
	Chromaticity	0.5%	0.5%	—	0.5%
	Chromatic Skew	0.0%	0.5%	—	0.5%
	Higher Order	0.5%	1.5%	—	1.5%
	Normal & Skew Sextupole	0.5%	1.5%	—	1.5%
Subtotal:		**26.9%**	**19.6%**	**4.1%**	**16.2%**
Other (*down time not included!*)		**7.0%**		**6.0%**	
	e^+-e^- arrival-time diff.	2.0%		1.0%	
	Losses from tuning scans	4.0%		4.0%	
	Time-out for bm-based alignment	1.0%		1.0%	
Total:		**46.6%**		**25.9%**	

Table 11-5. *Expected luminosity losses from dilutions in the beam delivery systems for 1-TeV-c.m. energy. See also Table 11-9 for a detailed account of low-order aberrations, and Table 11-8 for the effect of synchrotron radiation and higher-order aberrations.*

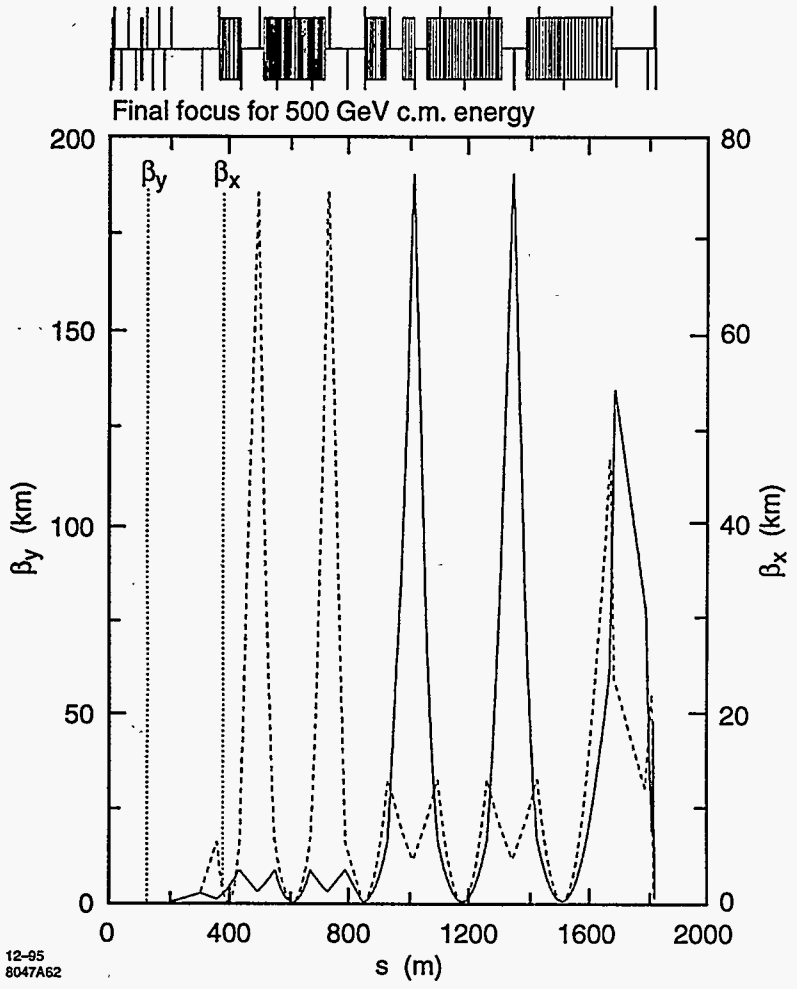


Figure 11-8. *Horizontal and vertical beta functions from BMS to IP, for the 500-GeV final focus.*

to allow matching of the betatron phase advance between the collimator section and the IP. These two new quadrupoles and the accompanying minor optics change of the BMS are not included in the following discussion.

The optics at 1-TeV-c.m. energy is almost the same as that for 500 GeV. Again assuming the parameters of Table 11-4 (Case IIa), the peak values of horizontal and vertical beta functions at the CCX sextupoles are 60 km and 190 km, respectively.

The upper part of Figure 11-8 indicates that more than half of the final focus is occupied by about 100 bending magnets. These magnets generate the dispersion required for chromatic correction. Their maximum field at 1-TeV-c.m. energy is only 160 G, in order to restrict the emittance growth due to synchrotron radiation. The length of the entire system, the maximum beta functions, and the maximum dispersion (hence the bending angles) were optimized for the original design parameters, not only with regard to the effect of synchrotron radiation, but also with regard to nonlinear aberrations, magnet-vibration and field-ripple tolerances. The optimization procedure is discussed in the next section, and in [Zimmermann 1995].

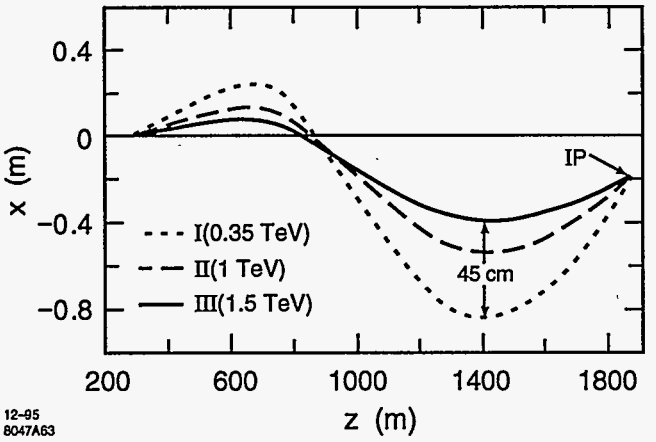


Figure 11-9. *Top view of the final-focus geometry. Magnet displacements by at most 45 cm are necessary during an upgrade from 500 (350) GeV to 1.5 TeV, while the IP-orbit angle changes by about 1.5 mr.*

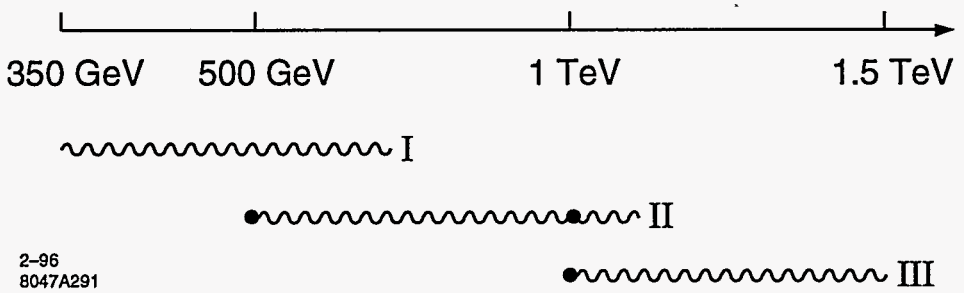


Figure 11-10. *Energy range of three different final-focus geometries. Dots show systems reported in the text.*

The bending in CCX and CCY is in opposite directions, so that the final focus exhibits an "S"-shape geometry. For a c.m. energy of 1 TeV, the accumulated absolute bending angle, $\sum_i |\theta_i|$, is 6.3 mr, the bending angle inside the CCY alone is 0.9 mr, and, due to the 'S'-geometry, the total net angle of the final focus is small: about 0.6 mr. At 350 GeV, due to shifted relative importance of aberrations and synchrotron radiation, the bending angles and the dispersion must be increased by about 50%, with an accompanying change of geometry. The maximum horizontal displacement of magnets required for the upgrade from 350 GeV to 1.5 TeV is about 45 cm (Figure 11-9). The first geometry change occurs at about 500 GeV, the second around 1.1–1.2 TeV. The energy range covered by the three final-focus geometries is illustrated in Figure 11-10. The final-focus design for 1 TeV, including the final doublet, can operate in the entire energy range from 500 GeV to 1 TeV, and possibly beyond. Between 350 GeV and 1.5 TeV, the orbit angle at the IP varies by about 1.5 mr, while the IP position itself is independent of energy. A constant IP position is desirable, since the detector needs to be strongly coupled to the ground, to preserve the coherence of magnet motion caused by ground waves.

To protect the shielding masks in the final doublet region from hard synchrotron radiation, soft bending magnets of field strength 12 Gauss are placed over a distance of 64 m in front of the final transformer. The present soft-bend configuration was devised by S. Hertzbach and implemented by R. Helm. The soft bends deflect the beam orbit at the entrance to the final doublet (quartet) by about 8 mm horizontally, so that high-energetic photons generated in upstream bends and quadrupoles will not hit the inner bore of the last two final-doublet quadrupoles [Hertzbach 1995].

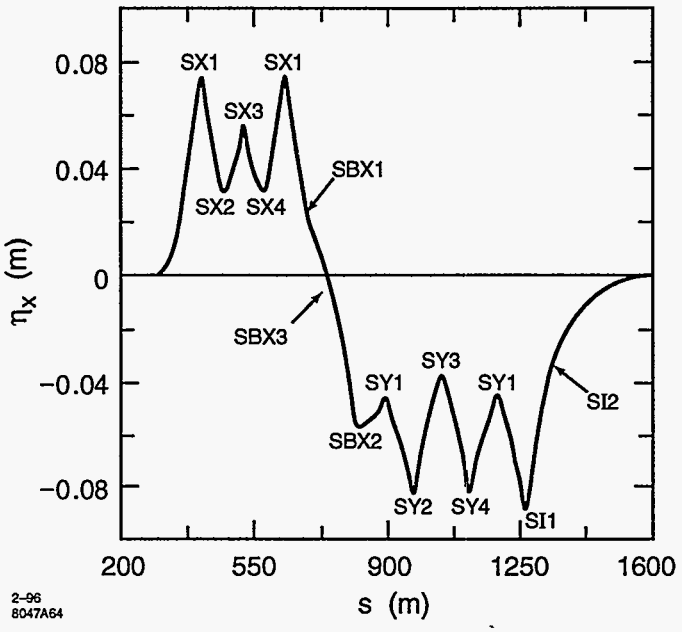


Figure 11-11. *Sextupole locations and dispersion for the 1-TeV design.*

The 44 quadrupoles between the BMS and IP are typically 0.5-m long, and, for 1-TeV-c.m. energy, their pole-tip field is 3–5 kGauss. A few magnets, with larger apertures, require pole-tip fields of about 8 kGauss. The final focus also comprises between seven and 16 sextupoles which cancel chromatic aberrations. First, there are two conventional $-I$ sextupole pairs located in the chromatic correction sections. These are used to compensate the first-order chromaticity of the system. In addition, between three and 12 weaker sextupoles are interspersed in the GAS, CCX, BX, CCY, and FT; all at positions with nonzero dispersion. The sextupole locations and the dispersion function for the 1-TeV final focus are illustrated in Figure 11-11.

The sextupole strengths are determined from tracking to optimize the momentum bandwidth of the system. This application of sextupoles for bandwidth-optimization was first proposed by Brinkmann at DESY [Brinkmann 1990]. Similar to Brinkmann's early results, the momentum bandwidth of the NLC final focus is at least doubled by means of the additional sextupoles. This beneficial effect of Brinkmann-sextupoles is explained by a reduced chromatic breakdown of the $-I$ sections between the main sextupoles and also of the FT: a Taylor-map analysis of the final-focus optics reveals a significant reduction of fifth-order chromo-geometric aberrations due to the additional sextupoles (Section 11.5.3).

Optimization

The length of the 1.5-TeV final-focus system was originally optimized with regard to nonlinear aberrations, such as third-order horizontal and vertical chromaticity, and chromo-geometric terms with generator $x'^2y'^2\delta$, and also with regard to the effect of synchrotron radiation in the bending magnets, octupole-like aberrations from long sextupoles, magnet vibration tolerances inside the CCY, and power-supply ripple. A general optimization procedure is described in [Zimmermann 1995], and is a modified version of an earlier proposal by Irwin [Irwin 1991]. Some specific side-constraints for the actual design are not included in this optimization

Parameter	1.5 TeV	
	Design	Theor. optimum
β_D^y (km)	300	80
η_D (mm)	30	24
k_D (m^{-2})	2.2	8.3
Δx (nm)	300	300
$\Delta k/k$	5×10^{-5}	8×10^{-5}
L_{tot} (m)	1800	900

Table 11-6. *Comparison of the 1.5-TeV final focus with a hypothetical design optimized for minimum length, assuming $\beta_x^* = 10\,mm$, $\beta_y^* = 125\,\mu m$, and $\delta \approx 3 \times 10^{-3}$. The length L_{tot} denotes the total distance from the start of the final focus to the IP, $\Delta k/k$ is the maximum field ripple of quadrupoles in the CCY, and Δx the tolerable orbit drift at the second Y-sextupole due to perturbations internal to the CCY.*

scheme. In particular, the optimization assumes only one operating energy and detector backgrounds are not taken into account.

The final focus has been designed such that its total length is constant in the entire energy range between 350 GeV and 1.5 TeV. The length of the system is then determined by the 1.5-TeV case, and it may be interesting to compare the design at this energy with the shortest possible, or optimum, final focus. Table 11-6 lists the theoretical optimum and the actual design values of vertical beta function and dispersion at the Y-sextupoles, sextupole strength, and length. The optimization assumes that the maximum tolerable horizontal orbit variation Δx at the second Y-sextupole caused by perturbations internal to the CCY is the same as for the actual design, *i.e.*, 300 nm.

The length of the 1.5-TeV final focus appears to be about a factor of 2 larger than the theoretical optimum. Not included in the optimization, however, are the geometry-adjustment section in front of the CCX, which holds the IP position constant at all energies, the soft-bend section, and the long final transformer which proved to be essential for reducing detector background due to synchrotron radiation. The geometry-adjustment section and final transformer in the 1.5-TeV design are about 230-m and 500-m long, respectively, and account for most of the additional length, compared with the theoretical optimum.

Apertures and Beam Stay-Clear

The collimation depth in the postlinac collimation section will be about 7 σ_x, 35 σ_y and $\Delta E/E \approx 4\%$. There are also four horizontal and vertical collimators in the final-focus system, close to the main sextupoles, with a slightly larger collimation depth. The beam stay-clear in the final-focus system has to exceed this collimation depth. A side-constraint is that there should be as few as possible variations of the vacuum-chamber dimension, since these generate geometric wakefields. The beam-pipe radius r of the present design fulfills

$$r^2 \geq \max\left[(20\sigma_{\beta,x})^2 + (\eta_x \times 0.04)^2\,, (45\sigma_{\beta,y})^2\right] \quad . \tag{11.8}$$

The contribution from the 4% energy spread is small compared with the 20-$\sigma_{\beta,x}$ betatron beam size. Figure 11-12 illustrates the variation of the 20 σ horizontal and 45 σ vertical beam envelopes, right-hand side of Eq. 11.8, along the 500-GeV (c.m.) final focus, assuming normalized emittances $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m

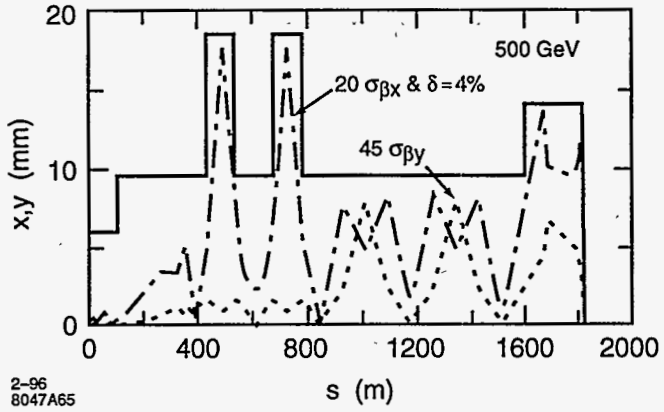


Figure 11-12. *Horizontal and vertical beam envelopes ($20\sigma_x$ and $45\sigma_y$) in the final-focus system for 500-GeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

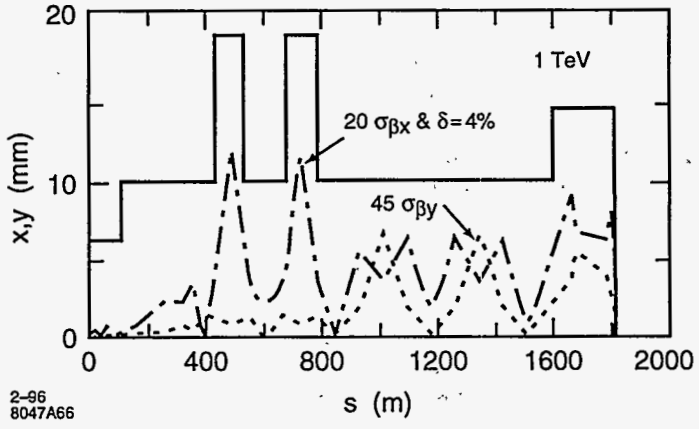


Figure 11-13. *Horizontal and vertical beam envelopes ($20\sigma_x$ and $45\sigma_y$) in the final-focus system for 1-TeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m, which are 10%–20% larger than the design value. The proposed beam-pipe radius is also depicted. In most of the system the radius is constant, equal to 9.5 mm. Magnet apertures are assumed to be 1 mm larger, *i.e.*, about 10.5 mm.

The largest beam size is encountered around the CCX sextupoles. Here the beam-pipe radius needs to be increased to 18.5 mm. Larger apertures of 19–20 mm are then required for the two SX1 sextupoles, for four quadrupoles (QE1, QX3A, QX3B, and QBX1) and for nine B2 bending magnets in the CCX. A larger beam-pipe radius, 14 mm, is also necessary in the final transformer, and in the adjacent soft bending magnets. In the figure, the beam-pipe radius was chosen as 14 mm for all soft bends.

In the 1-TeV final focus, the beam-pipe radii are still the same as for 500 GeV, and, because of the reduced beam sizes, the beam stay-clear becomes magnificent; see Figure 11-13.

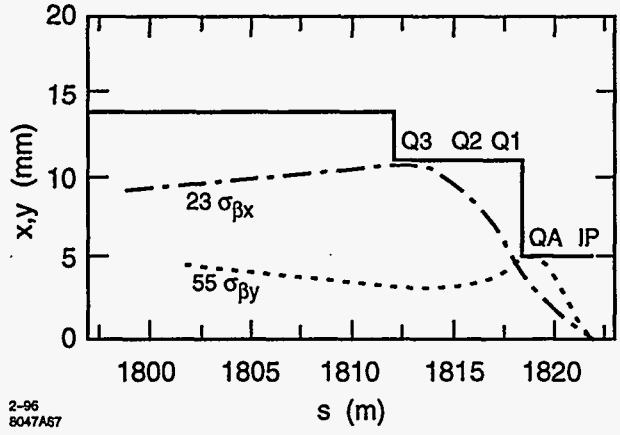


Figure 11-14. *Horizontal and vertical beam envelopes ($18\sigma_x$ and $45\sigma_y$) at the four quadrupoles of the final doublet, for 500-GeV-c.m. energy, assuming normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 8 \times 10^{-8}$ m. The beam-pipe radius is also depicted.*

An increased aperture could also be useful around the SY sextupoles, if the vertical emittance is much larger than the design value (for instance, during commissioning). In that case, the best luminosity would be achieved for constant vertical beta function (this is shown in Section 11.5.3), "Bunch Length," which implies larger IP divergence and larger aperture needs. Increased apertures in the CCY do make sense only, of course, if the aperture is not limited somewhere else, *i.e.*, in the final-doublet magnets.

The beam stay-clear at 500-GeV-c.m. energy for the 1-TeV final-doublet design is depicted in Figure 11-14. The beam-pipe radius is 11 mm inside the two conventional quadrupoles Q3 and Q2, as well as in the superconducting quadrupole Q1, and it decreases to 5 mm in the last, permanent-magnet quadrupole QA. The same doublet can operate in the entire energy range between 500-GeV and 1-TeV-c.m. energy. At 500 GeV the horizontal beam stay-clear in the final doublet is 23 σ_x; the vertical beam stay-clear is about 55 σ_y (see Figure 11-14). Again, at 1 TeV the beam stay-clear of the doublet, in terms of beam size, is much larger than at 500 GeV. For a detailed discussion of the final doublet see Section 11.6.

To reduce background and to ease the collimation, the final-focus design will likely evolve towards even larger beam stay-clear. An overall vertical stay-clear of 55 σ_y at 500-GeV-c.m. energy is readily achieved by slightly increasing the sextupole apertures. Either by modifying the final doublet or by using a different doublet at 500 GeV and at 1 TeV , the vertical aperture at 500 GeV may be further increased to 60 σ_y.

In conclusion, the apertures of the final-focus design are very large, when viewed in terms of beam size and compared with the SLC (where the stay-clear is about $15\sigma_{x,y}$). A further increase of the vertical beam stay-clear from 55 σ_y to 60 σ_y would also not be difficult.

11.5.3 Performance

Beam Size at the Interaction Point

In this subsection, we discuss contributions to the spot-size from synchrotron radiation and from higher-order optical aberrations. The spot size, or luminosity, is further diluted by residual uncorrected low-order aberrations, which are described later in this section, by crab-crossing errors and by timing offsets between

	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
	σ_x (nm)	σ_y (nm)	σ_x (nm)	σ_y (nm)	σ_x (nm)	σ_y (nm)
Linear	252	4.18	200	3.35	200	3.35
$\delta_{\rm rms} = 0.3\%$	264	4.32	212	3.40	208	3.55
+ s. rad. in bends	264	4.32	232	3.47	210	3.56
+ s. rad. in quads	264	4.33	232	3.59	211	3.71

Table 11-7. *RMS IP spot sizes for 500 GeV and 1 TeV, as obtained by tracking 10,000 particles with DIMAD.*

	500 GeV		1.0 TeV		1.0 TeV (1.5 TeV design)	
	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$	$\Delta\sigma_x/\sigma_{x0}$	$\Delta\sigma_y/\sigma_{y0}$
Chromo-geometric	3.5%	3.3%	**6.0%**	1.5%	4.0%	**6.0%**
Rad. in bends	0.0%	0.0%	**10.0%**	2.1%	1.0%	0.3%
Rad. in quads (tracking)	0.0%	0.2%	0.0%	**3.5%**	0.5%	**5.1%**
Total	**4%**	**3%**	**16%**	**7%**	**6%**	**11%**

Table 11-8. *Relative spot-size increases due to various effects, according to Table 11-7.*

the two beams: A crab-crossing error arises when the phases of the two crab cavities vary with respect to each other. In that case the two beams do not collide head-on. The crab-cavity tolerances, which are discussed in Section 11.7.2, allow for a 2% luminosity loss from imperfect crab crossing. If there is a timing (or phase) error between the two bunch compressors, the collision occurs longitudinally offset from the IP waist, and the spot size at the collision point is increased by the effective waist shift. The proposed tolerance budget also assigns a 2% luminosity loss to this timing error (tolerances on bunch-compressor rf phases are discussed in Chapter 5).

The IP-spot-size increase due to higher-order chromatic aberrations and due to synchrotron radiation in bending magnets and quadrupoles is shown in Table 11-7, for the IP beam parameters of Table 11-1. Each entry in Table 11-7 was obtained by tracking 10,000 particles using the program DIMAD [Servranckx 1990]. Spot sizes for the 1.5-TeV final focus (Version III in Figure 11-10) operated at 1 TeV are also shown, in the far right column of the table. They are slightly better than those obtained with Version II at 1 TeV (second column from the right).

Table 11-8 compiles the relative spot size increases due to different sources as deduced from the above tracking results. At 500 GeV, the chromo-geometric blow-up is about 3.5% both horizontally and vertically. There is hardly any spot-size increase caused by synchrotron radiation.

At 1 TeV, the spot-size increase is considerably larger: 16% horizontally and 7% vertically for the 1-TeV design, and 6%/11% for the 1.5-TeV version operated at 1 TeV (Table 11-8). For the 1-TeV final focus, the largest contribution to the vertical spot-size arises from synchrotron radiation in the final doublet (Oide effect), which causes an rms spot-size increase by about 3.5%. In general, the Oide effect decreases with increasing length of QFT2/QFT3. In the proposed design, the total length of these two quadrupoles is about 4 m, and further improvement for greater length is negligible. (The Oide effect is discussed later in this section and in 11.6.4: "The Oide Effect Including Horizontal Motion".)

The horizontal blow-up at 1 TeV is primarily caused by radiation in the bending magnets (10%) and by chromo-geometric aberrations (6%), assuming a Gaussian momentum distribution of rms value 0.3%. The blow-up due to synchrotron radiation is proportional to the R_{16} matrix-element from the bending magnets to the IP. This blow-up could be reduced by either lowering the dispersion or by weakening the bending magnets. On the other hand, the chromo-geometric aberrations are enhanced by a reduced dispersion (and larger sextupole strength). Thus the present value of dispersion, at which the effects of synchrotron radiation and aberrations are roughly comparable, is about the optimum. One possibility to further reduce the residual horizontal spot-size increase due to radiation and aberrations would be to increase the length of the final-focus system.

The impact of dispersion on spot size and bandwidth can be understood by noting that most of the aberrations arise from the chromatic breakdown of the $-I$ between the main sextupoles. The aberrations affecting the vertical spot size are then described by a Hamiltonian of the form

$$H \propto K_s^2 \eta_s^2 \left(y'^2 \delta^3 + 2\delta^2 x' y'^2 \frac{\beta_{x,s}^{\frac{1}{2}}}{\eta_s} + x'^2 y'^2 \delta \frac{\beta_{x,s}}{\eta_s^2} + \ldots \right) \quad , \tag{11.9}$$

where K_s is the sextupole strength, $\beta_{x,s}$ the beta function, and η_s the dispersion at the sextupole. The coordinates x', y' denote the normalized slopes at the IP (in units of $\sqrt{\mathrm{m}}$). The product $K_s \eta_s$ is a constant, determined by the doublet chromaticity. The Hamiltonian, Eq. 11.9, shows that most of the aberrations are reduced for increased values of η_s except for the first term which represents the third-order vertical chromaticity. Similarly, most of the important aberrations which influence the horizontal spot size are also smaller for increased dispersion. The third-order vertical chromaticity, which is independent of the dispersion, is canceled by the Brinkmann-sextupoles (Section 11.5.3, "Aberrations").

In summary, synchrotron radiation and high-order chromo-geometric aberrations lead to a total increase of the vertical rms spot size by about 3% and 7% at 500 GeV and 1 TeV, respectively. Since a large part of the blow-up at 1 TeV is due to the Oide effect, the actual luminosity loss will be smaller than quoted, because only few photons are radiated. This is elaborated in Section 11.6. Finally, the increase of the horizontal spot size due to synchrotron radiation could be compensated with a further reduction of β_x^*, but we have not made this assumption.

Momentum Bandwidth

A large momentum bandwidth of the final focus is very desirable considering beam dynamics in the linac, although a larger energy spread necessitates tighter tolerances in the final focus. The bandwidth of the final focus has been increased by introducing additional sextupoles, as suggested by Brinkmann. Examples of the achievable bandwidth are presented in Figure 11-15 and 11-16 for 500 GeV and 1 TeV, respectively, where in both cases three additional Brinkmann sextupoles have been used: one in each of CCX, CCY, and FT. A minimum of three sextupoles appears to be needed for achieving a decent bandwidth, but the further improvement in going from three to 12 Brinkmann-sextupoles is rather small. Therefore, the presently adopted design employs only three Brinkmann sextupoles.

Figure 11-15 shows the vertical and horizontal beam sizes at 500 GeV as a function of the momentum-offset Δ of a mono-energetic beam. The momentum bandwidth for a 10% blow-up of either spot size is $\pm 0.61\%$. As illustrated by Figure 11-16, at 1 TeV the bandwidth for a 10% spot-size increase is about the same. In both figures, the spot sizes are given in units of the values for zero momentum offset. If Brinkmann-sextupoles are not used, the momentum bandwidth is less than $\pm 0.35\%$.

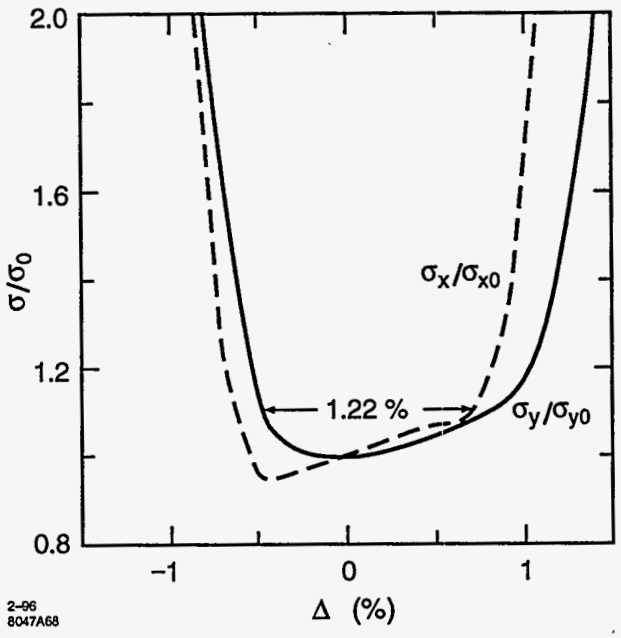


Figure 11-15. *Relative increase of spot sizes as a function of momentum-offset Δ for a monoenergetic beam at 500 GeV, using three Brinkmann-sextupoles. Effects of synchrotron radiation are not included.*

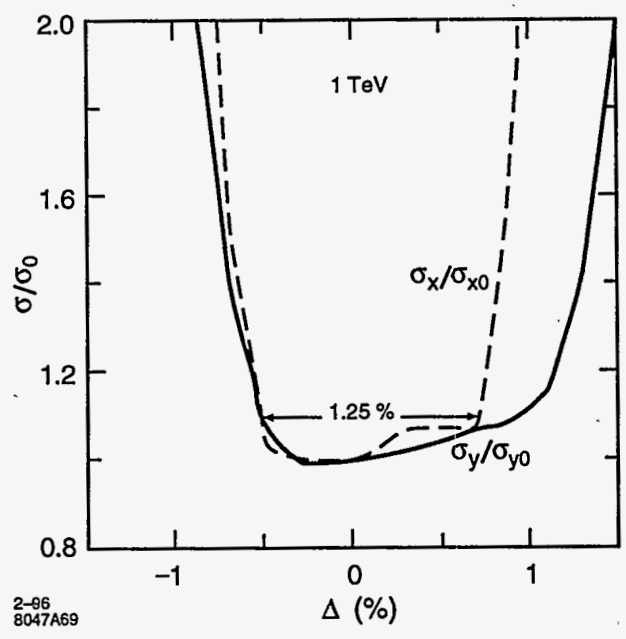


Figure 11-16. *Relative increase of spot sizes as a function of momentum-offset Δ for a monoenergetic beam at 1 TeV, using three Brinkmann-sextupoles. Effects of synchrotron radiation are not included.*

Aberrations

In this subsection, we first discuss the importance of low-order aberrations. Although they can be tuned, these aberrations turn out to be significant for proper estimating the achievable luminosity. Towards the end of this subsection, we also describe the dominant higher-order aberrations.

The perhaps most important, yet often overlooked source of IP spot-size dilution are residual low-order aberrations. These are aberrations, such as waist shift, skew coupling, etc., for which tuning schemes exist, but which cannot be fully compensated, due to the finite accuracy of the beam-beam deflection scans, from which the IP beam size is inferred during tuning. In general, we allocate a 1–2% spot-size increase to each first-, second- and third-order aberration, including the spot-size increase occurring between two tuning scans. This appears to be a fairly conservative estimate of tuning accuracy and stability. For comparison, the actual tuning accuracy in the SLC is about 0.5% per aberration.

The aberration tuning in the NLC final focus can further be improved in a number of ways: A fast and precise luminosity monitor, *e.g.,* a detector for low-angle Bhabha scattering, will provide a more accurate luminosity measurement than the beam-beam deflection scans and will, therefore, allow a better fine-tuning of aberrations. For all energy-dependent aberrations, the tuning may also be bettered by temporarily increasing the energy spread. The horizontal blow-up will be compensated by reducing the horizontal IP beta function. In addition, if first- and second-order dispersion are continually monitored and corrected, using measured correlations between the beam position at the IP or at the pre-image point and the centroid energy, these two aberrations will become insignificant.

The effect of the low-order aberrations is summarized in Table 11-9 which lists all aberrations that will be tuned, including tuning procedure, tuning frequency, luminosity loss attributed to each aberration, and associated maintenance systems. The table shows that the total increase of the vertical spot size due to residual uncorrected tunable aberrations is expected to be about 13%. Thus, imperfectly-tuned vertical aberrations amount to about half the total vertical blow-up originating in the beam-delivery system; most of the rest being due to synchrotron radiation and to uncorrectable higher-order aberrations. Note that the assumed value for the average dilution per aberration is somewhat ambiguous, since it depends on tuning accuracy, tuning frequency and BPM stability, and can be improved in numerous ways. The assumptions in Table 11-9 differ from those made earlier in Table 11-5. The latter table was more optimistic and predicted a vertical blow-up due to tuned aberrations of only about 6%.

Optical aberrations of order 5 or higher (in the Hamiltonian) exist by design, and at present no tuning scheme is contemplated for any of these. Higher-order aberrations can be calculated from a Taylor map M representing the final-focus system to arbitrary order in (x, x', y, y', δ). Such a map is extracted from a standard MAD input file by methods of differential algebra [Berz 1989] (also called truncated power series algebra), using the program DESPOT [Forest] [3]. The Taylor map can be rewritten as a Dragt-Finn factorization, from which an approximate nonlinear Hamiltonian is determined. Formally we have

$$M = R\exp(-H_{\rm non}) \; + \; \text{higher-order terms} \tag{11.10}$$

where the term R denotes the linear transfer matrix, and $H_{\rm non}$ represents the nonlinear Hamiltonian describing the aberrations. The spot-size increase caused by these aberrations is given by

$$\Delta\sigma_{x,y}^2 \approx \left\langle \left(\frac{\partial H_{\rm non}}{\partial p_{x,y}} \right)^2 \right\rangle \quad . \tag{11.11}$$

[3] In the current installation, the MAD input first has to be converted into a DESPOT input with the help of the code TRACY.

Time	Name	Hamil-tonian	Budget[c] (%)[a]	Tuning Knob	Diagnostic	Maint. System
1/5s	Steering	y'	.5+.5	Corrector at FD	IP kick	FD anchor + seism. + steering
		x'	.5+.5 [b]			none
1 hr	Dispersion	$\delta y'$	N/A[d]	Weak SD Sexts. asym. hor. move	IP spot size scans	FF orbit control
		$\delta x'$	N/A [b,d]	Weak SF Sexts. asym. hor. move		
	Waist	y'^2	.5+.5	Weak SD Sexts. symm. hor. move		
		x'^2	1+1 [b]	Weak SF Sexts. symm. hor. move		
	Skew1	$x'y'$	1+1	Weak SD Sexts. asym. vert. move		
1 wk	Sextupole	$x'y'^2$ x'^3	2 2 [b]	Sp. sext. in FT	IP spot size scans	Sext. strength & LT stability
	Skew sextupole	y'^3 x'^2y'	2 2	Sp. skew sexts. in FT		Sext roll, FD mult. & LT stability
	Skew2	xy'	2	Sp. skew quad		LT stability & solenoid strength
	Chromaticity	$\delta y'^2$ $\delta x'^2$	1 [c] 1 [b,c]	SD sext strength SF sext strength		& sext. strength & roll
	Chromatic Skew	$\delta x'y'$	1 [c]	CCY vert. disp.		Occasional FFS alignment studies
	2nd order disp.	$\delta^2 y'$	N/A [c,d]	QT6 and QT5 move		& LT stability
1 mo	Within range of aberration tuning knobs	All	N/A	Element movers	Beam-based alignment	FFS align-ment studies & XLT stability
1 yr	Within range of bm based align. (capt. tol.)	N/A	N/A	Coarse element movers	Survey	XXLT stability

[a] .5+.5 means the budget is 1/2% for the final doublet and 1/2% for the remainder of the system; 1+1 means the budget is 1% for the final doublet and 1% for the remainder of the system; system being either the final focus (for steering) or the whole beam line from linac to IP (for all other aberrations).
[b] The pure x-aberrations can be compensated by a slight reduction of β_x^*.
[c] Chromatic aberrations can be enhanced by increasing the energy spread. We assume that these aberrations are tuned out with a doubled energy spread. The chromatic skew can be eliminated by canceling the vertical dispersion. The second-order dispersion at the IP could also be detected directly if there were a beam size monitor available at the IP.
[d] Dispersion and second-order dispersion can be continually corrected, using the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy of both beams.
[e] Perhaps, because they are stable, the aberrations in the one-week timescale could, under steady beam conditions, be better compensated using a luminosity monitor. We have not made this assumption.

Table 11-9. *Low-order aberration tuning and maintenance. This table allocates a 1% spot-size increase to each δ-dependent aberration, except for dispersion which is assumed to be perfectly compensated, and it assigns a 2% spot-size increase to most other aberrations. The total increase of the vertical spot size is 13%. Note that in Table 11-5 different assumptions about tuning accuracy, frequency and stability were made (0.5% increase per aberration including dispersion), which resulted in about half the vertical blow-up.*

$(\Delta\sigma_y^2)^{\frac{1}{2}}/\sigma_{y0}$	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
Generator	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.
$x'y'^2\delta^2$	0.31	0.21	0.27	0.29	0.52	0.63
$y'^2\delta^3$	0.52	0.02	0.58	0.03	0.78	0.12
$(\Delta\sigma_x^2)^{\frac{1}{2}}/\sigma_{x0}$	500 GeV		1.0 TeV		1.0 TeV (1.5-TeV design)	
generator	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.	no B.-sext's.	3 B.-sext's.
$x'^4\delta$	0.12	0.08	0.06	0.22	0.25	0.30
$x'^2\delta^3$	0.25	0.27	0.33	0.47	0.28	0.33

Table 11-10. *Effect of fifth-order nonlinear aberrations in the final-focus system, with and without Brinkmann-sextupoles for the parameters listed in Table 11-1 and $\delta_{\rm rms}$ =0.3%. The contribution to the spot size quoted has to be added in quadrature.*

The angular brackets indicate an average over the (linear) beam distribution at the IP, usually assumed as Gaussian.

An example of such an analysis is given in Table 11-10, which shows contributions to the spot size of some important fifth-order aberrations, to be added in quadrature. The largest aberration is the third-order vertical chromaticity (generator $y'^2\delta^3$), which in the absence of Brinkmann-sextupoles would cause a spot-size increase of 16–25%. In the optimized design, the effect of this aberration becomes negligible (see right column). There is still a considerable amount of third-order horizontal chromaticity, which seems to give rise to a spot-size increase of 5–10%. However, it is somewhat misleading to look at this term in isolation: to produce the desired optimum bandwidth, the third-order horizontal chromaticity has (empirically) been balanced against a nonzero first-order chromaticity which is of similar size and opposite sign. For the optimization of higher-order terms, evidently, a compromise had to be found between a large momentum bandwidth and a small chromo-geometric blow-up for a beam of certain finite momentum spread.

In summary, uncorrected lower-order aberrations cause a 6%–13% increase of the vertical spot size. The horizontal spot size may be recovered by reducing β_x^*. The residual fifth-order chromo-geometric aberrations give rise to a vertical blow-up of 2%–10%.

Synchrotron Radiation

Besides the effect of lower- and higher-order aberrations, further increases of the spot size arise from synchrotron radiation in final doublet and bending magnets. These are the subject of this section. The Oide effect is also treated, much more comprehensively, in Section 11.6.4.

Synchrotron radiation in the final-doublet quadrupoles causes different particles to be focused at a different distance from the last quadrupole, and, thereby, gives rise to an ultimate limit on the achievable spot size. This has first been pointed out by K. Oide [Oide 1988], after whom it is called "Oide limit". Generalizing Oide's original formula for the spot-size increase [Oide 1988] to two planes (horizontal and vertical) and both quadrupoles of the final doublet, the increase of the vertical spot size due to synchrotron radiation in the last two quadrupoles is given by

$$\Delta\sigma_y^2 = \frac{55 r_e \lambda_e \gamma^5}{48\pi\sqrt{3}} \left[\int_0^{l_1+l_2} \left[\int_0^s g_y(s_1)^2 K(s_1) ds_1 \right]^2 \right].$$

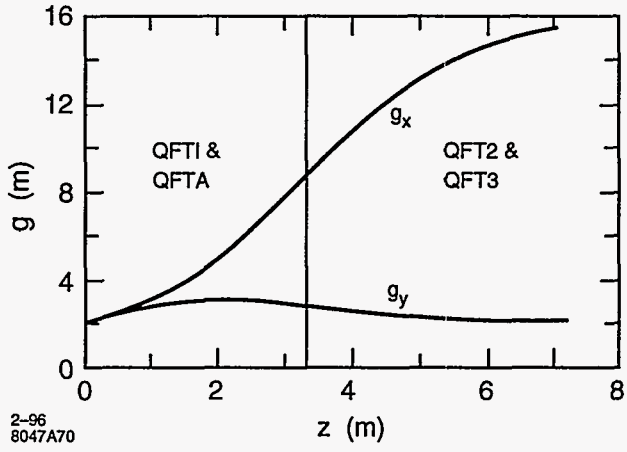


Figure 11-17. *Horizontal and vertical Green's functions as a function of longitudinal position s measured backward along the NLC final doublet. the zero point $s = 0$, refers to the exit face of the last quadrupole.*

$$\cdot|K(s)|^3 < \left[|g_x(s)^2 x_0'^{*2} + g_y(s)^2 y_0'^{*}2|\right]^{\frac{3}{2}} y_0'^{*2} > ds\Big] \quad , \tag{11.12}$$

where the angle-brackets denote an average over the bunch distribution, the variable s is the longitudinal position measured backward from the exit face of the last quadrupole, and $K(s)$ is the quadrupole strength at position s. The Green's functions g_x and g_y are depicted in Figure 11-17 for an approximation to the NLC final doublet.

The figure shows that most of the spot-size increase occurs in the second pair of quadrupoles viewed from the IP (QFT2 and QFT3), and is due to the focusing in the horizontal plane, rather than to the final vertical focusing. In general, increasing the length of QFT2 and QFT3, while keeping the integrated quadrupole field $K_2 \cdot l_2$ constant, reduces the effect of synchrotron radiation in the final doublet. The length chosen is close to the optimum, in the sense that further improvement for still greater length is negligible. With the present final-doublet design, the increase of the rms spot size due to the Oide effect is less than 0.1% at 500 GeV and about 3.5% at 1 TeV; compare Table 11-8. A modification of the final-doublet design may be required to reduce the blow-up at 1 TeV. The reader is referred to Section 11.6 for a more detailed discussion of the Oide effect.

Synchrotron radiation in the bending magnets affects the beam size at the IP in two different ways. First, radiation in dispersive regions increases the horizontal spot size because of a nonvanishing R_{16}-matrix element between dispersive bend regions and the IP. Synchrotron radiation in the bending magnets causes a horizontal spot-size increase by about 0% and 1–10% at 500 GeV and 1 TeV, respectively (Table 11-8). The blow-up at 1 TeV could be reduced by lengthening the system. Second, for particles losing energy in the final and central bending sections of the CCY, the vertical chromaticity of the final doublet is not or only partly compensated, and the particles are chromatically defocused, resulting in an increase of the vertical spot size. This second mechanism leads to a vertical blow-up by about 0.3–2% at 1 TeV (Table 11-8).

The spot-size increases due to synchrotron radiation in the final doublet and bending magnets have been listed in Table 11-8. Some related parameters such as the average energy loss ΔE_{rad}, the average number of photons emitted per electron N_γ, and the critical energy E_c in the bending magnets are given in Table 11-11.

We conclude that synchrotron radiation in bending magnets and quadrupoles is not important at 500 GeV and that at 1 TeV it reduces the luminosity by about 15%, assuming the same final-focus geometry as at

	Energy	500 GeV	1 TeV	1 TeV[a]
Number of photons	N_γ	33	66	34
Critical energy	E_c (MeV) in bends	0.3	2.7	1.6
Critical energy	E_c (KeV) in soft bends	25	200	173
Average energy loss	$\Delta E_{\rm rad}$ (MeV)	3.3	52	17
Induced energy spread	$\Delta\delta_{\rm rms}$	0.6×10^{-5}	3.2×10^{-5}	1.2×10^{-5}
Emittance growth	$\Delta(\gamma\epsilon_x)$ (mr)	8×10^{-9}	4.1×10^{-7}	4.1×10^{-8}
Rel. emittance growth	$\Delta(\epsilon_x)/(\epsilon_{x\,0})$ (%)	0.2	9.6	1.0
Hor. spot-size increase	$\Delta\sigma_x/\sigma_{x0}$ (%) (tracking)	0.0	10.0	1.0
Vert. spot-size increase	$\Delta\sigma_y/\sigma_{y0}$ (%) (tracking)	0	2	0.3

[a]This column refers to the 1.5-TeV final focus operated at 1 TeV.

Table 11-11. *Characteristics of synchrotron radiation in the bending magnets.*

500 GeV (Version II in Figure 11-10). The luminosity decrease due to synchrotron radiation is less than half that, *i.e.,* 7% , if the 1.5-TeV final focus (Version III) is operated at 1 TeV.

Long-Sextupole Effect

A long sextupole magnet gives rise to an octupole-like fourth-order aberration (see, *e.g.,* [Irwin 1992]). For a certain strength, the maximum sextupole length causing a 2% spot-size increase is given by

$$l_{sext} \le \frac{1.2}{\sqrt{3}\beta_x^s\beta_y^s\epsilon_{x,y} + \sqrt{15}\beta_{y,x}^{s\ 2}\epsilon_{y,x}}\frac{1}{k_{sext}^2} \tag{11.13}$$

where

$$k_{sext} \equiv \frac{2l_{sext}B_T}{(B\rho)a^2} \tag{11.14}$$

denotes the integrated sextupole strength in units of m^{-2}, and the beta functions are those at the sextupole. Applying this formula to the 1-TeV final focus for normalized emittances of $\gamma\epsilon_x \approx 5\times10^{-6}\,\mathrm{m}$ and $\gamma\epsilon_y \approx 8\times10^{-8}\,\mathrm{m}$, we find the maximum tolerable sextupole lengths

$$l_{SX1} \le 27\,\mathrm{m} \quad \text{and} \quad l_{SY1} \le 23\,\mathrm{m}\ , \tag{11.15}$$

For the same optics at 500 GeVand normalized emittances of $\gamma\epsilon_x \approx 5\times10^{-6}\,\mathrm{m}$ and $\gamma\epsilon_y \approx 10\times10^{-8}\,\mathrm{m}$, the limit on the sextupole length is

$$l_{SX1} \le 9\,\mathrm{m} \quad \text{and} \quad l_{SY1} \le 12\,\mathrm{m}\ . \tag{11.16}$$

Thus, the fourth-order aberrations induced by the finite sextupole length are insignificant. Note that the 500-GeV limit is tighter, because both the emittances and the horizontal beta function are larger than at 1 TeV. By contrast, due to increased dispersion and lowered sextupole strength at 350 GeV, the maximum sextupole length for the 350-GeV optics will be larger than at 500 GeV.

At 500 GeV, the sextupole lengths may be chosen as 0.4 m, which, for the assumed magnet apertures of 20 mm and 10.5 mm, translates into a pole-tip field of 3.3 kGauss and 1.9 kGauss for SX1 and SY1, respectively. The upgrade to 1 TeV could be performed by adding more sextupoles, to increase the effective length of the two SX1 and SY1 magnets to 0.8 m. Alternatively, longer sextupoles could be installed from the beginning.

We conclude that the present design is far from the long-sextupole limit.

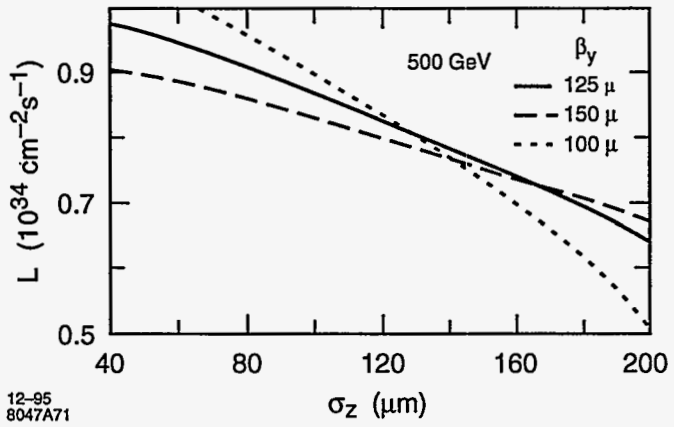


Figure 11-18. *Luminosity at 500-GeV-c.m. energy as a function of the bunch length, for three different values of β_y^*, 90 bunches with 0.65×10^{10} particles per bunch, and a repetition rate of 180 Hz; other parameters as listed in Table 11-1.*

Bunch Length

It is interesting to study the dependence of luminosity on bunch length and vertical beta function. A larger bunch length would relax requirements on the bunch compressor or on the main linac, while an increase of the vertical beta function would reduce the spot-size blow-up due to aberrations and due to synchrotron radiation.

Figure 11-18 shows the luminosity as a function of bunch length, for three different values of β_y^* and a c.m. energy of 500 GeV. The luminosity is estimated from a scaling law for luminosity enhancement, derived by P. Chen [Chen 1993], assuming linear IP spot sizes. The figure illustrates that an increase of the bunch length from 125 μm to 140 μm would cause a luminosity loss by about 3%. A similar luminosity loss occurs if the vertical IP beta function is raised from 125 μm to 150 μm.

The conclusion here is that both bunch length and IP beta function can be changed over a wide parameter range with only marginal loss of luminosity.

11.5.4 Tuning

Control over IP horizontal and vertical dispersion, waist position, demagnification, chromaticity, and linear coupling, is provided by orthogonal magnet- or mover-knobs similar to those suggested by Irwin for the SLC [Walker 1993, Zimmermann 1995b] or the FFTB. Laser wires will measure the beam size at several locations in the final focus. Valuable experience will be gained from a first laser wire installed at the SLC [Ross 1994]. Decoupling, match of the incoming dispersion and emittance-measurements are performed upstream of the CCX (*e.g.*, in big bend, SCS and DS). The BMS-quadrupoles are used to correct a mismatch in phase space of the incoming beam and to adjust the demagnification at a pre-image point of the IP in the BX (the tuning on the pre-image point is described in Section 11.4), or at the IP itself.

Symmetric or asymmetric displacements of the main-sextupole pairs correct waist shifts, coupling and dispersion at the IP. A mover-based tuning is attractive, since a precise adjustment and control of sextupole positions is required regardless. A summary of main-sextupole mover requirements, their effect on the IP

IP Effect	Magnitude	Mover Symmetry	Mover Plane	Δ/μm at SX1	Δ/μm at SY1
Skew	$r_{32} = 0.20$	symmetric	y	0.85	0
x-waist	1.6 mm	symmetric	x	1.66	0
y-waist	25 μm	symmetric	x	0	0.28
x-dispersion	19 μm	asymmetric	x	4.30	0
y-dispersion	1.1 μm	asymmetric	y	1.07	0

Table 11-12. *IP tuning with CCX and CCY main-sextupole movers. The minimum displacement requirements, Δ, in the right two columns and the magnitude, in the second column, are for a 2% IP single spot size increase ($E_{\rm beam} = 250$ GeV, $\beta_x^* = 8$ mm, $\beta_y^* = 125$ μm, $\gamma\epsilon_x = 5\times10^{-6}$ m-rad, $\gamma\epsilon_y = 5\times10^{-8}$ m-rad, $\sigma_\delta = 0.3\%$).*

beam and minimum step-size specifications is given in Table 11-12. In all cases the SX1 or SY1 sextupole pairs are moved either symmetrically or asymmetrically.

The minimum mover step size should be ~0.9 μm and ~0.3 μm for the SX1 and SY1 sextupoles, respectively. The accuracy of the corrections is determined by mover precision, and can be improved by adding movable weak sextupoles adjacent to the main ones. The minimum mover step size requirements can be increased a factor of ten by using a main to weak sextupole strength ratio of 9:1. This also allows the main sextupoles to remain aligned so that chromaticity adjustments do not couple to IP dispersion, skew or waist effects. The main-sextupole mover range should be ±300–500 μm (±3 mm for the weak sextupoles).

Alternatively, instead of movable weak sextupoles, one could correct waist, dispersion and skew coupling with three pairs of normal and skew quadrupoles, as are being used in the SLC. An advantage is that field changes of such tuning quadrupoles could be implemented much faster than the equivalent sextupole moves. The maximum pole-tip fields required to correct a 100% spot-size increase are of the order 10–30 G at a 2-cm radius. A magnet-time constant of 50–100 ms would be sufficient to limit the luminosity loss due to tuning: assuming that seven aberrations per beam are tuned once per hour, that each aberration tuning consists of seven beam-beam deflection scans at different quadrupole settings, and that a settle time of four time constants is allocated to each quadrupole-field change, the total luminosity loss due to time spent on aberration tuning is estimated to be about 1%.

A correction limit is set by additional aberrations generated by using sextupole movers or tuning quadrupoles in the CCX or CCY to correct optical errors induced by the final quadrupole. For example, a 100-μr rolled final quadrupole (QFT1) generates skew which is correctable by moving the SX1 main-sextupole pair vertically by 60 μm. However, the vertical beam size remains 4% larger than nominal due to the non-local correction. For this reason it will be useful to include a small air-core skew quadrupole near the final doublet. A skew quadrupole of 10-cm length and 50-mm pole radius will require ~4-Gauss pole-tip field step size and ~500-Gauss range at 250 GeV/beam to correct up to 250-μr QFT1 rolls. Similarly, some vernier on the final quadrupole gradients of ~2% will be useful for correcting large doublet induced waist errors.

Some preliminary tuning simulations have been performed on the final focus system simply to test the tuning concept. The incoming beam at the start of the BMS was unrealistically distorted ($\beta_x = 2\beta_{x0}$, $\Delta\alpha_y = 1$, $\Delta\eta_x = 10$ mm, $\Delta\eta_y = 1$ mm) and a final quadrupole gradient error of 0.3% was introduced. This degrades the luminosity by three orders of magnitude. Assuming an IP beam size and divergence measurement with precision of 5%, and using the BMS quadrupole strengths and main-sextupole movers, the final beam sizes achieved were within 10% of nominal. More realistic simulations will be performed in the future which include magnet misalignments, rolls and gradient errors. The beam-based alignment of quadrupoles and sextupoles

will be similar to that performed in the SLC [Emma 1992] or FFTB [Tenenbaum 1995]. To monitor and maintain the alignment and orbit, conventional strip-line BPMs are used as well as high-resolution radio-frequency beam-position monitors (rf BPMs) [Hartman priv] which could be integrated into the structure of about 10 quadrupoles in the CCY and the FT. Alternatively each quadrupole and sextupole could be equipped with these rf BPMs, but this is not strictly necessary. The rf BPMs can detect small dispersion-generating drifts of magnet-positions and orbits over a minute timescale (Section 11.5.6, "CCY and FT Orbit Feedback.")

11.5.5 Tolerances

Tolerance requirements are very tight for most parts of the NLC design. This is especially true for the final-focus system, in which the beta functions are largest and which has to generate and collide beams of minuscule transverse sizes. It is, therefore, necessary to evaluate all tolerances and to demonstrate that they can be met and how .

Variations of the incoming beam parameters, magnet motion, and changes of the magnetic fields will all affect the luminosity. Each change, drift, or vibration needs to be smaller than a certain tolerance value, in order that the design luminosity can be achieved. The timescale over which a tolerance has to be met depends on the type of aberrations generated and also on the tuning frequency or on the damping time of an automated feedback loop.

In accordance with Table 11-9, we distinguish between the following stability categories and associated timescales:

vibration/ripple:	$t < 1/5$ s
stability/drift:	$t < 1$ hr
long-term stability:	$t < 1$ week
very long term stability:	$t < 1$ month
extremely long term stability:	$t < 1$ year

Each timescale corresponds to a tuning frequency. All tolerances studied in this section are specified relative to these timescales. Specifically, the tolerances and sensitivities discussed in this section pertain to vibration and ripple (steering tolerances), to stability (skew, waist, and dispersion effects), and to long-term stability (second-order dispersion, chromaticity, geometric aberrations).

Unless noted otherwise, all tolerances presented in this section will refer to a c.m. energy of 1 TeV , normalized emittances of $\gamma\epsilon_x \approx 5 \times 10^{-6}$ m and $\gamma\epsilon_y \approx 10 \times 10^{-8}$ m, and to IP spot sizes $\sigma_y \approx 3.6$ nm, $\sigma_x \approx 225$ nm. Both emittances and spot sizes are about 10% larger than the linear design values quoted in Table 11-1.

In the following, we first determine how sensitive the IP spot size is to incoming orbit variations and to increased emittances. We then look at how displacing a single magnet affects the IP beam position and beam size. Next, the field stability tolerances are discussed for bending magnets, quadrupoles and sextupoles, in Section 11.5.5, "Power Supply Tolerances." In Section 11.5.5, "Scaling of Tolerances," we briefly describe how the tolerances scale with beta function, emittance and energy. In Section 11.5.5, "Tolerance Budget and Luminosity Loss," the overall tolerance budget and the resulting luminosity loss are described, including the different contributions to various aberrations. If the magnet motion is due to ground waves, the individual displacements are not uncorrelated, and the tolerances for incoherent magnet motion do not apply. Instead the wavelength-dependent response of the final-focus system has to be calculated. The luminosity loss

due to ground motion is computed in Section 11.5.5. "Ground Motion". Ground motion also plays a role in requirements on repetition rate of both IP spot-size tuning (waist correction, etc.) and beam-based alignment. These questions will be addressed in Section 11.5.5. "Capture Tolerances and Tuning Frequency". Collimator tolerances are discussed in Section 11.5.5, "Collimator Tolerances". Next, thermal stability and the required vacuum pressure are discussed in Sections 11.5.5, "Temperature", and 11.5.5, "Vacuum". Finally, some beam-loss mechanisms other than gas-scattering are described.

We will distinguish between sensitivities and tolerances. By sensitivity we mean the amount of displacement or field change of a single magnet which causes either a 0.1 σ_x (0.2 σ_y) steering error or a 1% spot-size increase. While the sensitivity can be calculated for every magnet, tolerances are assigned. The tolerances are chosen so as to restrict the total luminosity loss from each aberration, due to all magnets, to 1–2%. Thus tolerances translate directly into engineering specifications on power supplies or magnet supports, while the sensitivity has no immediate practical interpretation. The two notions are, however, closely related, since a tolerance is roughly equal to the sensitivity divided by the square root of the number of magnets.

Incoming Jitter Tolerance

Figure 11-19 shows the relative increase of the horizontal and vertical IP spot-size caused by a change of the incoming horizontal and vertical beam orbit. The orbit change is quoted in units of the design beam size (or divergence) at the entrance of the final focus, and the resulting blow-up represents an average over both position and slope changes of either sign.

The figure demonstrates that a horizontal or vertical orbit change by 0.5 σ would lead to an increase of the vertical or horizontal IP spot size by less than 1%. This sensitivity to orbit changes appears tolerable.

Sensitivity to Emittance Increase

Figure 11-20 illustrates the relative loss in luminosity, disruption and hourglass effect not included, at 500 GeV as a function of the vertical emittance, for either constant β_y^* or constant θ_y^*. It is clearly advantageous to keep the beta function constant, and the importance of the chromo-geometric aberrations, relative to the linear spot size, does not appear to increase for the emittances considered. The reason is that the Lie generators of the dominant aberrations, $x'y'^2\delta^2$ and $y'^2\delta^3$ (compare Table 11-10), are proportional to y'^2 and thus increase in the same way as the linear spot size. If the vertical emittance is three times larger than the design, the luminosity is reduced by about 40%. The bandwidth for twice the vertical emittance is still the same as the design value.

A similar result is found if both the vertical and the horizontal emittances are increased together. When both emittances are doubled, to $\gamma\epsilon_x = 10^{-5}$ m and $\gamma\epsilon_y = 17.5 \times 10^{-8}$ m, the horizontal and vertical spot sizes for a 0.3% energy spread at 500-GeV-c.m. energy are 416 nm and 7.16 nm, respectively, if the beta functions are held constant (8 mm and 125 μm). This corresponds to a luminosity loss by about a factor of 2, so that again contributions from the nonlinear aberrations do not significantly increase. The bandwidth does not deteriorate, as it is shown in Figure 11-21.

Tolerances on Magnet Vibration and Position Drift

Steering Tolerance. *a) Quadrupole Vibration:* According to the jitter-budget proposed by J. Irwin (Table 9-3), the maximum tolerable incoming vertical orbit-jitter is about $0.44\sigma_y$ corresponding to a 2.4%

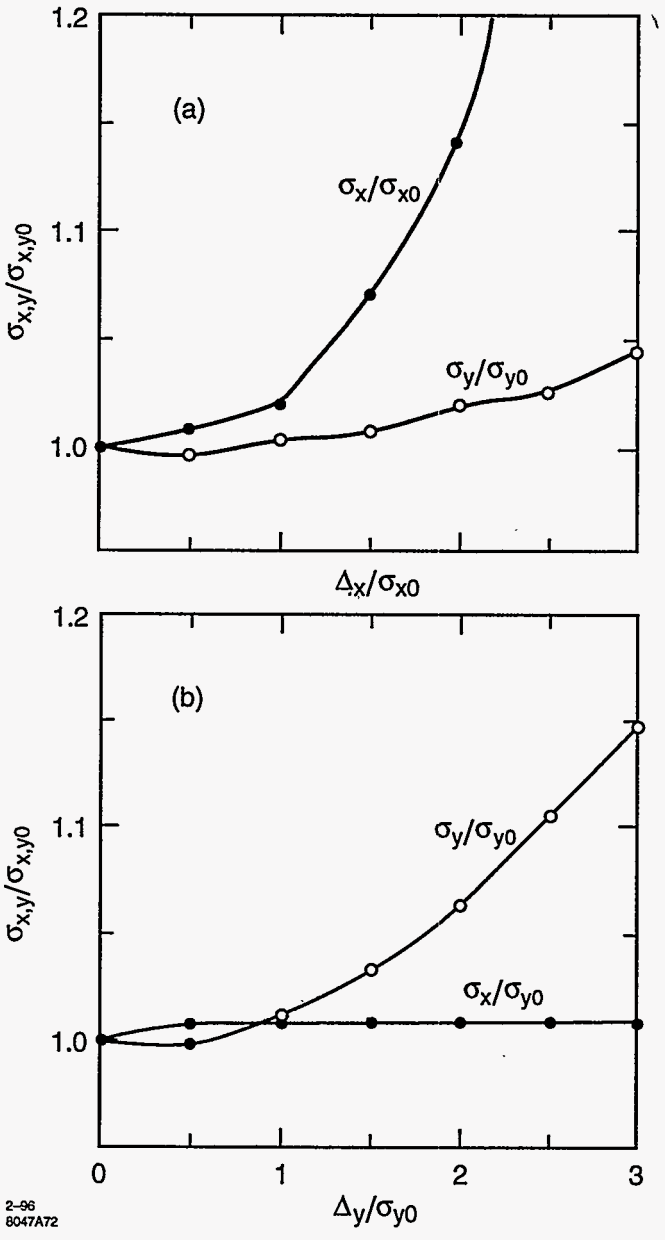


Figure 11-19. *Relative increases of the IP spot sizes as a function of horizontal or vertical incoming orbit variation in units of the design beam size. Effects of synchrotron radiation are not included.*

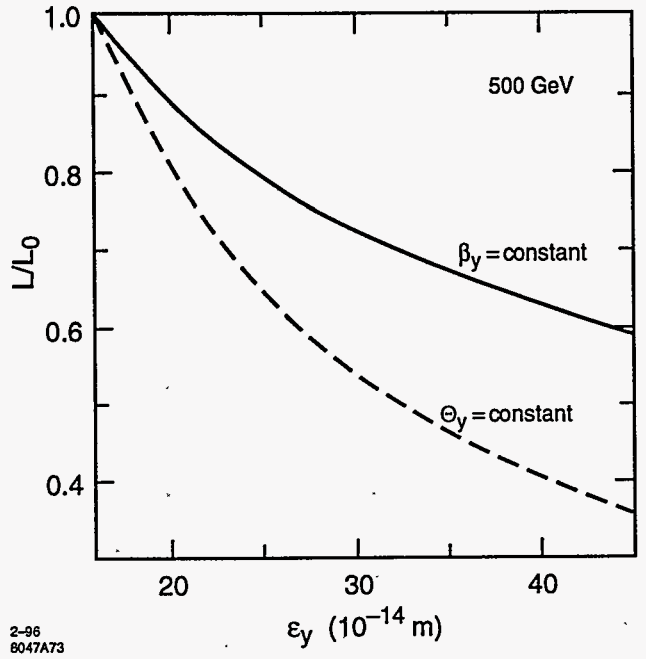


Figure 11-20. *Relative loss of luminosity for larger vertical emittances, keeping either IP beta function or divergence constant.*

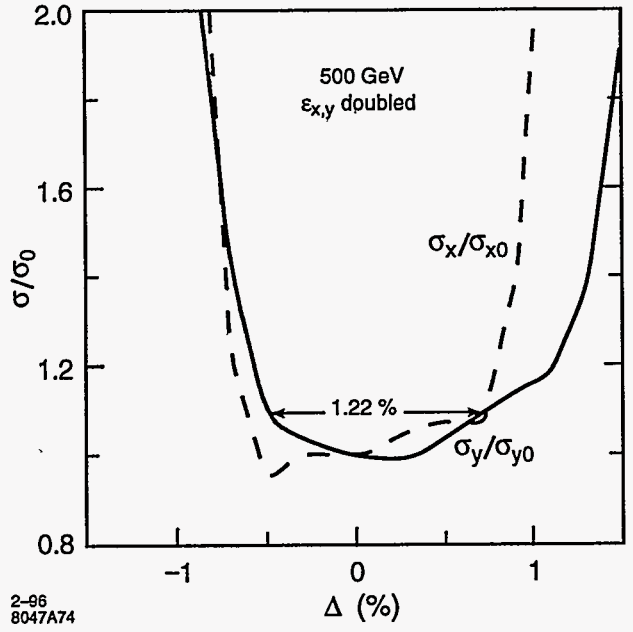


Figure 11-21. *Momentum bandwidth of the 500-GeV final-focus system, for increased emittances,* $\gamma\epsilon_x = 10^{-5}$ *m,* $\gamma\epsilon_y = 17.5 \times 10^{-8}$ *m, and constant IP beta functions,* $\beta_x^* = 8$ *mm,* $\beta_y^* = 125\ \mu$*m.*

loss of luminosity. The contributions to the IP position jitter from vertical and horizontal vibrations of magnets in the final focus (not including the final doublet) shall result in no more than an additional 0.5% luminosity reduction each.

The tolerance on the quadrupole vibration amplitude is thus calculated by requiring that the resulting change of the horizontal (vertical) IP beam position is not larger than $0.1\sigma_x$ ($0.2\sigma_y$). The vertical tolerance is looser by a factor of two, due to the large vertical disruption.

For a single quadrupole, the maximum vibration amplitude Δx, Δy is given by [Roy 1992]

$$\Delta x < \frac{\sigma_x^*}{10 R_{12} k_q} \tag{11.17}$$

$$\Delta y < \frac{\sigma_y^*}{5 R_{34} k_q} \tag{11.18}$$

where R_{12} and R_{34} are the R-matrix elements from the quadrupole to the IP and k_q denotes the integrated quadrupole strength. Sensitivities to both horizontal and vertical quadrupole vibrations are listed in Tables 11-13 and 11-14. Note that these numbers apply to each magnet individually, and that the actual vibration tolerance is given in Table 11-19. The tightest sensitivities (again not including the final doublet) for a 0.5 % luminosity loss, due to a single magnet, correspond to vibration amplitudes of about 70 nm horizontally and 14 nm vertically, for about seven magnets at 1-TeV-c.m. energy (Table 11-14, and Figures 11-23 and 11-24).

The sensitivity for skew-quadrupole vibrations can be calculated in a similar way. Some resulting tolerances for the correction skew-quadrupoles in the SCS can be found in Table 11-28.

b) Roll of Bending Magnet: If a bending magnet vibrates in the x-y plane, it can steer the beam vertically at the IP. The maximum allowed roll-angle θ_{xy}, for a $0.2\sigma_y$ change of the IP beam position, is

$$\theta_{xy} \leq \frac{\sigma_y^*}{5 R_{34}^{bIP} \theta} \tag{11.19}$$

where σ_y^* denotes the vertical IP beam size, R_{34}^{bIP} the (3,4)-R matrix element from the bending magnet to the IP, and θ the bending angle. The tolerances for most magnets are fairly tight (Table 11-16 and Figure 11-31), eg a change in roll by $2\,\mu$r of a single bend magnet in the BX can reduce the luminosity by 1%. This is a jitter tolerance, since steering errors at the IP will be corrected in less than 0.2s by a fast IP collision feedback, akin to that of the SLC.

Dispersion. A displaced or rolled magnet can generate dispersion at the IP. If the dispersion is tuned and corrected in hourly intervals, the corresponding tolerances are position-drift tolerances, which for some magnets may be difficult or expensive to meet. Fortunately, dispersion can be corrected continuously, if one utilizes the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy for both beams. For this purpose, the pulse-to-pulse energy variation could be intentionally increased, by lowering the gain of one or more energy feedback loops.

a) Quadrupole-Position Drift: A displaced quadrupole increases the IP spot size due to dispersion, which is generated both directly by the chromatic kick from the displaced quadrupole and by the orbit change in the downstream elements. The spot-size increases by 1% or less, if the quadrupole-position drift Δx or Δy satisfies

$$\Delta x < \frac{\sigma_x^*}{5\sqrt{2\delta}\,|-R_{12}+T_{126}|\,k_q} \tag{11.20}$$

$$\Delta y \;<\; \frac{\sigma_y^*}{5\sqrt{2}\delta \mid -R_{34}+T_{346}\mid k_q} \tag{11.21}$$

where δ is the rms momentum spread ($\delta \approx 0.3\%$), and T_{126}, T_{346} denote the T-matrix elements (in TRANSPORT notation) from the quadrupole to the IP. For magnets inside or behind the CCY, the largest contribution to these two coefficients arises from the uncompensated chromaticity of the final doublet.

At 1 TeV, the maximum tolerable vertical dispersion at the IP is 160 nm, corresponding to a 1% blow-up of the vertical spot size for an intrabunch energy spread of $\delta \approx 0.3\%$. The vertical displacement corresponding to a 1% luminosity loss is about 40 nm for the first two quadrupoles in the FT (QFT5 and QFT6) and for QEI4, further upstream, and it is 140 nm for the two quadrupoles in the center of the CCY (Table 11-14). These sensitivities correspond to an orbit change by about 110 nm at the end of the first doublet magnet (QFT3), which needs to be detected and corrected by a fast orbit feedback. The orbit variation to be measured is small compared with the vertical rms beam size of 60 μm. Nevertheless, simulations show that orbit correction with the desired accuracy is possible in less than 2 s, based on four 1-μm-resolution BPMs. These simulations are presented in Section 11.5.6, "CCY and FT Orbit Feedback." The feedback response time could be further improved by using rf BPMs, which have much higher resolution.

Horizontal displacement sensitivities due to dispersion are 140 nm or larger.

b) Sextupole-Position Stability: Sextupole motion at a location with nonzero dispersion generates dispersion at the IP. The corresponding maximum motion for a 1% spot-size increase is

$$\Delta x_s^\eta \;<\; \frac{\sqrt{2}\sigma_x^*}{5k_s\delta\eta_s R_{12}} \tag{11.22}$$

$$\Delta y_s^\eta \;<\; \frac{\sigma_y^*}{5\sqrt{2}k_s\delta\eta_s R_{34}} \tag{11.23}$$

where δ denotes the rms momentum spread, $R_{12,34}$ are the R-matrix elements from the sextupole to the IP, $\sigma_{x,y}$ the IP beam sizes, η_s is the nominal dispersion at the sextupole, and k_s the integrated sextupole strength.

In case the dispersion is not corrected continually, this is both a vibration and a stability tolerance, which has to be met over 30–60 minutes.

c) Bend Roll: Also a rolled bending magnet generates dispersion at the IP. The bend roll causing a 1% luminosity loss is given by

$$\theta_{xy}^b \leq \frac{\sigma_y^*}{5\sqrt{2}|T_{346}-R_{34}|\theta\delta} \tag{11.24}$$

where θ is the bend angle, δ the rms momentum spread, and R_{34}, T_{346} denote the R- and T- matrix elements from the bending magnet to the IP.

Skew Coupling and Waist Shift. *a) Sextupole-Position Stability:* In addition to dispersion, sextupole motion causes waist shift or skew coupling, and also thereby increases the IP spot size. The corresponding maximum motion for a 1% luminosity loss is:

$$\Delta x_s^w \;<\; \frac{1}{5\sqrt{2}k_s \max(\beta_x,\beta_y)} \tag{11.25}$$

$$\Delta y_s^s \;<\; \frac{1}{5\sqrt{2}k_s\sqrt{\beta_x\beta_y}}\sqrt{\frac{\epsilon_y}{\epsilon_x}} \tag{11.26}$$

where ϵ_x, ϵ_y denote the emittances and k_s the integrated sextupole strength. Values for the 15 sextupoles in the 500-GeV final focus are listed in Table 11-15. The tightest sensitivities for a 1% luminosity loss are about 500 nm horizontally and vertically for the two main Y-sextupoles (Figures 11-28 and 11-29). Again the actual tolerances that have to be achieved are tighter than those quoted, in this case by roughly a factor of two assuming uncorrelated sextupole motion.

b) Quadrupole-Position Stability (or Drift): The same skew coupling or waist shift is generated when a quadrupole moves inside the CCX or CCY steering the beam off-center through the second sextupole. The maximum drift amplitudes due to this effect for a 1% luminosity loss are:

$$\Delta y_q \leq \frac{\sqrt{2}\Delta y_s^s}{R_{34}^{qs} k_q} \tag{11.27}$$

$$\Delta x_q \leq \frac{\sqrt{2}\Delta x_s^w}{R_{12}^{qs} k_q} \tag{11.28}$$

where Δx_q and Δy_q denote the motion tolerances, due to waist shift or skew coupling, R_{34}^{qs} and R_{12}^{qs} the R-matrix elements from the quadrupole to the next sextupole, and k_q the integrated quadrupole strength. The term Δy_s^s is the tolerance for sextupole movement giving rise to a skew aberration, and Δx_s^w is the tolerance for sextupole movement giving rise to a waist aberration. For a few quadrupoles in the CCY, the vertical (horizontal) displacement sensitivity due to this effect can be as tight as 110 (500) nm, for a spot size increase by 1% (Figures 11-25 and 11-26 present sensitivities for 1% luminosity loss).

c) Quadrupole Roll: A quadrupole rolling in the x-y plane may generate skew coupling at the IP. The tolerance on the x-y roll angle θ_{xy} for a 1% spot size increase is given by

$$\theta_{xy} \leq \frac{1}{10\sqrt{2} k_q \sqrt{\beta_x \beta_y}} \sqrt{\frac{\epsilon_y}{\epsilon_x}} \tag{11.29}$$

where β_x, β_y are the horizontal and vertical beta function at the quadrupole, and k_q is the integrated quadrupole strength. Roll sensitivities for all quadrupoles are listed in Table 11-14. The tightest actual tolerances are those for the final-doublet magnets QFT3–QFT1: about 1.5 μr in 1 h (compare Figure 11-27).

d) Bend Roll: A rolled bending magnet may steer the beam vertically at the next sextupole, thus generating skew coupling. This effect gives rise to a second drift tolerance on the bend roll angle, in addition to that arising from vertical dispersion. The bend roll drift causing a 1% luminosity loss due to skew coupling is

$$\theta_{xy}^b \leq \frac{\Delta y_s^s}{\sqrt{2} 5 R_{34}^{bSX} \theta} \tag{11.30}$$

where θ is the bend angle, Δy_s^s the drift tolerance for the downstream sextupole due to skew coupling, and R_{34}^{bSX} the R-element between bending magnet and sextupole.

Second-Order Dispersion. *a) Quadrupole Alignment:* The maximum tolerable second-order dispersion at the IP is about 35 μm, for a 1% luminosity loss and an rms energy spread of $\delta_{\rm rms} \approx 0.3\%$. If no tuning scheme for this aberration is implemented, the number above translates into a vertical alignment tolerance for quadrupole magnets in the final focus. Second-order dispersion is primarily generated by interaction of a chromatic kick from a displaced quadrupole with the doublet chromaticity. In the absence of a correction scheme, the absolute alignment tolerance for a single final-focus quadrupole due to generated second-order dispersion would be

$$\Delta y < \frac{\sigma_y^*}{10\,\delta^2\,|T_{346}|\,k_q}. \tag{11.31}$$

The tightest sensitivities of about 12 μm are those for the first two magnets in the FT (QFT6 and QFT5). Sensitivities for all quadrupoles are given in Tables 11-13 and 11-14.

One possibility for compensating the second-order dispersion is by vertical moves of quadrupoles QFT6 and QFT5. A quadrupole move will also create first-order dispersion, which is of similar magnitude as the second-order dispersion and which, therefore, will dominate the spot size during a scan. Specifically, we have the following relation between first- and second-order dispersion $\eta_y^{(1)}$, $\eta_y^{(2)}$ at the IP and the vertical quadrupole positions Δy_{QFT5}, Δy_{QFT6}:

$$\eta_y^{(1)} = 3.8 \times \Delta y_{QFT5} - 3.0 \times \Delta y_{QFT6} \tag{11.32}$$

$$\eta_y^{(2)} = -4.0 \times \Delta y_{QFT5} + 1.7 \times \Delta y_{QFT6} \tag{11.33}$$

In principle, the two quadrupoles can be moved simultaneously such that only second-order dispersion is generated, but no first-order dispersion. As an example, if QFT5 is moved by $-18\,\mu$m and QFT6 by $-23\,\mu$m, a second-order dispersion of $35\,\mu$m is produced, and the first-order dispersion is not changed. However, in order to realize such a pure $\eta_y^{(2)}$ tuning knob, the mover positions would need to be controlled with an unrealistic precision of 30 nm. Otherwise the measured spot sizes during an $\eta_y^{(2)}$-scan are still affected by residual first-order dispersion. Hence, the second-order dispersion cannot be tuned by a simple scan, but instead an iterative procedure is necessary: after each QFT5/QFT6 move the first-order dispersion is corrected by the standard procedure using sextupole movements in the CCY, before the vertical spot size can be measured. Such a 2-dim. scan (or a "scan of scans") would be a new type of procedure, not used at the SLC.

This complication can be avoided by utilizing the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy, as it has already been proposed for correcting the regular dispersion. In this case, the second-order dispersion can be continually tuned, and it then causes no significant spot-size dilution.

Power Supply Tolerances

Steering Tolerance. *a) Bending Magnets:* A field change of a bending magnet steers the beam horizontally at the IP. To give a horizontal steering error smaller than 0.1 σ_x, the relative field-ripple of the bending magnets must satisfy

$$\frac{\Delta\theta}{\theta} \leq \frac{\sigma_x^*}{10\, R_{12}^{bIP}\theta} \tag{11.34}$$

where θ denotes the nominal bending angle, and R_{12}^{bIP} the R-matrix element from the bending magnet to the IP. Sensitivities for bending magnets in the different bending sections are listed in Table 11-16 (Figure 11-30). The actual tolerances which need to be achieved are quite tight, on the order of 3×10^{-6}. This is a "jitter" tolerance, which has to be met over 1/5 s.

Waist Shift. *a) Bending Magnets:* Similar to the effect of horizontally drifting quadrupoles (Eq. 11.27), a field change of the bending magnets inside the CCX or the CCY steers the beam horizontally off-center through the next sextupole, which causes a waist shift at the IP. The maximum allowed field ripple for a 1% luminosity loss is given by

$$\frac{\Delta\theta}{\theta} = \frac{\Delta x_s^w}{R_{12}^{bs}\theta} \tag{11.35}$$

where Δx_s^w denotes the horizontal vibration tolerance due to waist shift of the next sextupole (SX1b or SY1b), θ the nominal bending angle, and R_{12}^{bs} the R-matrix element from the bending magnet to the sextupole. Some numbers may be found in Table 11-16. Assuming 10-m-long bend sections inside the CCX or the CCY are fed by independent power supplies, the tolerable relative field ripple is about 2×10^{-5}. The tolerance decreases to 4×10^{-6}, if all magnets are powered as a string. This is a stability tolerance (timescale 1 hr), unless there is an orbit stabilization system such as discussed in Section 11.5.6: "CCY and FT Orbit Feedback."

b) Quadrupoles: The change of a quadrupole field also causes a waist shift at the IP. For an associated spot-size increase smaller than 1%, the maximum field change is written [Roy 1992]

$$\frac{\Delta k_q}{k_q} = \frac{1}{5\sqrt{2}\max(\beta_x, \beta_y)k_q} \tag{11.36}$$

where k_q is the integrated quadrupole strength, and $\beta_{x,y}$ the beta function at the quadrupole. The sensitivities for the quadrupoles of CCX, CCY, and FT are listed in Table 11-17. The tightest sensitivities (except for the final doublet) for a 1% luminosity loss are about 10^{-4}. Again, the final tolerance that actually needs to be achieved will be smaller than the quoted numbers, each of which refers to ripple in a single magnet only.

Dispersion. Field changes of magnets may give rise to dispersion at the IP. It is contemplated to continually monitor and correct the IP dispersion, using the correlation of beam-beam pulse-to-pulse deflection with pulse-to-pulse energy of both beams, so as to render the dilution due to dispersion as insignificant.

a) Quadrupoles: In addition to causing a waist shift, a field change of quadrupole magnets at locations with nonzero dispersion also generates horizontal dispersion at the IP. The relative field change for a 1% loss of luminosity due to the induced dispersion is given by

$$\frac{\Delta k_q}{k_q} = \frac{\sigma_x^*}{5\sqrt{2}\delta k_q \eta_x^q R_{12}} \tag{11.37}$$

where k_q is the integrated quadrupole field, R_{12} the R-matrix from the quadrupole to the IP, and η_x^q the value of the dispersion at the quadrupole. Sensitivities are listed in Table 11-17.

The relative field-drift sensitivity of four CCX quadrupoles is 4×10^{-4} for a 1% luminosity loss due to generated dispersion. Their sensitivity due to induced waist shift is about twice as tight.

b) Bending Magnets: Horizontal dispersion is also generated by a field change of a bending magnet which alters the downstream orbit. The change of bend angle $\Delta\theta$ causing a 1% luminosity loss is given by

$$\Delta\theta \leq \frac{\sigma_x^*}{5\sqrt{2}|T_{126} - R_{12}|\delta} \tag{11.38}$$

where θ denotes the bending angle, δ the rms momentum spread, σ_x^* the horizontal IP spot size, and T_{126}, R_{12} the T- and R-matrix elements from the bend to the IP (in TRANSPORT notation).

Chromaticity. *a) Sextupoles:* The sextupole field-drift sensitivity due to the induced chromaticity follows from

$$\frac{\Delta k_s}{k_s} \leq \frac{1}{5\sqrt{2}k_s \max(\beta_{s,x}, \beta_{s,y})\eta_{s,x}\delta} \tag{11.39}$$

where $\beta_{s,x,y}$ denotes the horizontal or vertical beta function at the sextupole, $\eta_{s,x}$ the dispersion and k_s the integrated sextupole strength in units of m^{-2}. Sensitivities are not very tight (Table 11-18).

Geometric Aberrations. *a) Sextupoles:* The sextupole field-drift sensitivity due to induced third-order geometric aberrations is approximately

$$\frac{\Delta k_s}{k_s} \leq \frac{1}{5\sqrt{2}k_s\sqrt{\beta_{s,x}\epsilon_x}\max\left(\frac{2}{\sqrt{3}}\beta_{s,x},\beta_{s,y}\right)} \tag{11.40}$$

where $\beta_{s,x}$ and $\beta_{s,y}$ denote the horizontal and vertical beta function at the sextupole, and k_s the integrated sextupole strength. The resulting tolerances are loose and similar to those imposed by chromaticity (Table 11-18).

Magnet	Steering		Spot Size			
			disp. and skew		2nd o. disp.	roll
	Δx (nm)	Δy (nm)	Δx (μm)	Δy (μm)	Δy (μm)	θ_{xy} (mrad)
QDD	1700.	150.	7.	0.14	33.	4.9
QDF	600.	650.	11.	0.26	61.	7.6
QDD	660.	440.	37.	0.30	70.	4.3
QDF	400.	220.	6.	1.92	470.	2.5
QDD	640.	100.	6.	0.49	110.	2.2
QDF2	340.	80.	2.	0.20	46.	1.6
QDD2	2300.	50.	12.	0.09	22.	3.0
QDF2	240.	90.	4.	0.15	35.	1.3
QDD	640.	150.	6.	0.18	42.	4.0
QDF	550.	650.	3.	0.36	85.	10.0
QDD	1400.	440.	5.	0.51	120.	8.4
QDF	9200.	220.	4.	7.54	1600.	4.7
QDD	1700.	100.	7.	0.41	95.	2.7
QDF	600.	130.	8.	0.37	87.	1.9
QDD	660.	90.	65.	0.20	47.	1.8
QDF	400.	150.	12.	0.29	69.	22.3
QDD	640.	150.	9.	0.23	54.	4.0
QDF2	340.	720.	3.	0.35	81.	6.4
QDD2	2300.	240.	12.	0.49	116.	7.2
QDF2	240.	160.	11.	13.16	2290.	1.6
QDD	640.	100.	11.	0.52	120.	2.2
QDF	550.	130.	5.	0.52	120.	2.6
QDD	1400.	90.	6.	0.29	67.	3.6
QDF	9200.	150.	4.	0.45	104.	4.3
QDD	1700.	150.	7.	0.37	86.	4.9
QDF	600.	650.	7.	0.87	204.	7.6
QDD	660.	440.	17.	2.62	632.	4.3
QG1	430.	240.	122.	2.41	545.	2.6
QG2	1200.	70.	11.	0.36	82.	2.8
QG1	2300.	250.	3.	1.73	395.	4.5
QG2	1700.	3000.	8.	1.36	320.	6.8
QB1	2100.	1200.	39.	8.71	1993.	13.0
QB2	2700.	950.	147.	15.65	3450.	8.8
QB3	880.	370.	199.	19.44	3762.	1.3
QB4	280.	70.	32.	5.19	947.	0.2
QB5	84.	50.	9.	3.75	668.	0.1

Table 11-13. *Vibration, drift and roll sensitivity of quadrupoles between the SCS and BMS in the 1-TeV final focus; each number corresponds to a horizontal (vertical) position error of $0.1\sigma_x$ ($0.2\sigma_y$), or to a 1% spot-size increase, respectively.*

Magnet	Steering		Spot Size			
			disp. and skew		2nd o. disp.	roll
	Δx (nm)	Δy (nm)	Δx (μm)	Δy (μm)	Δy (μm)	θ_{xy} (μrad)
QE2	430.	89.	280.0	9.8	1600.	200.
QE2	400.	89.	240.0	9.8	1600.	200.
QE1	76.	97.	26.0	7.7	1400.	63.
QX3	76.	97.	26.0	7.7	1400.	63.
QX2	250.	57.	4.0	2.3	830.	120.
QX2	260.	58.	4.0	2.3	830.	130.
QX1	30000.	6800.	1.0	3.5	1300.	800000.
QX1	15000.	3400.	1.0	3.5	1300.	300000.
QX2	260.	58.	3.0	1.8	670.	130.
QX2	250.	57.	3.0	1.8	680.	120.
QX3	76.	98.	28.0	9.4	1600.	63.
QBX1	76.	97.	32.0	8.1	1400.	63.
QBX2	250.	57.	8.0	1.2	260.	120.
QBX2	260.	58.	7.0	1.2	260.	130.
QBX3	30000.	6800.	2.0	2.0	480.	800000.
QBX4	22000.	4900.	3.0	2.9	680.	480000.
QBX5	260.	58.	2.0	8.3	4800.	91.
QBX5	260.	56.	2.0	8.4	5200.	88.
QBX6	440.	17.	46.0	6.8	2400.	45.
QY3	440.	17.	47.0	7.0	2300.	45.
QY2	260.	57.	0.6	0.7	2500.	89.
QY2	260.	59.	0.6	0.7	2400.	92.
QY1	43000.	9600.	1.0	0.2	52.	—
QY1	22000.	4800.	1.0	0.2	52.	480000.
QY2	260.	58.	0.6	0.3	90.	90.
QY2	260.	56.	0.6	0.3	93.	88.
QY3	440.	17.	46.0	8.4	1800.	45.
QEI1	440.	17.	96.0	5.9	3100.	45.
QEI2	260.	57.	2.0	0.3	88.	89.
QEI2	260.	59.	2.0	0.4	85.	92.
QEI3	43000.	9700.	3.7	0.1	25.	—
QEI4	85000.	19000.	14.0	0.4	99.	—
QFT6	50.	11.	1.0	0.1	13.	9.
QFT5	70.	8.	1.5	0.1	11.	9.
QFT4	120.	17.	48.0	0.3	68.	20.
QFT3	12.	3.	2.2	0.7	85.	3.
QFT2	17.	2.	3.2	0.7	600.	3.
QFT1	21.	1.	4.6	0.1	79.	4.
QFTA	37.	1.	9.0	0.1	67.	9.

Table 11-14. *Vibration, drift, and roll sensitivity of quadrupoles between the CCX and IP in the 1-TeV final focus; each number corresponds to a horizontal (vertical) position error of $0.1\sigma_x$ ($0.2\sigma_y$), or to a 1% spot-size increase, respectively.*

Magnet	Δx (μm)	Δy (μm)
SX1	3.0	1.9
SX3	260000.0	46000.0
SX1	3.0	1.9
SY1	0.5	0.5
SY3	460000.0	81000.0
SY1	0.5	0.5
SI2	320000.	57000.0

Table 11-15. *Sextupole vibration sensitivity for the 1-TeV final focus due to induced skew coupling or waist shift, for a 1% spot-size increase.*

Name	$\Delta\rho/\rho$ (10^{-5}) (steering)	$\Delta\rho/\rho$ (10^{-4}) (beam size)	θ_{xy} (μrad)
BG	16.5	—	96.0
B1A	6.5	—	15.0
B1B	5.2	—	11.0
B2	1.7	2.3	11.0
B3A	3.5	—	8.0
B3B	3.5	—	1.7
B4	3.3	0.8	3.4
B5A	5.7	—	6.7
B5B	4.4	—	10.0
B5C	10.0	—	23.0

Table 11-16. *Field ripple and roll sensitivity for the bending magnets. Each number corresponds to a steering error by $0.1\sigma_x$, $0.2\sigma_y$ or to a spot-size increase by 1%, respectively.*

Magnet	Waist Shift $\Delta k/k$ (10^{-3})	Dispersion $\Delta k/k$ (10^{-3})	Total $\Delta k/k$ (10^{-3})
QE2	2.10	8.5	2.08
QE2	2.20	7.6	2.07
QE1	0.20	0.4	0.18
QX3	0.20	0.4	0.18
QX2	1.40	3.5	1.30
QX2	1.40	3.7	1.30
QX1	11000.00	240.0	240.00
QX1	4200.00	120.0	120.00
QX2	1.40	3.8	1.30
QX2	1.40	3.6	1.30
QX3	0.20	0.5	0.18
QBX1	0.20	0.5	0.18
QBX2	1.40	5.0	1.34
QBX2	1.40	5.0	1.37
QBX3	11000.00	72000.0	11000.00
QBX4	5900.00	25000.0	5800.00
QBX5	1.00	1.9	0.90
QBX5	1.00	1.9	0.86
QBX6	0.09	4.0	0.09
QY3	0.09	4.0	0.09
QY2	1.00	1.3	0.77
QY2	1.00	1.3	0.80
QY1	16000.00	460.0	460.00
QY1	5900.00	230.0	230.00
QY2	1.00	1.3	0.80
QY2	0.96	1.3	0.80
QY3	0.09	4.0	0.09
QEI1	0.09	4.0	0.09
QEI2	0.98	1.4	0.80
QEI2	1.00	1.4	0.80
QEI3	16000.00	890.0	890.00
QEI4	23000.00	1800.0	1800.00
QFT6	0.10	330000.0	0.10
QFT5	0.05	720000.0	0.05
QFT4	0.15	—	0.15
QFT3	0.04	200000.0	0.04
QFT2	0.02	390000.0	0.02
QFT1	0.01	990000.0	0.01
QFTA	0.02	—	0.02

Table 11-17. *Field ripple sensitivity for quadrupoles between the CCX and IP due to induced waist shift or due to horizontal dispersion generated at the IP. Each listed value corresponds to a 1% spot-size increase.*

Name	$\Delta k_S/k_S$ (10^{-3})		
	Chromaticity	Geom. aberrations	Total
SX1	12.0	4.7	4.4
SX3	—	—	—
SX1	12.0	4.7	4.4
SY1	2.9	3.4	2.2
SY3	—	—	—
SY1	2.9	3.4	2.2
SI2	—	—	—

Table 11-18. *Field ripple sensitivity due to chromaticity or due to geometric aberrations for the normal sextupoles in the CCX, BX, CCY, and FT: each number listed corresponds to a 1% (horizontal or vertical) beam-size increase.*

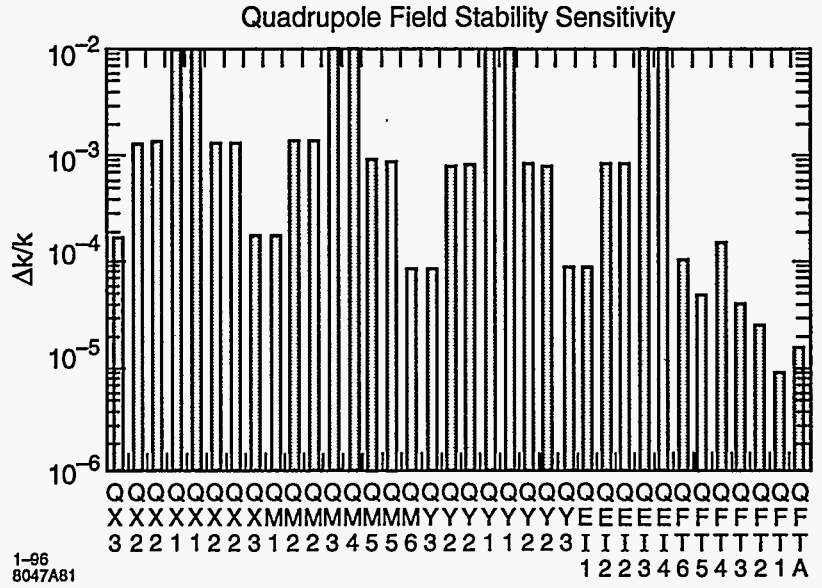


Figure 11-22. *Field-stability sensitivity of quadrupoles in the BMS, CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

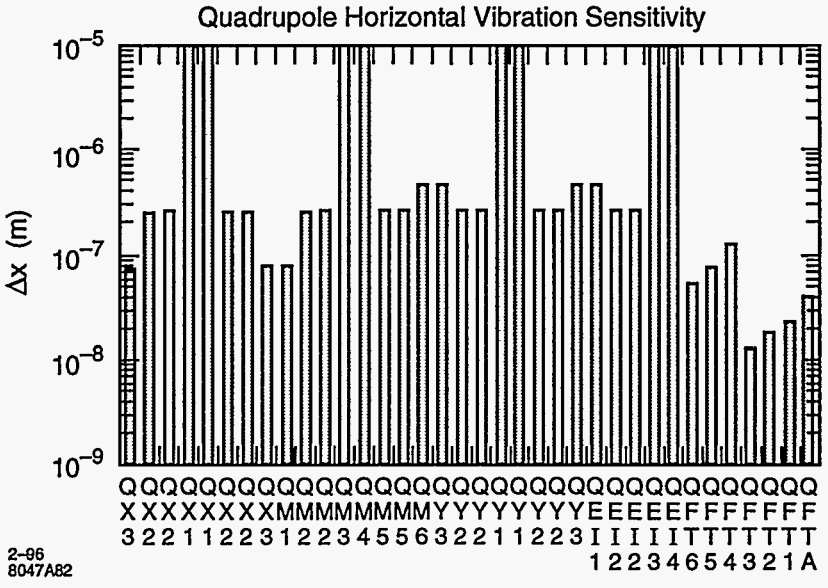


Figure 11-23. *Horizontal vibration sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

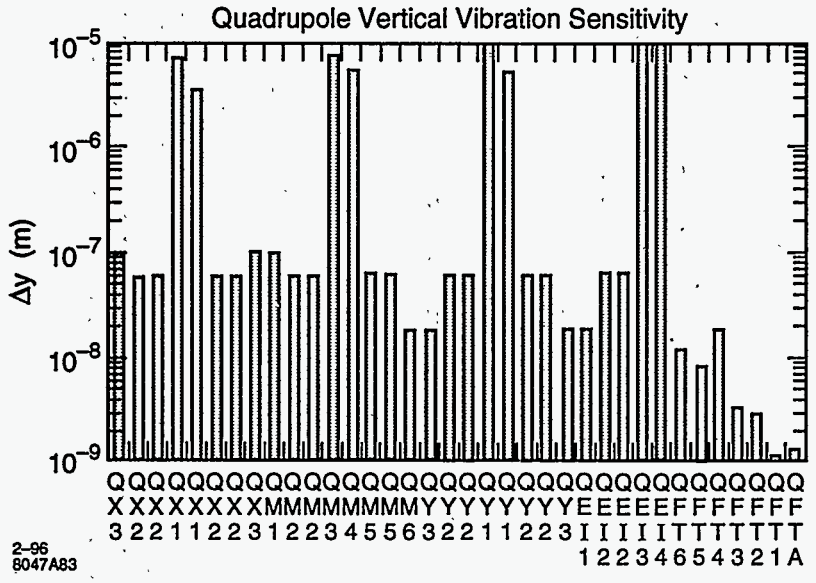


Figure 11-24. *Vertical vibration sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

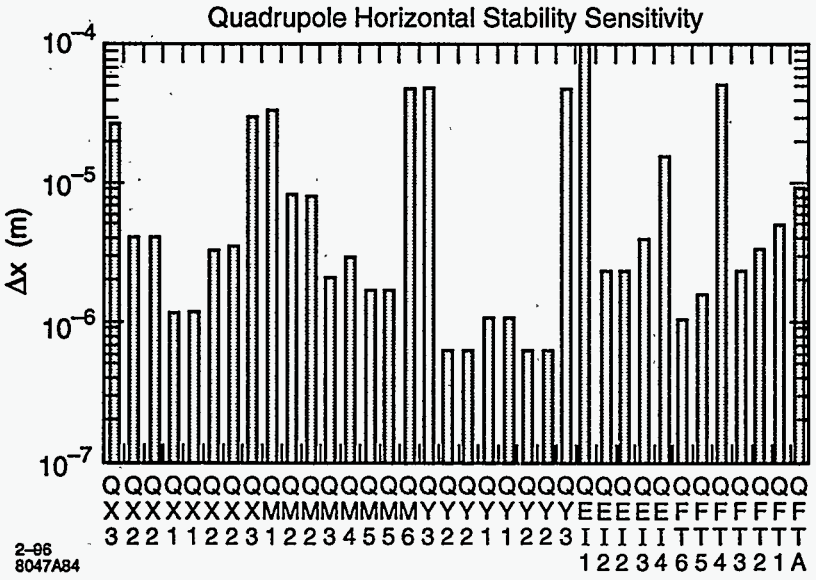


Figure 11-25. *Horizontal stability (or position-drift) sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

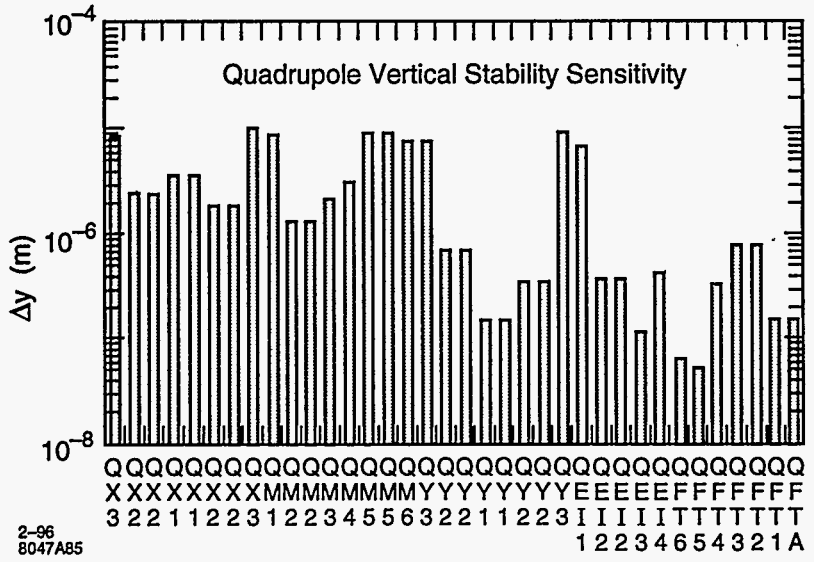


Figure 11-26. *Vertical stability (or position-drift) sensitivity of quadrupoles in CCX, BX, CCY and FT; each corresponding to a luminosity loss of 1%.*

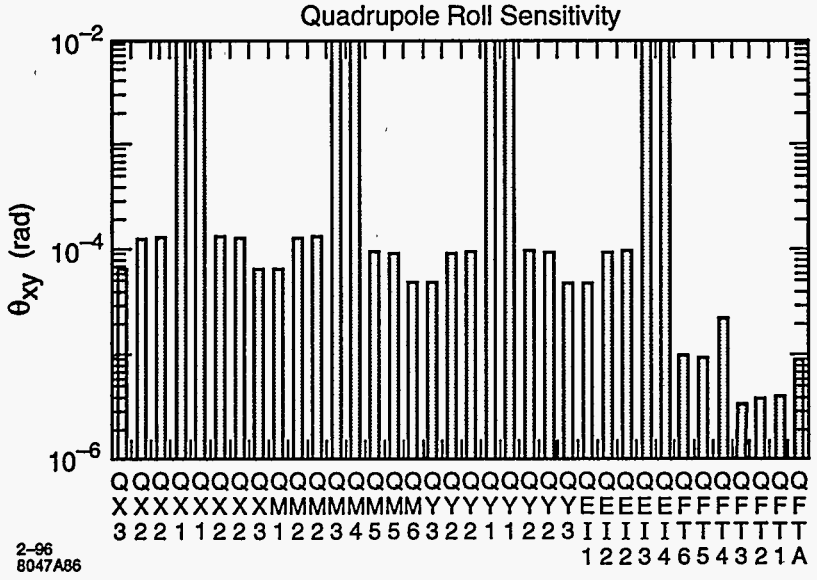


Figure 11-27. *Roll sensitivity of quadrupoles in the CCX, BX, CCY, and FT; each corresponding to a luminosity loss of 1%.*

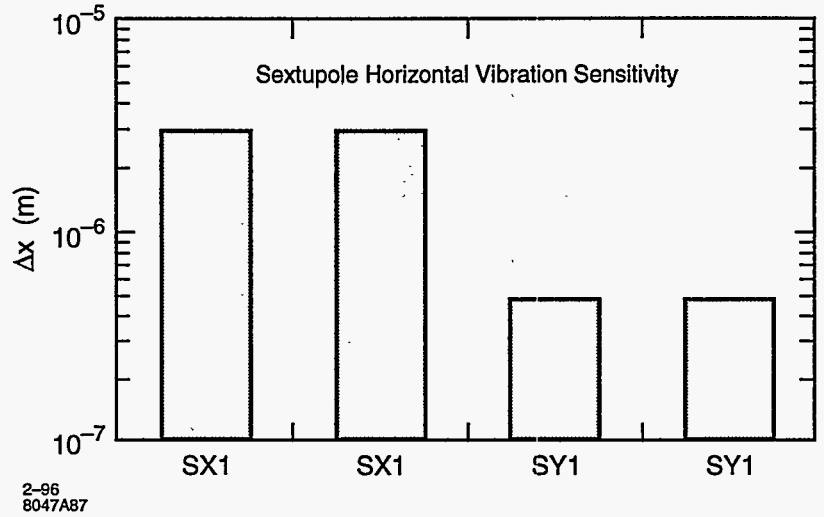


Figure 11-28. *Horizontal vibration sensitivity of main sextupoles in the final focus; each corresponding to a luminosity loss of 1%.*

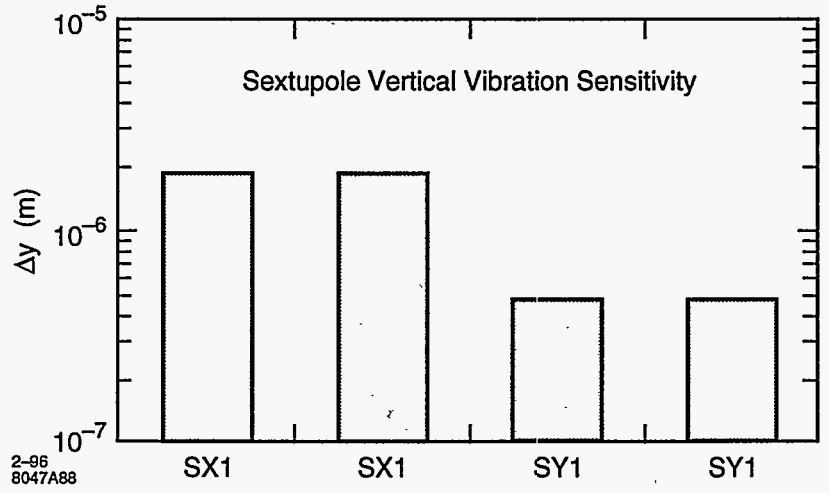


Figure 11-29. *Vertical vibration sensitivity of main sextupoles in the final focus; each corresponding to a luminosity loss of 1%.*

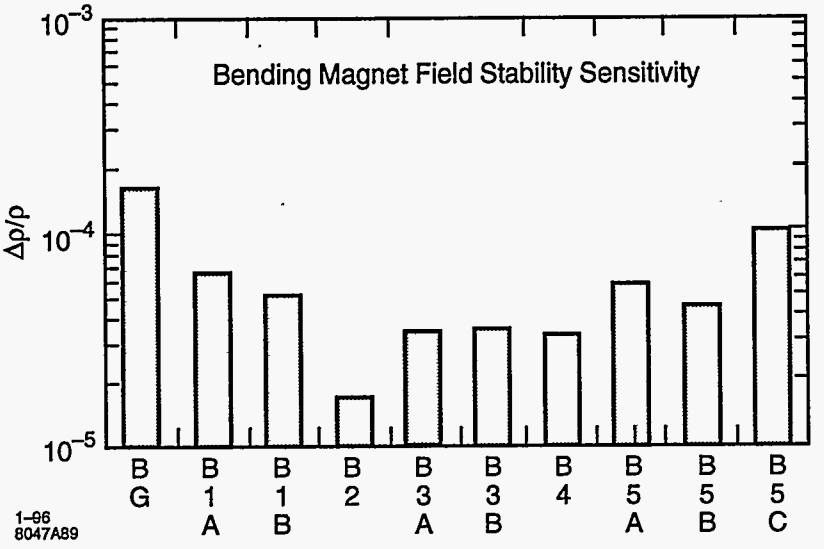


Figure 11-30. *Field-stability sensitivity for bending magnets in the final focus, assuming uncorrelated field ripple; each corresponding to a luminosity loss of 1%.*

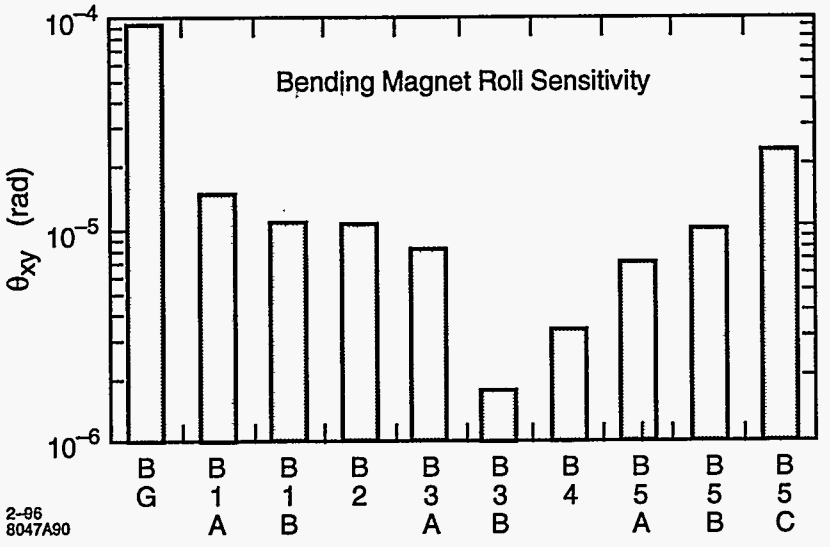


Figure 11-31. *Roll sensitivity of bending magnets in the final focus; each corresponding to a luminosity loss of 1%.*

Scaling of Tolerances

All vibration and stability (drift) tolerances presented in this section refer to a particular parameter set for 1-TeV-c.m. energy (Table 11-1). Most of the sensitivities scale directly with the IP beam size, independently of whether the latter changes due to an emittance variation or due to a different beta function. An exception are quadrupoles in front of the BMS, whose sensitivity is not affected by a change of the IP beta function. From a tolerance point of view, the operation at 500 GeV is equal to an emittance increase. Thus, almost all tolerances at 500 GeV will be a factor of $\sqrt{2}$ looser than those quoted in this chapter.

Tolerance Budget and Luminosity Loss

Beam-beam deflection scans will be used for correcting low-order aberrations which affect the IP spot size, such as waist shift, skew coupling etc.. Because of limited resolution of the deflection scans, these corrections cannot be perfect, but they are accurate to about 1%, implying a corresponding spot-size increase. Since this spot-size increase is dictated by the tuning method, it sets a lower limit on practicable tolerances, *e.g.*, for element strength changes and position drifts. Between two scans of the same aberration, the spot-size increase due to this aberration is assumed to grow from 1% to about 3%, because of magnet field and position drifts. On average (over time) the residual spot-size increase is then expected to be on the order of 2% for each tunable aberration, although various possibilities exist to reduce this number to 0.5% (see discussions in Section 11.5.3 and also Table 11-5). Thus, for consistency, a set of tolerances on magnet motion and field stability was chosen which roughly assigns a 2% spot-size increase caused by all components between linac and IP, or 1% generated in the final focus alone, to each such perturbation. Figure 11-32 illustrates that, for the tolerances chosen, various different effects may contribute by similar amounts to each aberration.

To facilitate design and cost estimates, we have grouped the magnets into different classes. The same tolerances then apply to all magnets of one class. Specifically, the quadrupoles are divided into four classes. The majority of the final-focus magnets is in Class 1. Four quadrupoles in the CCY with tighter tolerances form Class 2. Three quadrupoles at the entrance of the final transformer, with still tighter tolerances, are in Class 3, and the final-doublet magnets define Class 4. The sextupoles are split into two classes: main sextupoles and Brinkmann-sextupoles. The bending magnets are not further subdivided. A list of the different magnet types, along with relevant tolerances and the number of magnets of each type, is presented in Table 11-19. Table 11-20 describes the individual tolerances for each magnet class and how they can be met. The overall loss of luminosity corresponding to these magnet tolerances is 11%; see Table 11-21. This number assumes that dispersion is corrected at hourly intervals only, and not continually in which case the luminosity loss would be 1–2% smaller. Taking into account other residual aberrations, such as skew sextupole components and additional sources of aberrations, the estimated total loss of luminosity due to uncorrected tunable aberrations is about 13% (Table 11-9).

Ground Motion

Transverse quadrupole displacements due to ground motion may lead to a serious reduction of luminosity if the two beams are steered off collision faster than feedback systems are able to correct the IP beam position. There are several different approaches to treat the effect of ground motion on the final-focus system. Here, we base the analysis on a two-dimensional power spectrum following Juravlev and co-workers [Juravlev 1995]. A complete treatment of ground motion can be found in Appendix C.

If the vertical betatron phase advance from the entrance point e to the IP is a multiple of π, the offset of the two beams at the IP $\Delta y \equiv y_{\text{right}} - y_{\text{left}}$ due to an arbitrary vertical displacement $y(s)$ of the final-focus

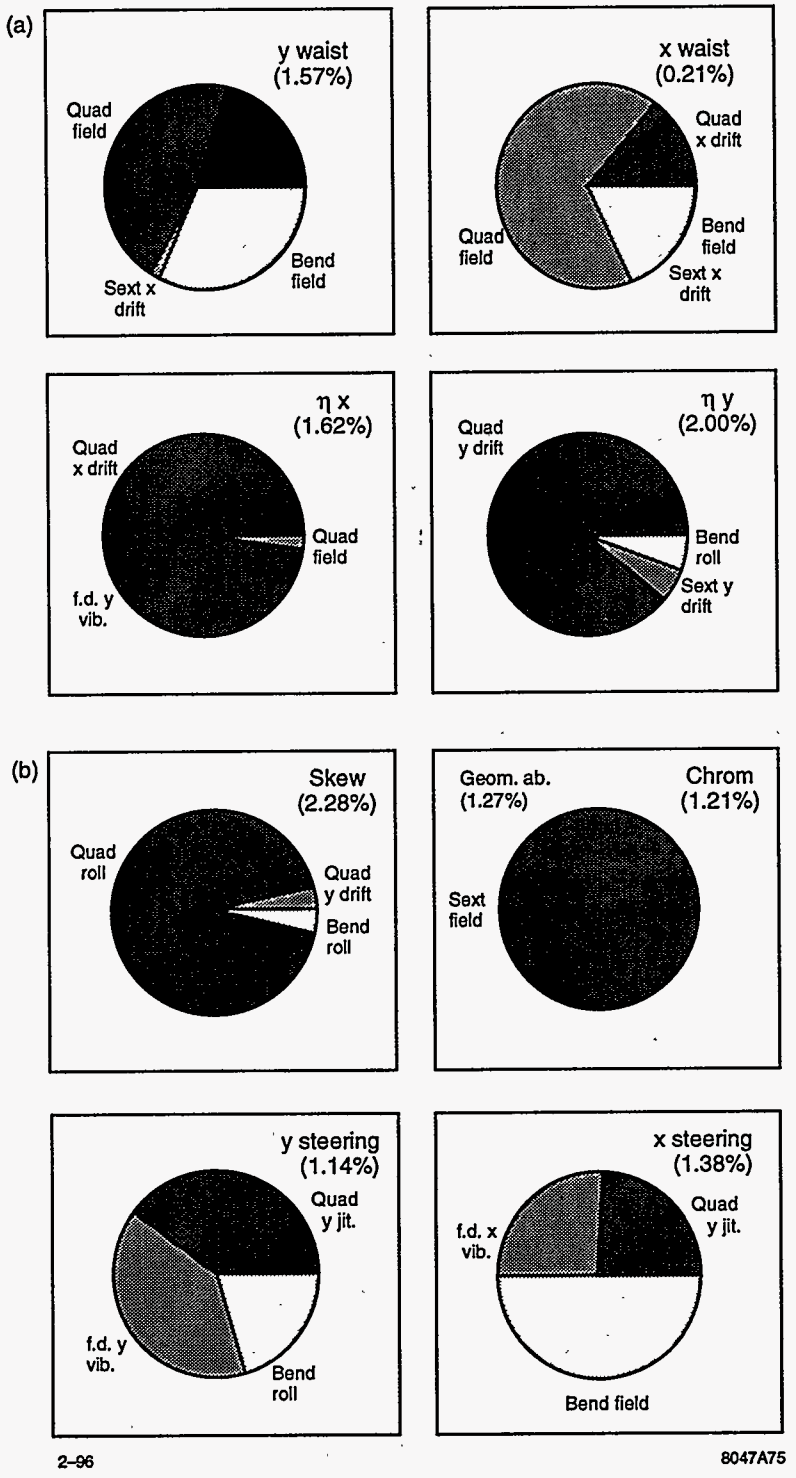


Figure 11-32. *Sources of important aberrations in the 1-TeV final-focus design, for the tolerances of Table 11-19. Numbers in parentheses refer to the overall luminosity loss due to a particular aberration. f.d.: final doublet; jit.: jitter; field: field ripple or drift.*

Type	P #	Radius a (mm)	Length L (m)	Drift (10^{-5}) $\Delta K/K$	Vibration[a] (nm) Δx	Δy	Drift[a] Δx (nm)	Δy (nm)	θ_{xy} (μr)
Quad 1	33	10.50	0.5/1.0	3.3	6.8	7.3	144.0	33.0[b]	11.3
Quad 2	4	10.50	0.5/0.5	1.8	54.0	4.8	15000.0	2000.0	15.7
Quad 3	3	10.50	1.0/1.0	1.6	9.0	3.4	480.0	21.0[b]	3.4
Quad 4	4	4.00	1.5/2.0	0.4	3.6	0.7	1600.0	98.0	1.8
Skew q. 1	4	6.00	0.4/0.4	100.0	50.0	30.0	15000.0	15000.0	500.0
Sext. 1	4	10.50	0.4/0.4	220.0	50.0	50.0	50.0	50.0	20.0
Sext. 2	12	10.50	0.4/0.4	900.0	1000.0	1000.0	10000.0	10000.0	20.0
Bend 1	100	—	5.0/16.0	0.3	100.0	100.0	500.0	500.0	0.5

[a] With respect to ground.

[b] The vertical drift tolerance will be looser at least by a factor of 10, if the dispersion at the IP is continually corrected utilizing the correlation of beam-beam deflection and beam energy.

Table 11-19. *Magnet types and tolerances for 1-TeV-c.m. energy. Tolerances result in the luminosity loss of Table 11-21.*

magnets is given by

$$\Delta y = -\sum_{i,\text{right}} k_i R_{34}^{i\to IP} y(s_i) + R_{33}^{e\to IP} y(s_{e,\text{right}}) + \sum_{i,\text{left}} k_i R_{34}^{i\to IP} y(s_i) - R_{33}^{e\to IP} y(s_{e,\text{left}}) \equiv \sum_j \mu_j y(s_j) \quad (11.41)$$

where $R_{34}^{i\to IP}$ denotes the (3,4) TRANSPORT-matrix element from the ith magnet to the IP, $R_{33}^{e\to IP}$ is the (3,3) matrix element from the entrance of the final focus to the IP, and k_i the integrated quadrupole strength of quadrupole i. For simplicity, we have replaced the $\pm k_i R_{34}$ and $\pm R_{33}$ by the generalized lattice parameters μ_j. The subindex i runs over only one side of the final focus system, while the subindex j runs over both sides. If we square the sum in Eq. 11.41, we will find mixed expressions of the form $y(s_i)y(s_j) \equiv y(s_i)y(s_i + \Delta s_{ji})$, whose expectation value over position s and over time t is given by

$$< y(s_i)y(s_j) >_{s,t} = \lim_{S\to\infty} \frac{1}{ST} \int_{-\frac{S}{2}}^{\frac{S}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} y^*(s_i,t) y(s_i + \Delta s_{ji}, t)\, ds\, dt \quad (11.42)$$

$$= \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty \frac{dk}{2\pi} P(\omega,k) \cos(k\Delta s_{ji}) \quad (11.43)$$

where $\Delta s_{ji} \boxminus s_j - s_i$. The term $P(\omega,k)$ denotes the two-dimensional power spectrum (regarding frequency and wave number) of the ground motion. The expectation value for the change of the beam-offset squared after time T can be written

$$< (\Delta y(t) - \Delta y(t+T))^2 >_t = 2 \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty \frac{dk}{2\pi} G(k) P(\omega,k)(1 - \cos(\omega T)) \quad . \quad (11.44)$$

The function $G(k)$ describes the response to a harmonic displacement of quadrupoles as a function of wavelength. It equals the average squared ratio of the beam-beam offset at the IP and the ground-motion amplitude at that wavelength, and is called the lattice response function. In terms of the lattice parameters μ_j and the positions s_j, this function can be written as:

$$G(k) \equiv \sum_{j1,j2} \mu_{j1}\mu_{j2} \cos(k(s_{j1} - s_{j2})) = 4 \left(\sum_j \mu_j \sin k s_j \right)^2 \quad (11.45)$$

Ripple:	
Quad1	Conventional
Quad2	10^{-5} commercially available, standard FFTB supply
Quad3	10^{-5} commercially available, standard FFTB supply
Quad4	Permanent magnet QA: SmCo with Erbium, see Sec. 11.6. Superconducting Q1: under design Q2 and Q3: if conventional, need special supply
Skew q1	Conventional power supply
Sext1	Conventional power supply
Sext2	Conventional power supply
Bend1	Special supply
Vibration:	
Quad1	8 nm, FFTB support: ground to quad diff. 1 nm
Quad2	5 nm, FFTB support: ground to quad diff. measured as 1 nm
Quad3	3.5 nm, FFTB support: ground to quad diff. 1 nm
Quad4	1 nm, optical anchor, plus seismometer on ground, to keep in line, See IR section
Skew q1	Conventional mount
Sext1	Conventional mount
Sext2	Conventional mount
Bend1	Conventional mount
Drift:	
All	Rf BPMs or orbit stabilization system. See Sec. 11.5.6. Thermal stabilization of tunnel ($<0.1\,\mu$m/hr achieved in FFTB)
Drift (roll):	
All	Rf BPMs or orbit stabilization system. See Sec. 11.5.6.

Table 11-20. *Tolerances for magnet types from Table 11-19 and how they can be achieved.*

The lattice response function $G(k)$ for the present NLC final-focus design is shown in Figure 11-33. For large wave numbers k (above 0.6 m^{-1}) the response function $G(k)$ is about constant, equal to ten. The oscillations are due to the discrete distances between the supports of the four final-doublet magnets. The calculation assumes that all magnets are supported beneath their center.

For small values of the wave number k, that is for $k < 0.01$ m^{-1}, $G(k)$ increases as the sixth power of k. There are three reasons for this: First and second, each side of the final-focus optics fulfills the Irwin sum rules [Irwin 1995a]:

$$-\sum_i k_i R_{34}^{i\to IP} + R_{33}^{e\to IP} = 1 \tag{11.46}$$

$$-\sum_i k_i R_{34}^{i\to IP} s_i + R_{33}^{e\to IP} s_e = 0, \tag{11.47}$$

where Eq. (11.46) expresses that a constant displacement of the entire system, including incoming beam orbit, has no effect on the beam-beam offset at the IP, and Eq. 11.47 means that a displacement of constant

Name	Tolerance	$\Delta L/L$ (%)
Quads	hor. vibr.	0.33
	vert. vibr.	0.50
	hor. stab.	1.00
	vert. stab.	1.00
	field ripple	0.50
	roll	1.10
	Total:	**4.23**
Final Doublet	hor. vibr.	0.36
	vert. vibr.	0.50
	hor. stab.	0.95
	vert. stab.	1.00
	field ripple	0.38
	roll	0.96
	Total:	**3.90**
Sextupoles	hor. vibr.	0.03
	vert. vibr.	0.05
	field ripple	2.40
	Total:	**2.50**
Bends	field ripple	0.69
	field stability	0.67
	roll vibration	0.24
	roll drift	0.13
	Total:	**1.71**
Overall	**Total:**	**10.9**

Table 11-21. *Overall luminosity loss corresponding to imposed magnet tolerances. Reference design luminosity (without pinch) is* 0.96×10^{34} cm^{-2} s^{-1}. *Numbers correspond to tolerances listed in Table 11-19.*

slope also does not affect the beam-beam offset. Note that, in Eq. 11.47, the betatron phase advance from the final-focus entrance e to the IP was assumed to be a multiple of π, but this assumption is not essential. The two conditions (11.46) and (11.47) imply at least a k^4 behavior at small k. In addition and third, the final-focus system consists of several paired π-modules, *i.e.*, of identical modules with phase advance π on either side of the IP. For large wavelengths, the effects of these modules cancel each other, giving rise to the observed k^6 asymptotics. In an intermediate wave number range, between 0.01 and 0.6 m^{-1}, the response function increases about linearly with k (Figure 11-33).

The frequency-dependent correlation of ground motion at different positions can be characterized by

$$R(\omega, \Delta s) \equiv \int_0^\infty \frac{dk}{2\pi} \mu(\omega, k)(1 - \cos k\Delta s) \tag{11.48}$$

where

$$\mu(\omega, k) \equiv \frac{P(\omega, k)}{P(\omega)}, \tag{11.49}$$

with

$$P(\omega) \equiv \int \frac{dk}{2\pi} P(\omega, k) \tag{11.50}$$

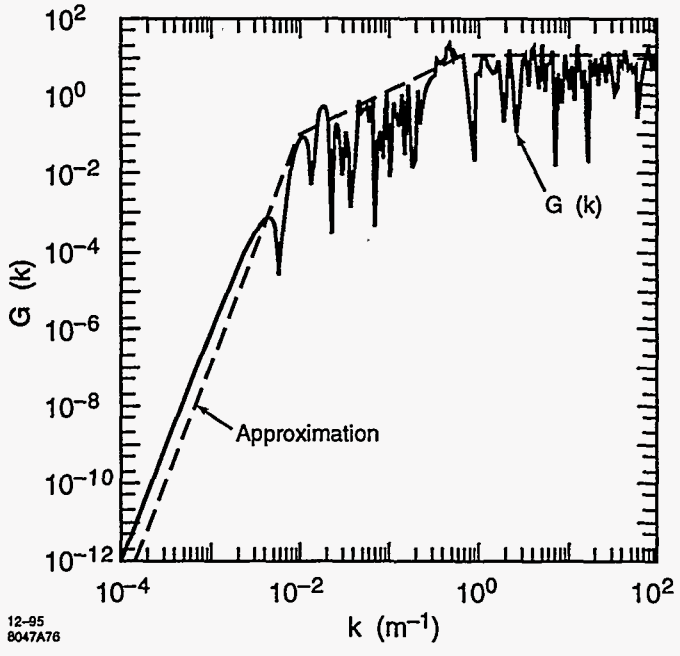


Figure 11-33. *Lattice response function for the NLC final-focus system.*

representing the local power density of the ground motion. The function $R(\omega, \Delta s)$ can be extracted from recent ground motion measurements at SLAC, performed with two STS-2 seismic sensors on the linac tunnel floor. The measurements suggest that R is parameterized quite well by the expression [Adolphsen 1995]

$$R(\omega, \Delta s) \approx 1 - J_0\left(k(\omega)\Delta s\right) \tag{11.51}$$

where J_0 denotes the zeroth-order Bessel function, $k(\omega) = \omega/v(\omega)$ with

$$v(\omega)\ [\,\mathrm{m\,s^{-1}}] \approx 450 + 1900 \exp\left(-\frac{\omega}{4\pi}\right) \tag{11.52}$$

and SI units are used throughout. The quantity $v(\omega)$ may be interpreted as the velocity of ground waves at frequency $f = \omega/(2\pi)$. The correlation function $R(\omega, \Delta s)$ is converted into the two-dimensional power spectrum $P(\omega, k)$ via Fourier transform:

$$P(\omega, k) = 4P(\omega) \times \int_0^\infty (1 - R(\omega, L)) \cos(kL) dL. \tag{11.53}$$

or

$$P(\omega, k) = \begin{cases} \frac{4}{\sqrt{k(\omega)^2 - k^2}} P(\omega) & \text{if} \quad k(\omega) > k \\ 0 & \text{else} \end{cases} \tag{11.54}$$

Interestingly, the same functional dependence is obtained if the ground motion is modeled by an ensemble of isotropic plane surface waves, for which each frequency is associated with a certain wavelength; longer wavelengths corresponding to lower frequencies [Irwin 1995b].

According to measurements at various locations [Juravlev 1995], a reasonable approximation to $P(\omega)$ for a "quiet" site is

$$P(\omega)[\,\mu\mathrm{m}^2/\mathrm{Hz}] \approx \frac{15.6 \cdot 10^{-3}}{\omega^4}. \tag{11.55}$$

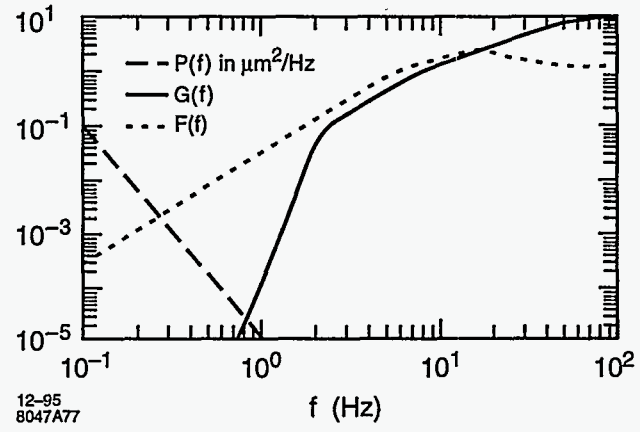


Figure 11-34. *Three functions which determine the rms beam-beam separation due to plane-wave ground motion: $F(f)$ — the feedback response [Hendrickson 1995];$P(f)$ — the local power density; $G(f) \equiv \tilde{G}(2\pi f)$—defined by Eq. 11.57. The integral over the product of these functions gives the square of the rms beam-beam separation.*

where the angular frequency ω is given in s^{-1}, and only frequencies $\omega > 0$ are considered. The power density $P(\omega)$ decreases rapidly at high frequencies, while at low frequencies the response to ground motion is suppressed by the lattice-response function $G(k)$ in conjunction with the dispersion relation, Eq. (11.54). A further reduction may be achieved by orbit-feedback systems.

The ground-motion spectral density, Eq. (11.54), determines the rms separation $\Delta y_{\rm rms}$ of the two beams at the IP, as caused by the plane-wave ground motion. The rms separation can be written as

$$\Delta y_{\rm rms}^2 = \int_0^\infty \frac{d\omega}{2\pi} \tilde{G}(\omega) P(\omega) F(\omega), \tag{11.56}$$

where $F(\omega)$ denotes the feedback response, $P(\omega)$ is the power density, Eq. (11.55), and $\tilde{G}(\omega)$ represents the lattice response function converted into the frequency domain. It is defined as an integral of the product of $G(k)$ and the k-dependent part of $P(\omega, k)$ over k, *i.e.*,

$$\tilde{G}(\omega) \equiv \int_0^{k(\omega)} G(k) \frac{4}{\sqrt{k(\omega)^2 - k^2}} \frac{1}{2\pi}\, dk \quad . \tag{11.57}$$

The three functions $P(\omega)$, $\tilde{G}(\omega)$, and $F(\omega)$ are illustrated in Figure 11-34. The shape of $\tilde{G}(\omega)$ shows that the effect of ground motion is strongly suppressed by the lattice response. Numerical integration yields a value for the rms beam-beam separation of

$$\Delta y_{\rm rms} \approx 0.2 \text{ nm} \qquad \text{(with feedback)} \tag{11.58}$$

$$\Delta y_{\rm rms} \approx 0.3 \text{ nm} \qquad \text{(without feedback)} \tag{11.59}$$

and, thus, the plane-wave ground motion is insignificant, provided the magnet centers move exactly as the ground beneath them. If plane-wave ground motion were the only source of orbit-perturbation, no orbit-feedback would be needed at a quiet site. Above 6 Hz, the power spectrum of the SLAC measurements is considerably larger than that of Eq. (11.55), due to resonances of the linac support structures and also due to cultural noise. Assuming the SLAC power spectrum, one finds an expected rms beam-beam separation of

about 1.1–1.3 nm, which still seems acceptable, but which is considerably larger than the 0.3 nm for a quiet site.

Some authors have argued that at low frequencies a component of ground motion exists which cannot be cast into the above framework, and which is of pure random character [Juravlev 1995]. They characterize this part of the ground motion by a so-called ATL law, according to which the change of the average squared relative displacement of two points is proportional to the distance between the points and to time. (An entirely different interpretation of ground motion as a systematic process was suggested by R. Pitthan [Pitthan 1995].) The two-dimensional spectral density describing the ATL law can be written

$$P_{ATL}(\omega, k) = \frac{4A}{\omega^2 k^2} \tag{11.60}$$

where $k > 0$, $\omega > 0$ is assumed, and $A \approx 10^{-8}$–10^{-5} $\mu m^2\ s^{-1}\ m^{-1}$ depending on location and on relevant timescale. The SLAC measurements over a timescale of seconds indicate $A < 6 \times 10^{-7}$ $\mu m^2\ s^{-1}\ m^{-1}$.

Folding in the feedback response function, it is possible to determine the rms beam-beam separation due to ATL-like ground motion in the frequency range 0–0.01 Hz, where the ATL law might be applicable. For $A = 10^{-6}$ $\mu m^2\ m^{-1}\ s^{-1}$, this rms separation is found to be 15 pm. Assuming as a worst case, that the size of uncorrelated motion is equal to the noise floor of the STS seismometer, the rms separation in the frequency range of 0.01–6 Hz is estimated to be no larger than 242 pm, and the contribution from frequencies above 6 Hz is about 124 pm at a quiet site. In total then, the rms separation from uncorrelated ATL-like ground motion does not exceed 0.3 nm, and the entire separation due to all ground motion (ATL-like and plane-wave motion) at a quiet site is about 0.4 nm, which would result in a luminosity loss of only 0.1%. The conclusion is that one can use the ground (bedrock) as a reference for stabilization.

Thus, magnet supports need to be designed which neither amplify nor damp the ground motion, but couple the magnet firmly to the ground. As discussed previously, the tightest tolerances on uncorrelated quadrupole-to-ground vibration above 5 Hz (vibration at lower frequency is compensated by the orbit feedback) are 1 nm (rms) for the final doublet and 3–4 nm for a few other quadrupoles (Table 11-19). These tolerances correspond to a total luminosity loss of about 2%. At DESY (SBLC-TF), relative rms vibrations above 1 Hz were measured to be smaller than 1 nm, which would meet all the NLC tolerances. Quadrupoles at the FFTB were found to vibrate by about 4 nm with respect to the ground, excited mainly by bad cooling pumps. These vibration amplitudes would still satisfy the tolerance criteria for all NLC magnets other than the final doublet. For the latter a special stabilization system based on an optical anchor and piezo-electric movers has been devised, which is described in Chapter 12.

Capture Tolerances and Tuning Frequency

The magnet pre-alignment tolerances required to steer the first beam through the final focus onto the dump will be loose. They are essentially determined by the magnet-mover ranges. A pre-alignment with an accuracy of 500 μm is expected to be sufficient. This is easily achieved by standard alignment techniques, which provide a typical accuracy of 100–200 μm. Once the beam reaches the NLC main dump, the actual alignment is performed using beam-based techniques, similar to those employed at the SLC or FFTB [Emma 1992, Raimondi 1993, Tenenbaum 1995].

More interesting is the question how often the beam-based alignment has to be performed, and how accurate it should be. The answer to this depends on the maximum range of the tuning elements, and on the IP spot-size increase due to ground motion. The recent progress in understanding ground motion (see previous section and Appendix C), and new work on tuning schemes shed light on this problem. A related question refers to the required frequency of aberration tuning and whether it is acceptable.

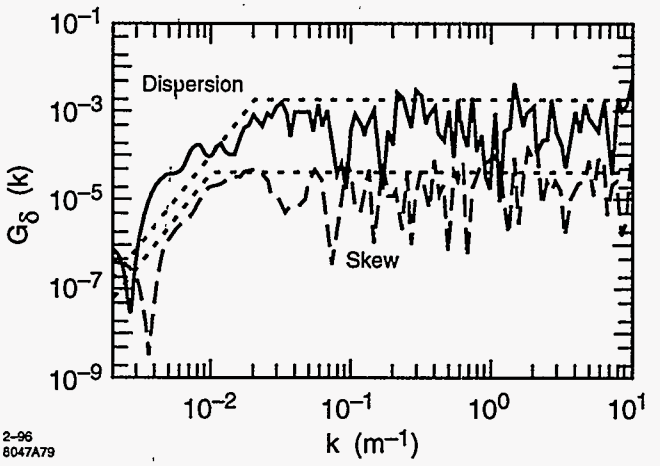


Figure 11-35. *Lattice response functions for vertical dispersion and skew coupling. Dotted lines represent approximations which are employed in the later analysis.*

In complete analogy to the steering effect of ground motion, lattice-response functions may be defined for vertical dispersion and skew coupling, in an obvious way. The only difference is that, in this case, the two final foci on either side of the IP are treated independently. More precisely, the lattice-response functions are of the form

$$G_\delta(k) = \sum_{i,j} \mu_i^\delta \mu_j^\delta \cos(k\Delta s_{ij}) \tag{11.61}$$

where the subindices i, j run over only one side of the IP only, and the multipliers $\mu_{i,j}^\delta$ characterize the IP spot-size increase due to a displacement of quadrupole (or sextupole) i. The term Δs_{ij} denotes, as before, the distance between the two magnets i and j. The lattice response functions for dispersion and skew coupling are depicted in Figure 11-35, along with simple approximations which are used in the following studies. If the wavelengths are small, *i.e.*, at large k, the response is about constant, since the magnets move incoherently. For small k, the lattice response is strongly suppressed and decreases as the 4th power of k. As for the steering effect discussed earlier, this suppression represents the invariance of the luminosity with regard to displacements and tilts of the entire beam line. Unlike the steering, however, there is no cancellation between the two sides of the IP or between paired $-I$ modules, and this is the reason why the response here decreases as the 4th power of k, instead of k^6.

The spot-size increase due to ATL ground motion as a function of time T is described by the equation

$$< \Delta\sigma_y(T)^2 >= 2 \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi} G_\delta(k) P_{ATL}(\omega, k)(1 - \cos(\omega T)) \tag{11.62}$$

where $G_\delta(k)$ denotes the lattice-response function for dispersion or skew coupling (Figure 11-35). Figure 11-36 presents the increase of the spot size due to dispersion as a function of time, for three different magnitudes of ATL-type ground motion.

Even though this figure does not assume any orbit stabilization other than a beam-centroid correction at the IP, there would still be several minutes time to correct the dispersion. In reality, feedback systems which stabilize the orbit will reduce the generated dispersion. If the frequency-response curve of a typical SLC-style orbit feedback is included in the calculation, an ATL coefficient $A = 10^{-6}$ μm^2 m^{-1} s^{-1} is assumed, and the product spectrum is integrated over frequencies below 0.01 Hz, the spot size is found to reach a maximum value after about 100s and then to stay constant. The spot-size increase at that moment is less

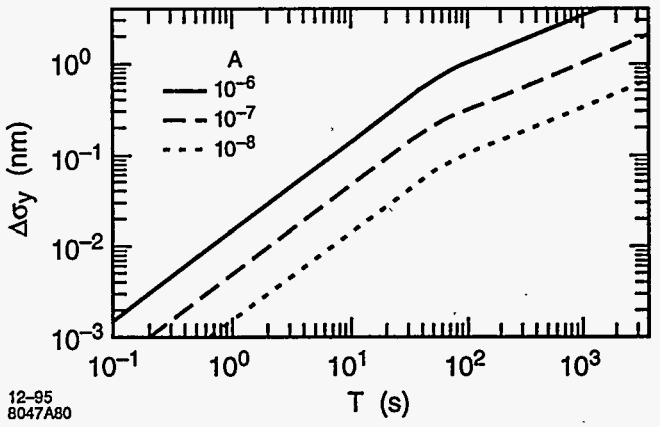


Figure 11-36. *Incoherent spot-size increase due to vertical dispersion, caused by ATL ground motion, as a function of time, for three different values of A (in units of $\mu m^2\ s^{-1}\ m^{-1}$), and assuming $\delta \approx 3 \times 10^{-3}$.*

than a picometer, to be added in quadrature. Thus the effect of ATL-like ground motion on the IP spot size appears to be insignificant, assuming the orbit is stabilized by a typical SLC-style feedback system. The frequency at which the different aberrations need to be tuned, in order to keep the average luminosity loss per aberration below 1–2%, is thus not determined by the ground motion, but by the stability of the feedback BPMs. If the BPM stability were the same as measured in the FFTB, the aberrations would have to be tuned less than once per hour. Since the NLC BPMs will be located underground, and temperature and humidity much better controlled, the BPM stability is expected to surpass that in the FFTB.

Collimator Tolerances

As a supplement to the two-phase collimation section upstream of the big bend, a last collimation of the final-doublet betatron phase is performed in the final focus itself. Its purpose is to reduce the number of electrons which may hit the final doublet to below ten, so that the detector background remains acceptable. The final-focus collimators remove the beam halo which is newly generated by gas scattering, bremsstrahlung and inverse Compton scattering downstream of the collimation section and by edge scattering in the IP-phase collimation. The ideal position for final-focus collimators are the high-beta points in CCX and CCY. Therefore, the two horizontal and two vertical collimators will be located close to the main horizontal and vertical sextupoles, respectively.

In addition to these four tungsten collimators, the final focus accommodates almost 200 wide-aperture titanium spoilers. These are placed upstream of each quadrupole and each dipole section, and protect the magnets against damage by a missteered beam.

Collimators do not only improve the background, but, as SLC experiments have demonstrated [Bane 1995], they can also adversely affect the beam quality by generating resistive-wall and geometric wakefields. To reduce the geometric wake, most or all of the NLC collimators will be tapered at a shallow angle, as discussed in Chapter 9, Adjustable collimators are usually flat and consist of two jaws (top and bottom, or left and right). Protection collimators are round.

Specifically, there are four main effects of collimators:

- If the incoming beam executes a betatron oscillation, the collimator wakes will cause an amplification of the oscillation in the orthogonal betatron phase. This effect is called jitter amplification, and occurs even when the collimators are perfectly aligned. If the collimator is located in the final-doublet phase, the position change at the IP Δy caused by the collimator is

$$\frac{\Delta y}{\sigma_y^*} = \xi \frac{y'}{\sigma_{y'}} \tag{11.63}$$

for an initial betatron oscillation of size y', where all coordinates and rms values refer to the IP. The coefficient ξ is the relative jitter amplification, and is expressed in terms of collimator and beam parameters as

$$\xi \equiv \frac{\beta_y^{\rm coll} 2Nr_e}{\gamma\sigma_z}\left[\left(8 < f_r >< g_r > \frac{(b^2-g^2)(b-g)}{b^3}\right)^{\frac{1}{2}}\left(\frac{\lambda\sigma_z}{g^6}\right)^{\frac{1}{4}} + \frac{L}{g^3} < f_r > \sqrt{\lambda\sigma_z}\right] \tag{11.64}$$

where $< f_r > \approx 0.3$, $< g_r > \approx 0.82$, g is the half-aperture, b the beam-pipe radius, $\lambda = \rho/(120\pi)$ (ρ resistivity in Ωm), and L the length of the flat part of the collimator.

- In addition to jitter amplification, there is also an emittance growth arising from the variation of the wakefield kick across the bunch. Typically, this effect is smaller than the former.
- If the collimator jaws (of a flat collimator) are misaligned with respect to the center of the beam, they will also cause emittance growth, even if the incoming orbit is unperturbed. This emittance growth gives rise to an alignment tolerance. For a spot size increase by 2% or less the alignment tolerance of a single collimator is

$$\Delta y \approx \frac{\sigma_y^{\rm coll}}{4.5\,\xi} \quad . \tag{11.65}$$

- Finally, even if a flat collimator is perfectly aligned, and the beam unperturbed, the collimator still delivers a quadrupole wake-field to the beam, changing the focusing of particles in the bunch-tail halo and causing an emittance dilution. The magnitude of the quadrupole wakefield kick from a flat collimator is about a third of that of the dipole wake. Thus, for an expected orbit jitter larger than $0.5\,\sigma_y$, the quadrupole effect is small compared with the dipole kick.

The main horizontal and vertical final-focus collimators are located at the high beta points in CCX and CCY, respectively, where $\beta_x \approx 70$ km or $\beta_y \approx 200$ km. If we assume $N = 10^{10}$ particles per bunch, and $\sigma_z = 100\,\mu$m, the jitter amplification ξ for a 7-cm-long (about 20 r.l.) W collimator, and collimation at 17 σ_x or 43 σ_y, is 0.06 (0.01) in $y(x)$. The vertical alignment tolerance, Eq. (11.65), for a 2% spot-size increase caused by the two vertical collimators is 1.6 mm; the horizontal alignment tolerance for the CCX collimators is even larger. Thus, the required alignment accuracy will not be determined by emittance growth, but by the collimation efficiency. The latter probably demands that the vertical collimators are aligned to within $150\,\mu$m ($1\sigma_y$).

Temperature

The tolerances on position and field strength translate into tolerances on the temperature stability of magnets and magnet supports. The tolerances on medium-term stability are considerably relieved by the existence of an orbit stabilization system (*i.e.*, feedback) which corrects orbit distortions resulting from quadrupole drifts, bend-field changes and bend rolls.

In the FFTB, quadrupole motion due to temperature variation was measured to be smaller than 100 nm/h. This stability is an order of magnitude better than what is needed for the NLC.

Vacuum

The required vacuum pressure is determined by the tolerable background in the detector. The following discussion shows that the largest contributions to the vacuum-related background are due to bremsstrahlung. At an average pressure of 10 nTorr, less than 10 particles per bunch train will hit apertures in the final-doublet region. The background from these particles needs to be compared with that expected from synchrotron radiation generated in the bending and quadrupole magnets, and with that caused by beamstrahlung and pair-creation at the IP.

Coulomb Scattering. If an electron undergoes hard Coulomb scattering on a residual-gas nucleus, it may hit the pole faces or inner beam pipe of the final doublet. Following Irwin (Chapter 9), the relative number of scattered particles hitting the doublet is given by

$$\frac{\Delta N}{N} = 2\pi\rho\left(\frac{r_e}{\gamma a_D}\right)^2\left(\sum_i Z_i^2\right)\left[\int_{SX1b}^{QFTA} ds\, R_{12}^2(s) + \int_{SX2b}^{QFTA} ds\, R_{34}^2(s)\right] \tag{11.66}$$

where $R_{12}(s)$ and $R_{34}(s)$ denote the R-matrix elements from point s to the entrance face of QFTA, r_e the classical electron radius, and a_D the inner radius of QFTA (assumed as 5 mm). The integrations are performed between the last sextupole in CCX and CCY, respectively,—these are the collimator positions—and QFTA. The sum over i runs over the different atoms of a molecule, and $\rho \approx 3.2\times 10^{13}\mathrm{m}^{-3}p/\mathrm{nTorr}$ is the molecule density. Evaluation yields

$$\left[\int_{SX1b}^{QFTA} ds\, R_{12}^2(s) + \int_{SX2b}^{QFTA} ds\, R_{34}^2(s)\right] \approx 7\times 10^6\ \mathrm{m}^3. \tag{11.67}$$

Assuming carbon monoxide molecules and a beam energy of 500 GeV, we find

$$\frac{\Delta N}{N} \approx 4\times 10^{-14}\frac{p}{\mathrm{nTorr}}\ . \tag{11.68}$$

If we allow 10 electrons in each bunch train of 6×10^{11} particles to hit the doublet in the absence of final-focus collimators, a pressure of 400 nTorr is sufficient. The collimators in the final-focus system will intercept most of the scattered particles.

Bremsstrahlung. An electron may also get lost due to bremsstrahlung in the field of a nucleus. The cross section for an energy loss between ΔE_1 and ΔE_2 is [Piwinski 1985]

$$\sigma_{\mathrm{brems}} \approx \sum_i \frac{4r_E^2 Z_i^2}{137}\frac{4}{3}\ln\left(\frac{\Delta E_2}{\Delta E_1}\right)\ln\left(\frac{183}{Z_i^{\frac{1}{3}}}\right)\ . \tag{11.69}$$

For CO molecules and an energy loss between 6% and 25%, we find $\sigma_{\mathrm{brems}} \approx 2$ barn. The fraction of particles suffering this energy loss is given by

$$\frac{\Delta N}{N} = \rho L\sigma_{\mathrm{brems}} \approx 6.4\times 10^{-12}\frac{p}{\mathrm{nTorr}} \tag{11.70}$$

where $L \approx 1000$ m. In order that less than 10 particles per train lose 6–25% of their energy, we need to achieve a pressure of 4 nTorr. This value is two orders of magnitude smaller than that calculated before for particle loss due to hard Coulomb scattering. However, a large fraction of the electrons lost by bremsstrahlung do not hit the final-doublet apertures and will not contribute to the detector background. As a conservative estimate, we assume a pressure of 10 nTorr to be sufficient.

Inelastic Scattering. For simplicity, let us confine ourselves to hydrogen molecules, *i.e.,* we consider electron-proton collisions. In this case, using the Weizsäcker-Williams approximation, the differential cross section for leptoproduction $d\sigma_{ep}/dy$ reads [ZEUS/H1]

$$\frac{d\sigma_{ep}}{dy} = \sigma_{tot}^{\gamma p}(W_{\gamma p})\frac{\alpha}{2\pi}\frac{1+(1-y)^2}{y}\ln\frac{Q^2_{\max}}{Q^2_{\min}} \tag{11.71}$$

where $y \equiv 1 - E'/E$ is the photon energy in units of the incident electron energy. Approximating $\sigma_{tot}^{\gamma p} \approx 150\ \mu$barn, $Q^2_{\max} \approx 4E^2$, $Q^2_{min} = (m_e c^2 y)^2/(1-y)$, and integrating over photon energies from 500 keV to 250 GeV, we derive an upper bound for the relevant total cross section:

$$\sigma_{ep} \leq 180\ \mu\text{barn} \quad . \tag{11.72}$$

Over a distance l of 100 m, the fraction of electrons undergoing inelastic reactions with the residual gas is then

$$\frac{\Delta N}{N} \approx \sigma_{ep} l \rho \leq 6 \times 10^{-17}\frac{p}{\text{nTorr}} \tag{11.73}$$

or $\Delta N \leq 0.003$ per bunch train for a pressure of 100 nTorr—a very small number.

Other Sources of Background

Inverse Compton Scattering. Electrons (or positrons) do not only scatter off residual-gas atoms, but can also suffer inverse Compton scattering on thermal radiation photons. The latter effect can limit the beam lifetime in high-energy electron storage rings [Telnov 1987], and it has been measured at LEP [Bini 1991] and at HERA [Lomperski 1993]. In this section, we estimate its importance for the NLC final-focus system.

The density of thermal photons increases as the third power of the temperature. For a temperature of 300 K, it is about

$$n_\gamma \approx 5 \times 10^{14}\ \text{m}^{-3}, \tag{11.74}$$

and thus equal to the density of residual-gas atoms at a pressure of 17 nTorr. The total Compton cross section can be expressed in terms of the dimensionless parameter x [Telnov 1987].

$$x \equiv \frac{4E\omega_0}{m_e^2 c^4}\cos^2\frac{\alpha_0}{2}, \tag{11.75}$$

where α_0 denotes the angle at which photons and electrons collide in the laboratory frame ($\alpha_0 = 0$ means head-on collision), E the beam energy, and ω_0 the photon energy. Assuming 500 GeV beam energy, an average photon energy of $\bar{\omega}_0 = 0.07$ eV and $\alpha_0 = 0$, we find $x \approx 0.56$, and a total Compton cross section not much different from the Thomson cross section $\sigma_t \approx 0.7$ barn. This cross section is an order of magnitude smaller than that for beam loss due to bremsstrahlung on residual-gas nuclei. However, the maximum energy loss of a Compton-scattered electron is quite large,

$$\left.\frac{\Delta E}{E}\right|_{\max} = \frac{x}{1+x} \approx 36\%, \tag{11.76}$$

which implies that almost all scattered electrons are lost so rapidly, that they do not make it to the final doublet. The total number of particles lost per bunch train is estimated as

$$\Delta N \approx \bar{\sigma}_T L n_\gamma N \approx 36. \tag{11.77}$$

Here $N \approx 6 \times 10^{11}$ is the number of particles in the bunch train, and $L \approx 2000$ m the length of the final-focus system. The number of electrons suffering inverse Compton scattering on thermal photons is at least comparable to, if not larger than those lost due to bremsstrahlung or Coulomb scattering off the residual gas.

Touschek Effect. The scattering of two particles inside the same bunch off each other may lead to a particle loss due to the introduced change of longitudinal momentum. This effect limits the beam lifetime in many electron storage rings. It was first observed at AdA, in 1963 [Bernadini 1963].

The number of electrons scattered to energies larger than $\eta \equiv \left(\frac{\Delta E}{E}\right)_{\max}$ is given by [Walker 1987]

$$\Delta N = \frac{r_e^2 n_b N_b^2 L}{\gamma^2 \left.\frac{\Delta E}{E}\right|_{\max} 2\pi^{\frac{1}{2}} \sigma_x \sigma_y \sigma_z} J(\eta, \delta q) \quad . \tag{11.78}$$

Here, δq denotes the rms transverse momentum in units of $m_e c$, r_e the classical electron radius, n_b the number of bunches, N_b the number of particles per bunch, and L the length of the beam line considered. The loss rate is proportional to the second power of N_b, and inversely proportional to the square of the energy and to the beam size. An approximation to the function $J(\eta, \delta q)$ for $\delta q/\eta \geq 100$ is given in Eq. (7) of ref. [Walker 1987]. For the parameters of interest here, J is nearly equal to 1.

In the NLC final focus at 500-GeV-c.m. energy, transverse beam size and transverse momentum are $< \delta q >= 6.4\ \mu\text{rad} \times \gamma \approx 3.2$ and $< \sigma_x \sigma_y > \approx 148\ \mu\text{m}^2$. Assuming $N_b \approx 8.5 \times 10^9$ and $n_b = 90$, the number of scattered particles per bunch train is:

$$\Delta N \approx 2 \text{ for } \Delta E/E \geq 0.5\%, \qquad \text{and} \qquad \Delta N \approx 0.1 \text{ for } \Delta E/E \geq 2\% \tag{11.79}$$

This is a rather small effect.

11.5.6 Feedback and Stability

Overview of Feedback Systems

Several orbit feedback systems are envisioned to stabilize the beam orbit throughout the beam delivery system. In addition to feedbacks in the collimation section, the big bend, and the diagnostics section, there will be a "launch" feedback correcting the orbit in the beta-match section and CCX, and a feedback controlling the CCY and FT orbit. Each feedback system will use orbit readings from at least four BPMs to correct the orbit by moving quadrupoles of by adjusting dipole correctors. The CCY- and FT-orbit feedback, which is the most important of these feedback loops, is discussed in the next section.

A fast IP collision feedback is essential for correcting beam position and preserving collisions at the IP. This feedback will be very similar to the IP feedback at the SLC, which uses fast dither coils to maintain head-on collisions. The required dither changes are deduced from the measured beam-beam kick angle and the slope of the beam-beam deflection curve. The main difference to the SLC is that, because of the smallness of the required corrector strengths, the NLC feedback will use electrostatic vertical dither coils rather than magnetic. In addition, it may be possible to use a device similar to the crab cavity for compensating orbit changes along the bunch train, should that be desired.

Regarding control of emittance or spot size, a slow automatic skew-correction feedback is being contemplated, which may be based on DS wire scans and SCS skew correctors. Such a feedback will only be necessary if there are significant changes to the incoming x-y coupling over a few hours, and it would be straightforward to implement.

CCY and FT Orbit Feedback

The IP beam size is extremely sensitive to orbit variations which originate within the final focus. For example, a vertical displacement of QFT3 (see location on Figure 11-37; note that the design and the magnet names have changed since this was written and drawn!) by 44 nm kicks the vertical orbit in QFT1 by ~200 nm which, for 0.3% rms energy spread, generates enough vertical dispersion at the IP to increase the beam size by 2%. A similar effect occurs for quadrupole displacements within the CCY where the orbit is steered off in the second CCY sextupole (SY1B). For these reasons it will be important to monitor and correct local orbit variations in the CCY and FT to a very high precision. The following describes a conceptual monitor and feedback system for CCY and FT local orbit distortions. Throughout, it is assumed that the IP beam-beam steering errors introduced by these orbit variations are corrected with a much faster (~30 pulse) IP collision feedback which is not addressed here.

The vertical beam size at the final doublet is quite large in comparison to the required orbit monitoring precision (60 μm compared with 200 nm). Therefore, local orbit distortion monitoring will be made in the presence of a large background signal of incoming trajectory jitter—which has much less impact on the IP beam size. The feedback system must suppress this incoming jitter signal reliably so that non-jitter related y-BPM reading variations of ~200 nm are discernible. In order to accomplish this the feedback system must monitor BPMs in the CCX, CCY and FT. In the simplest case, two BPMs are placed in the CCX—each one adjacent to an SX1 sextupole (SX1A and SX1B)—in order to monitor both betatron and energy jitter (BPM-4 and BPM-3). A third BPM is placed in the CCY adjacent to the SY1B sextupole (BPM-2) to monitor CCY induced orbit changes, and a fourth BPM is placed just upstream of QFT1 (BPM-1) to monitor FT induced orbit changes. More BPMs may be added for redundancy.

A simplifying characteristic of the final focus is that only the IP angle betatron phase is visible with BPMs throughout the CCX, CCY and FT—a one-sigma IP angle oscillation generates a 60-μm BPM-1 reading while a one-sigma IP position oscillation generates a ~3-nm BPM-1 reading. Figure 11-37 shows a one-sigma betatron oscillation at the IP angle phase in x and y, while Figure 11-38 shows a one-sigma oscillation at the IP position phase. The invisibility of the IP position betatron oscillation is clearly demonstrated.

For this reason each BPM reading can be broken down into a component of incoming IP angle jitter, energy jitter (for horizontal plane or vertical dispersion error), a CCY orbit kick, an FT orbit kick and a static offset. All other sources of BPM reading variations need to be <200 nm (BPM-1) over the scale of a few hours (time between re-tuning). A simple algorithm can then be devised where the four (or more) BPMs are used in a linear combination to determine the CCY- and FT-induced orbit changes so that corrections may be applied in both the CCY and the FT. The linear combination is determined by using the normal trajectory jitter to calibrate the BPM to BPM coefficients (assuming local orbit distortions are constant over the ~200 pulses necessary for calibration). Each ith BPM is then fitted to the form

$$y_i(j) = \eta_i \cdot \delta(j) + \alpha_i \cdot y_\beta(j) + c_i \quad , \tag{11.80}$$

where $\delta(j)$ is the fractional energy deviation of the jth pulse, y_β is the incoming betatron amplitude (proportional to IP angle), and the three coefficients, α_i, η_i and c_i are the fit results. The energy deviation and betatron amplitude can then, given their $-I_{4\times4}$ transfer matrix separation, be conveniently parameterized as follows (actually any linear combination will suffice),

$$\delta(j) \equiv \frac{y_4(j) + y_3(j)}{R_{16}} \quad , y_\beta(j) \equiv \frac{y_4(j) - y_3(j)}{2} \quad . \tag{11.81}$$

where R_{16} (~150 mm) is a linear transfer matrix element between BPM-4 and 3.

This calibration is then performed after each IP re-tune. After a time necessary to generate local orbit distortions (must not be less than ~10 seconds) a calculation is made using ~200 more pulses and the same

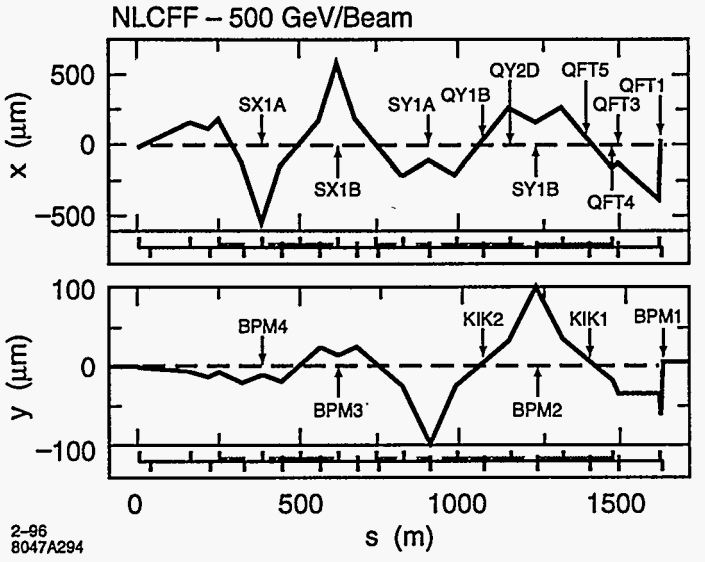


Figure 11-37. *Horizontal and vertical one-sigma IP angle betatron oscillation through the final focus (i.e., at the IP: $\Delta x = 0$, $\Delta y = 0$, $\Delta x' = \sigma'_x$, $\Delta y' = \sigma'_y$).*

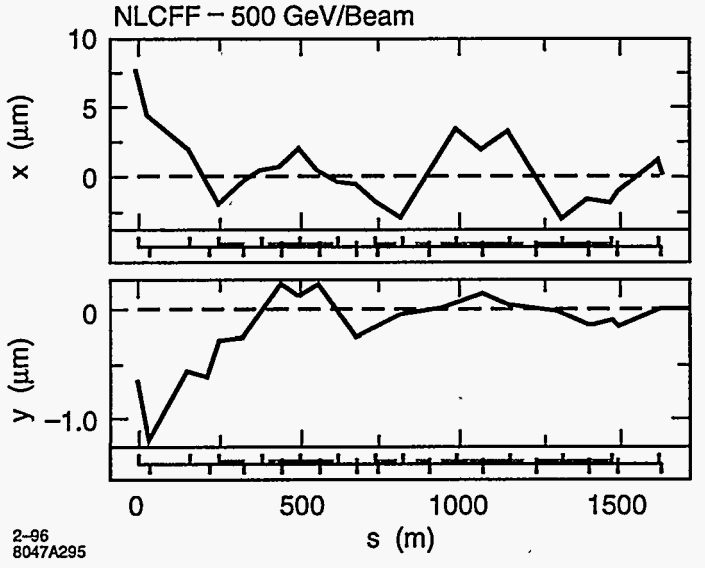


Figure 11-38. *Horizontal and vertical one-sigma IP position betatron oscillation through the final focus (i.e., at the IP: $\Delta x = \sigma_x$, $\Delta y = \sigma_y$, $\Delta x' = 0$, $\Delta y' = 0$).*

Number of pulses (cal. & cor.)	IP pos. rms jitter ($\sigma^*_{x,y}$)	IP ang. rms jitter ($\sigma^*_{x',y'}$)	Energy rms jitter (σ_δ)	BPM rms noise (μm)	BPM rms gain errors (%)	BPM non-linearity (% at 1 mm)	BPM rms offsets (μm)
200	0.5	0.5	0.25	1	5	2	250

Table 11-22. *Simulation conditions for CCY and FT local orbit distortion feedback system. With these conditions the feedback simulation successfully restored the IP beam size to within 2% for two CCY and two FT quadrupoles drifting 5 times their tolerance in x and y.*

coefficients.

$$\begin{bmatrix} y_4 - c_4 \\ y_3 - c_3 \\ y_2 - c_2 \\ y_1 - c_1 \end{bmatrix} = \begin{bmatrix} \eta_4 & \alpha_4 & 0 & 0 \\ \eta_3 & \alpha_3 & 0 & 0 \\ \eta_2 & \alpha_2 & 1 & 0 \\ \eta_1 & \alpha_1 & \alpha_1/\alpha_2 & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta \\ y_\beta \\ \Delta y_2 \\ \Delta y_1 \end{bmatrix} \tag{11.82}$$

Here the previously determined coefficients—assumed constant over the hour timescale—make up the matrix and the critical orbit distortion results, Δ_{y2} and Δ_{y1}, represent the local orbit distortions at BPM-2 (the second CCY sextupole SY1) and BPM-1 (the last quadrupole QFT1). This treatment assumes that all local orbit distortions are along the IP angle phase. This is generally true for the quadrupoles with the tightest tolerances (see Table 11-14). Corrections are then applied in order to maintain the CCY and FT orbit with respect to the incoming orbit. This is the only point in the algorithm where some knowledge of short sections of the optics is required. Static optical errors and BPM gain errors have been removed in the calibration process.

Simple simulations of this feedback have been run which include random BPM noise, IP position and angle jitter, energy jitter, BPM gain errors, BPM non-linearities and BPM offsets. Table 11-22 summarizes the simulation conditions.

For the simulation, two CCY quadrupoles and two FT quadrupoles (QY1B, QY2D, QFT4 and QFT3) were each displaced by 1 μm which, for QFT3, is >20 times tolerance. In the case of adjacent and opposite strength quadrupoles such as QFT3 and QFT4 the misalignments chosen were the worst case combination (*i.e.*, QFT3 $\Delta_{x,y} = +1\,\mu$m and QFT4 $\Delta_{x,y} = -1\,\mu$m). These sudden, uncompensated misalignments dramatically reduce the luminosity by a factor of ~10 (assuming this drift occurs in only one of the two final-focus beam lines). In this case the vertical IP beam size increases to ~13 times larger than nominal (2.5 nm) while the horizontal is only 6% larger. For the simulation, correctors were placed with one x and y pair (KIK2) at QY1B and one pair (KIK1) at QFT5 (each $\sim n\pi$ from the IP). Only one correction interval was applied (*i.e.*, no iterations were allowed). For these conditions the feedback successfully restored the luminosity to within 1% of nominal. Since there are only two correctors per plane and four misaligned quadrupoles the orbit is not everywhere restored. It is restored primarily at QFT1 (BPM-1) and SY1B (BPM-2) which, for this drift magnitude, is adequate. Figure 11-39 shows the difference orbit through the final focus, including the effect of the four displaced quadrupoles, before feedback correction is applied. Figure 11-40 shows the orbit after feedback correction. The horizontal feedback for this case is hardly necessary. However, further simulations have shown the horizontal feedback to work well at scales of up to ~10 times the horizontal drift tolerances for these quadrupoles (10–30 μm).

The corrector strengths here are too small to be reliably applied by a simple dipole magnet (~0.2 Gauss for a 10-cm length dipole). However, at this timescale (>10 seconds) a quadrupole mover could be used. In this case a 19-μm and 1.4-μm vertical displacement of QFT5 and QY1B, respectively will provide the necessary correction.

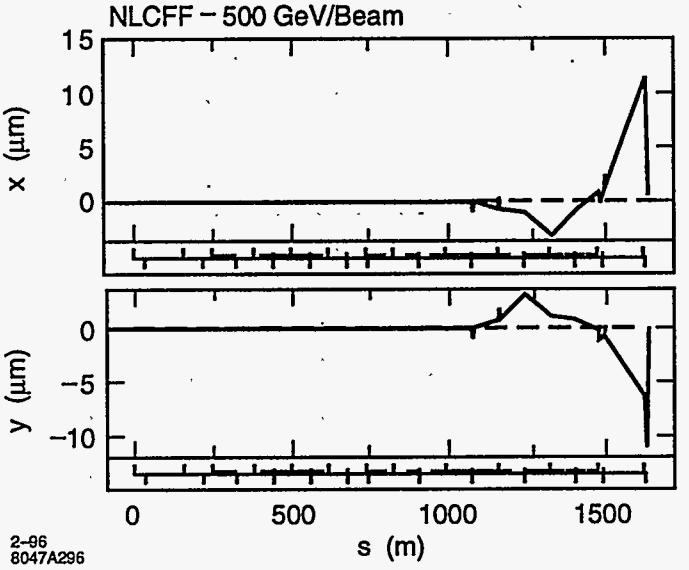


Figure 11-39. *Horizontal and vertical orbit through final focus after 1-μm alignment drift of four quadrupoles (before feedback correction). The luminosity is degraded by a factor of 10.*

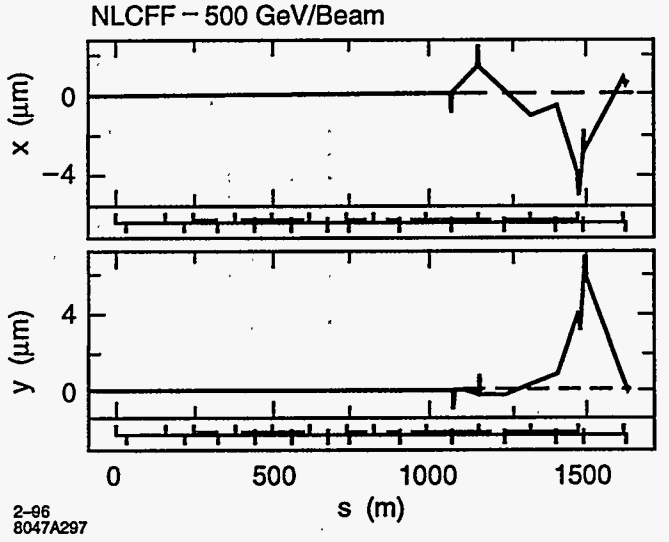


Figure 11-40. *Orbit through final focus after feedback correction. The feedback primarily restores the orbit at the final doublet and second y-sextupole. With more drifted quadrupoles than correctors, some orbit errors remain. The luminosity is restored to within 1% of nominal.*

Note that a BPM with 1-μm rms random noise was used successfully in this case. However, BPM offsets need to remain constant to <200-nm drift over a few hours which is probably the most challenging requirement and may be somewhat inconsistent with a 1-μm resolution BPM.

If high-resolution rf BPMs are required (a few nanometers), the dynamic range of these BPMs is limited, so that, in order to determine and correct magnet motion in the final focus, their readbacks must be averaged over a subset of orbits for which the incoming orbit jitter is small.

For these studies we have chosen to place only one corrector (per plane) in the CCY and one in the FT. It may also be possible to use more correctors so that the corrections are more closely associated with their original kick. This scheme probably allows a wider correction range, however it places much more stringent requirements on BPM resolution requirements in order that individual quadrupole drifts become measurably separable. This possibility has not been studied.

Finally, the linearity of the final-focus optics over both energy and betatron variations has been verified to be adequate (*i.e.*, <200-nm orbit change at BPM-1) over energy deviations of up to $\pm 0.25\%$ and betatron oscillations more than twice the beam size. Therefore, if sampled orbit data is cut at $\pm 0.1\%$ in energy deviation and $\pm$1-sigma in betatron jitter, the optics is completely linear.

The feedback system explored here looks promising. In the limited conditions simulated here the extraordinarily tight alignment drift tolerances in the final transformer can be greatly loosened with a local orbit feedback system by more than a factor of 20 which sets the quadrupole drift scale tolerances at the more livable level of 1 μm rather than 50 nm. Some of the important remaining issues not addressed here are summarized below.

- BPM offsets and gains need to be stable to <200 nm over a few hours.
- Tails in the beam distribution must not introduce a systematic position error.
- Optics need to be stable enough over a few hours (can be measured).
- BPMs must be insensitive to particle backgrounds.

11.5.7 Operations and Controls

Control System and Software

In general, the NLC control system will be based on the SLC experience and will adopt many grown features of the SLC control system. In certain aspects, however, it needs to surpass the SLC control system. Three features are particularly desirable: The control system has to be

- flexible, *i.e.*, accommodate to any unforeseen or changing demands,
- easy to modify and to augment, *i.e.*, allow accelerator physicists and operators to write applications themselves on the fly, and
- fast, *i.e.*, contain the luminosity loss due to tuning, magnet trimming, etc.

The last item is very important. As an illustration, let us consider the time required for scanning an aberration: if we only count a minimum number of 50 pulses per beam-beam deflection scan and take a

settle time of 200 ms equal to four magnet time constants for each magnet trimming, we find that a seven-step aberration scan can be performed in less than 5 s. Since during an aberration scan the luminosity is close to zero, a tuning time of 5 s results in an average luminosity loss of 1.0%, assuming 10 scans per hour. To contain the luminosity loss, the processing time of the control system during a scan should be shorter than the 5-s time needed for tuning and steering. At the SLC, a tuning scan presently takes about a minute, which is almost entirely devoted to processing by the control system. The NLC control system needs to allow for a much faster tuning speed; if necessary this can be accomplished by postponing any analysis or data processing to after the aberration scan is completed and the tuning quadrupoles or sextupoles have been reset to their nominal value.

11.5.8 Components

Main Magnets

The skew correction, diagnostic and geometry sections contain 31 normal and four skew quadrupoles. Parameters are listed in Tables 11-23 and 11-28, respectively, along with power-supply ripple, vibration, and slow drift (stability) tolerances. BMS, CCX, BX, CCY, and FT comprise 40 normal quadrupoles, whose parameters are listed in Table 11-24, and the four quadrupoles of the final doublet, in Table 11-25. Field-stability sensitivities for these magnets are graphically displayed in Figure 11-22, vibration sensitivities in Figures 11-23 and 11-24, stability (or position-drift) sensitivities in Figures 11-25 and 11-26, and, finally, roll sensitivities in Figure 11-27.

Most of the quadrupoles are 0.5-m-long, and their pole-tip field is 3–7 kGauss at 1 TeV, for a typical half-aperture of 10.5 mm. Four quadrupoles in the CCX need a larger half-aperture of 19.5 mm to ensure sufficient beam stay-clear. The largest field gradients are those in the final doublet. Assuming that the half-aperture of the four final-doublet quadrupoles decreases from 11 to 5 mm (Figure 11-14), the required pole-tip fields are about 11 kGauss for the first two, conventional quadrupoles, 34 kGauss for the superconducting magnet QFT1, and 13.5 kGauss for the last quadrupole, QFTA, which is a permanent magnet (see also Section 11.6 and Chapter 12).

At least seven sextupole magnets are foreseen for correcting the horizontal and vertical chromaticity and for increasing the momentum bandwidth of the system. Parameters are given in Table 11-26. The main-sextupole vibration sensitivities are depicted in Figures 11-28 and 11-29. Note that the tolerances for the Brinkmann-sextupoles are extremely loose. The main CCX sextupoles SX1 and SX1b need reversible power supplies, since they will be used to minimize the vertical or horizontal spot size at the pre-image point during initial tune up.

More than half of the final focus is occupied by bending magnets (Table 11-27). Field-stability and roll sensitivities for 7–16-m-long sections of bending magnets are shown in Figures 11-30 and 11-31, respectively. The field-ripple sensitivities appear fairly tight for most of them. However, these ripple sensitivities correspond to jitter tolerances, which have to be met only over a time period of about 1 s, since thereafter orbit-stabilization feedbacks will correct the steering errors.

All sextupoles, quadrupoles, and dipole sections are installed on remotely controlled movers and fed by independent power supplies to facilitate beam-based alignment, orbit feedback and tuning.

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K(10^{-4})$	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QDD	7.00	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD2	4.96	7.	2.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD2	4.96	7.	2.	10.	25.	15.	550.	30.00	200.00
QDF2	5.85	7.	2.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDF	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QDD	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG1	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG2	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG1	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00
QG2	7.14	7.	1.	10.	25.	15.	550.	30.00	200.00

[a] (nm), with respect to ground.

Table 11-23. *Normal quadrupoles in skew correction, diagnostic, and geometry sections.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-5})	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QB1	2.09	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB2	0.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB3	0.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB4	2.95	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QB5	6.54	10.5	1.0	3.3	6.80	7.30	144.0	33.00	11.30
QE2	2.71	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QE2	2.71	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QE1	7.89	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX3	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX1	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX1	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QX3	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX1	7.88	19.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX2	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX3	4.24	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX4	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX5	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX5	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QBX6	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QY3	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY1	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY1	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QY3	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QEI1	3.01	10.5	0.5	1.8	54.00	4.80	15000.	2000.00	15.70
QEI2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI2	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI3	3.01	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QEI4	0.76	10.5	0.5	3.3	6.80	7.30	144.0	33.00	11.30
QFT6	4.12	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40
QFT5	4.01	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40
QFT4	2.51	10.5	1.0	1.6	9.00	3.40	480.	21.00	3.40

[a] (nm), with respect to ground.

Table 11-24. *Normal quadrupoles in the CCX, BX, CCY, and FT.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-5})	Vibration[a] Δx	Δy	Drift[a] Δx	Δy	Roll θ_{xy} (μrad)
QFT3	11.20	11.	2.00	.42	3.60	.70	1600.	98.00	1.80
QFT2	11.20	11.	2.00	.42	3.60	.70	1600.	98.00	1.80
QFT1	34.70	11.	1.50	.42	3.60	.70	1600.	98.00	1.80
QFTA	13.50	5.	1.50	.42	3.60	.70	1600.	98.00	1.80

[a] (nm), with respect to ground.

Table 11-25. *The quadrupoles of the final doublet. QFTA is a permanent magnet, QFT1 superconducting, and QFT2 and QFT3 are conventional.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-3})	Vibration/drift Δx	Δy
SX1	6.66	20.00	0.40	2.20	50.	50.
SX3	1.80	10.50	0.40	9.00	1000.	1000.
SX1	6.66	20.00	0.40	2.20	50.	50.
SY1	3.71	10.50	0.40	2.20	50.	50.
SY3	1.02	10.50	0.40	9.00	1000.	1000.
SY1	3.71	10.50	0.40	2.20	50.	50.
SI2	1.47	10.50	0.40	9.00	1000.	1000.

[a] (nm), with respect to ground.

Table 11-26. *Normal sextupoles in the CCX, BX, CCY, and FT.*

Name	#	B [G]	L (m)	θ (μrad)	$\Delta\rho/\rho$ (Δx) (10^{-6})	$\Delta\rho/\rho$ ($\sigma_{x,y}$) (10^{-6})	$\theta_{x,y}$ (μrad)
BG	2	538.7	5.00	0.5	2.8	15.	0.5
B1A	5	157.5	7.00	0.5	2.8	15.	0.5
B1B	5	157.5	7.00	0.5	2.8	15.	0.5
B2	28	157.5	7.00	0.5	2.8	15.	0.5
B3A	8	157.5	8.28	0.5	2.8	15.	0.5
B3B	4	157.5	9.19	0.5	2.8	15.	0.5
B4	24	112.5	10.00	0.5	2.8	15.	0.5
B5A	12	65.0	10.00	0.5	2.8	15.	0.5
B5B	8	65.0	11.46	0.5	2.8	15.	0.5
B5C	4	12.0	16.00	0.5	2.8	15.	0.5

Table 11-27. *Bending magnets in the final focus. In reality, each bending magnet listed here represents 3–5 shorter magnets on a common power supply, and possibly common support.*

Name	Pole-tip field B_T (kGauss)	Radius a (mm)	Length L (m)	Ripple $\Delta K/K$ (10^{-4})	Vibration/drift Δx (nm)	Δy (nm)
SQ1	4.	6.	0.4	10.	50.	30.
SQ2	4.	6.	0.4	10.	50.	30.
SQ3	4.	6.	0.4	10.	50.	30.
SQ4	4.	6.	0.4	10.	50.	30.

[a] With respect to ground.

Table 11-28. *Skew quadrupoles in skew correction section.*

Type	#	Location
Geom. sextupoles	4	FT
Tuning sextupoles	4	CCX, CCY
Skew quad	1 (4)	FT (SCS)
Magnetic dither	1	at IP
Electrostatic dither	1	at IP
Steering correctors (x or y)	20	at each bend function

Table 11-29. *Steering and tuning elements in the final focus.*

Steering and Tuning Elements

All magnets in the final focus are mounted on remotely-controlled movers, which is exploited for beam-based alignment, for orbit stabilization (feedback), and for steering and tuning procedures. Regardless, there is still need for certain special correction elements: one additional skew quadrupole upstream of the final doublet will be used for correcting the residual y–x' coupling at the IP. Two normal and two skew sextupoles in the final transformer allow complete control over those second-order geometric aberrations which affect the vertical spot size. In addition, four small movable tuning sextupoles are foreseen adjacent to the main sextupoles. For this purpose, used FFTB sextupoles could be employed.

A magnetic dither coil is needed for horizontal beam-beam deflection scans. For vertical beam-beam scans and vertical orbit-feedback at the IP, an electrostatic dither is envisioned, since only tiny deflection angles are required. In addition, two horizontal and two vertical steering correctors are installed at each bend function (for instance, between the two main sextupoles of the CCY) to correct for bend-field drifts and bend rolls.

The tuning and correction elements are summarized in Table 11-29.

Diagnostics, Protection and Tune-Up Elements

The region between the entrance of the SCS and the IP comprises 75 main quadrupoles, 15 sextupoles, and five skew quadrupoles. This probably requires the installation of 75–95 conventional beam position monitors with wide dynamic range (*i.e.*, a few mm) for global alignment, initial tuning, and to diagnose "flyers." In addition about 10 rf BPMs [Hartman priv] will be integrated into the structure of some critical magnets to allow stabilization of the average orbit. This is not strictly necessary (see discussion of FT and CCY feedback in this chapter), but the rf BPMs promise very high resolution (*i.e.*, 10 nm) and correspondingly faster feedback response. A particular rf BPM will be installed at the IP pre-image point in the beta-

Type	#	Resolution	Comment
Conventional BPMs	84	1 μm	at each magnet
Rf BPMs	10	10–100 nm	at critical locations
Laser-wire scanners	8	10 % of beam size	3 wires at each scanner
Laser-interferometer	1 (or 2)	10 % of beam size	in BX (and close to IP)
Conv. wires	8	20 % of beam size	
Toroids	2	1 % of beam current	in SCS and FT
Profile monitor	2	30 % of beam size	in front of dump

Table 11-30. *Diagnostics elements between the SCS and IP.*

Type	#	Comment
Collimators W	4	at sextupoles
Prot. collimators 1/4 r.l. Ti	176	one per quad and dipole
Insertible single-bunch stopper	1	in FT and after DS
Muon spoilers	3	at strategic locations
Magnet sensors	176	in each quadrupole and dipole

Table 11-31. *Protection and tune-up elements between SCS and IP.*

exchanger. This BPM will detect betatron oscillations in the IP phase and, thereby, will allow to correct the beam-beam deflection scans for pulse-to-pulse orbit variations. Performance and reliability of exemplary rf BPMs is being tested at the FFTB.

Beam sizes and emittances are measured in the diagnostic section by means of six laser-wire scanners, each equipped with laser wires at three different angles. For redundancy, an equal number of conventional wire scanners will be installed. The conventional wires can be recycled from the SLC and/or FFTB. They will be useful only during tune up with single bunches and enhanced emittances. A further laser wire is used to infer the beam divergence at the pre-image point of the IP in the BX section (see the discussion in Section 11.4), and a last laser wire, located upstream or inside of the final doublet, measures the IP beam divergence. Finally, a laser-interferometer is foreseen for the pre-image point itself. It needs to be determined if a variant of such a monitor could also be installed inside the detector close to the IP.

Two insertible single-bunch stoppers will be convenient during commissioning and for tune up. One is located behind the DS and facilitates coupling correction prior to bringing beam through the final focus. The second insertible stopper, in the final transformer, allows tuning of the final focus, using single bunches with enlarged emittance at 10-Hz repetition rate, before the nominal beam is sent through final doublet and detector. A conventional profile monitor will be placed in front of each beam stopper.

Magnet sensors in each quadrupole and each dipole detect fast field changes and are part of the machine-protection system. Machine protection is also served by two toroids which measure beam losses in the final-focus region. At the high-beta points close to the main sextupoles, four tungsten collimators are installed. These collimators are important for controlling the detector background. There are also about 176 1/4 r.l. Ti spoilers which will protect quadrupoles and bends from missteered beams. Finally, three muon spoilers located at appropriate positions ensure that the muon background remains acceptable [Keller 1993]. Muon spoilers are treated in Chapter 12.

Tables 11-30 and 11-31 summarize the diagnostics and protection elements, respectively, between the SCS and the IP.

11.5.9 Summary

The proposed final-focus design delivers spot sizes not too far from the linear values, in the entire ZDR operating plane, and it has a sufficiently large momentum bandwidth of 1.2–1.3%, both at 500 GeV and at 1 TeV. Motivated by SLC experience, considerable effort was devoted to evaluating and budgeting all possible sources of spot-size dilution. The budgets for tolerances and aberrations are an integral part of the NLC design. Dedicated stabilization systems and tuning procedures throughout the system ensure that most tolerances can easily be met and maintained. The requirements on ground motion and on magnet-to-ground vibrations are less severe than what has already been achieved at the FFTB. A possible exception is the final doublet, whose vibration tolerances are comparable to measured quadrupole-vibration amplitudes. Therefore, the final doublet will be further stabilized by means of an optical anchor and a seismometer. In general, the final-focus system is very tolerant with regard to parameter changes, and it lends itself to an uncomplicated upgrade from 350 GeV to 1.5-TeV-c.m. energy.

Outstanding questions that still need some work include a simulation of beam-based alignment, and a description of commissioning and operation. None of these items is expected to be difficult. Further studies may also be devoted to the odd-dispersion design proposed by Oide, which could offer cost savings due to its smaller number of magnets and shorter length without compromising the performance. Tunability, tolerances, and operational flexibility of this scheme, in comparison with the present design, remain to be evaluated.

11.6 The Final Doublet

It is the function of final-focus systems to demagnify the beam. Systems using only quadrupoles are limited by a fundamental, unavoidable chromatic aberration in each plane. In present generation final-focus systems, this aberration is compensated by adding sextupoles at dispersive points in the beam line (K. Brown). The amount of chromatic aberration from quadrupoles, to be compensated by the sextupoles, remains a fundamental parameter determining the characteristics of the system, including length, optical functions, and tolerances. One wishes to start with a quadrupole configuration that has the smallest possible chromaticity.

Given a specification for the demagnification, one can show that the smallest chromaticities arise from placing strong quadrupoles as close as physically possible to the focal point. The strength of the quadrupoles is determined by magnet technology and aperture considerations. The free distance from the first quadrupole to the focal point is denoted by the symbol L^*. Factors determining L^* involve background and detector solid-angle considerations.

The function of the final strong quadrupoles then is to match rays having divergences that, on the one side are determined by the proximity of the focal point, and on the other side, are characterized by the remainder of the final focal system. Since the focal system modules are usually much larger than L^* (hundreds of meters compared to one or two meters), the function of the final quadrupoles is simply to focus to a point those particles whose trajectories are parallel to the beam direction.

The simplest such system consists of two quadrupoles and is referred to as the final doublet. The systems described in this section function as a doublet though, for reasons that will be clarified, it is beneficial to use four separate magnets, which we will call the final quartet. Since from an optical point of view the important parameters will be the horizontal and vertical chromaticities, we will seek final doublet designs that meet other criteria and limitations, keeping the chromaticities as small as possible.

The other criteria to be met include:

- L^* is large enough to meet detector solid-angle and background requirements.
- The magnetic material is suitable for a large solenoidal field environment (3 T).
- The magnetic field strength is able to meet stability requirements.
- The magnets can be constructed, supported, and monitored so as to meet alignment tolerances.
- The system satisfies geometric constraints arising from crossing exit and entrance beam lines.
- Synchrotron radiation is below the Oide limit.
- The system has acceptable energy adjustability.

The purpose of this section will be to show how chromaticity varies with doublet parameter choices, and how one may meet the above criteria with a minimum impact on chromaticity. Then, the important tolerances of the system will be explored, and some operational guidelines presented.

Since the final quartet in the beam line will not exactly focus parallel to point, they will vary somewhat from the designs presented in this chapter. If one wishes to know the parameters exactly, consult the lattice "decks." The purpose of this chapter is to clarify the principles underlying the design of the final doublet (quartet), and to give a sense of the trade-offs that were made in arriving at the present design, and could be made if further changes in quartet parameters were desirable to meet additional or modified criteria.

11.6.1 Doublet Parameters

For the purposes of studying the final doublet (FD) as a system independent of the final-focus system, it is a very good approximation to assume that its function is to focus parallel rays to a point at the IP. This can be confirmed by looking at the plots of $\sqrt{\beta(s)}$ for final-focus systems. The ratio of the slope of this function before and after the FD is less than 0.1, both horizontally and vertically.

To begin we will limit pole-tip fields for final doublet quadrupoles to 1.2 T, (the permanent-magnet material of choice, SmCo with Erbium, for a ratio of inner to outer radius of 1/4, has a pole-tip field of 1.35 T), and assume the pole-tip radius is greater than 3 mm. The latter limit arises from resistive-wall wakefields. This means for the quadrupole $Q1, \kappa_1 = B_{T1}/(a_1 B\rho) \leq 120/E(\text{GeV}) = \{0.48, 0.24, 0.16\}\,\text{m}^{-2}$ for beam energies of {250, 500, 750} GeV respectively. In general, $\kappa_2 < \kappa_1$, since a somewhat larger Q2 aperture is usually desirable, but more importantly the contribution to the Oide effect that arises from horizontal betatron motion through the final doublet can be reduced by taking a smaller κ_2.

Given κ_1, κ_2, and the free length, L^* (from the IP face of Q1 to the IP), the lengths L_1 and L_2 of Q1 and Q2 are determined by the parallel-to-point focusing condition. (There is one other parameter, the free space between Q1 and Q2 which we will denote by d. For purposes of this discussion we take $d = 0$.) Since κ_1 will be chosen as large as possible, and its limit is being fixed by magnet technology and resistive-wall limits, we use $k_1 = \sqrt{\kappa_1}\,\text{m}^{-1}$ to scale all lengths and define $\ell^* = k_1 L^*$, $\ell_1 = k_1 L_1$, $\ell_2 = k_1 L_2$, and $\tilde{k} = k_2/k_1$.

There are now two parameters determining the system: L^* and $\tilde{\kappa}$. A sketch of a model final doublet is shown in Figure 11-41.

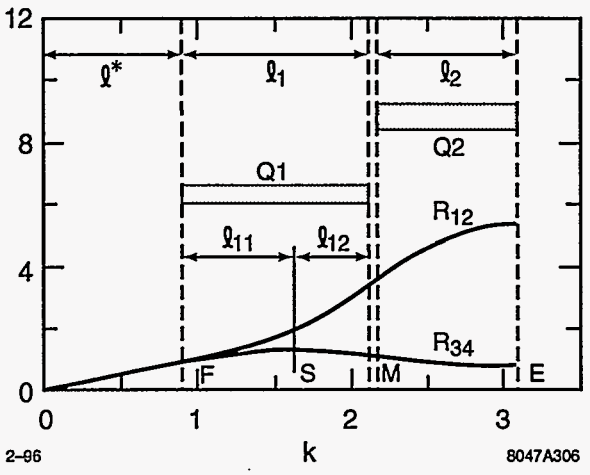


Figure 11-41. *A diagram of a model final doublet.*

The IP end of Q1 is denoted by F (front face), the split between Q1 and Q2 by M (middle), and the entrance to Q2 by E (entrance). Rays have been drawn with unit slope at the IP. This implies that these rays will be the R_{12} and R_{34} functions for horizontal and vertical motion respectively. The magnet Q1 has been split into two parts at S where the R_{34} function is a maximum, defining lengths ℓ_{11} and ℓ_{12}. Note that with $k_1 \approx$ {0.7, 0.5, and 0.4} and $L^* = 2$m, we have $l^* \approx$ {1.4, 1.0, and 0.8} for beam energies $E =$ {250, 500, and 750} GeV, respectively.

11.6.2 An Analytical Model of the Doublet

Element Lengths and Strengths

Letting $x = R_{12}$ and $y' = R_{34}$ be the rays originating from the IP, the condition $y/y' = \ell^*$ at F implies $\ell_{11} = \cot^{-1}\ell^*$ and, at S, $y_{\max} = \sqrt{(\ell^{*2}+1)}$. We have

$$\ell_{11} = \cot^{-1}\ell^* \approx \frac{\pi}{4} - \frac{(\ell^*-1)}{2} \quad \text{for} \quad \ell^* \approx 1 \quad . \tag{11.83}$$

At boundary M, we may derive that the ratio $x/x' = (1+m)/(1-m)$ where $m = (\ell^*-1)/(\ell^*+1)\exp(-2\ell_1)$. From the nature of the y function in the ℓ_{12} region, we can make a rough estimate of $\ell_{12} \approx \pi/6$ and adding this to $\ell_{11} \approx \pi/4$ from above, find $\ell_1 \approx 1.3$, implying an estimate for m < 0.01 even for $\ell^* = 1.4$. (Even if $\ell_{12} \to 0$ which can occur for very small k, we have $m < 0.05$.) In other words, for all parameters under consideration, x/x' is within a few percent of unity at M. This fact makes it possible to find an analytical model for the doublet by setting $x/x' = 1$ at M.

Since for $\tilde{k} = 1$, and m small, $\ell_2 = \pi/4 = 0.79$, it follows that at M $y/y' = \coth(\pi/4) = 1.52$, hence $\ell_{12} = 0.58$. For general $\tilde{k}$, and m small, $\tilde{k}\ell_2 = \cot^{-1}\tilde{k}$. At M, $y/y' = \coth(\tilde{k}\ell_2)/\tilde{k} = \cot(\ell_{12})$ determines ℓ_{12}.

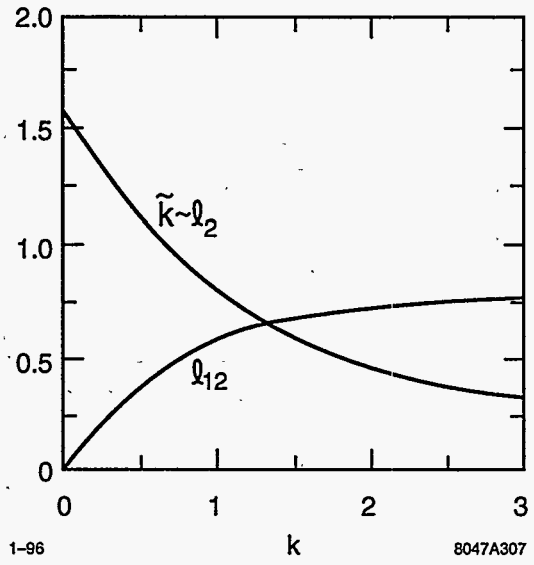


Figure 11-42. *The functions $\ell_{12}(\tilde{k})$ and $\tilde{k}\ell_2(\tilde{k})$.*

We will define a function $c(\tilde{k}) = \coth(\tilde{k}\ell_2) = \coth(\cot^{-1}\tilde{k})$. As $\tilde{k} \to \infty$, $c(\tilde{k}) \to \tilde{k}$, and as $\tilde{k} \to 0$, $c(\tilde{k}) \to \coth(\pi/2)$.

$$\ell_{12}(\tilde{k}) = \cot^{-1}\left[\frac{\coth(\tilde{k}\ell_2)}{\tilde{k}}\right] = \cot^{-1}\left[\frac{c(\tilde{k})}{\tilde{k}}\right] \tag{11.84}$$

The functions $\ell_{12}(\tilde{k})$ and $\tilde{k}\ell_2(\tilde{k})$ are shown in Figure 11-42. As $\tilde{k} \to 0$, $\tilde{k}\ell_2 \to \pi/2$ and $\cot\ell_{12} \to \infty$, hence $\ell_{12} \to 0$. As $\tilde{k} \to \infty$, $\tilde{k}\ell_2 \to 1/\tilde{k}$.

The R_{12} and R_{34} Functions

The values of R_{12} and R_{34} at M and E can now be determined.

$$\begin{aligned} R_{34}^M(\tilde{k}) &= \sqrt{\ell^{*2}+1}\frac{c(\tilde{k})}{\sqrt{c(\tilde{k})^2+\tilde{k}^2}} & R_{34}^E(\tilde{k}) &= \sqrt{\ell^{*2}+1}\sqrt{\frac{c(\tilde{k})^2-1}{c(\tilde{k})^2+\tilde{k}^2}} \\ R_{12}^M(\tilde{k}) &\approx \tfrac{1}{2}(\ell^*+1)e^{\ell_{11}}e^{\ell_{12}(\tilde{k})} & R_{12}^E(\tilde{k}) &\approx \tfrac{1}{2}(\ell^*+1)e^{\ell_{11}}e^{\ell_{12}(\tilde{k})}\frac{\sqrt{\tilde{k}^2+1}}{\tilde{k}} \end{aligned} \tag{11.85}$$

These functions are shown in Figure 11-43a and 11-43b.

Chromaticity

An important parameter for final-focus system design and operational tolerances is the chromaticity, which is given by

$$\int ds\,\kappa_2(s)\beta_z(s) = \frac{1}{\beta_2^*}\int ds\,\kappa_2(s)R_{12}^{z^2}(s) \equiv \frac{L_z^c}{\beta_z^*} \quad . \tag{11.86}$$

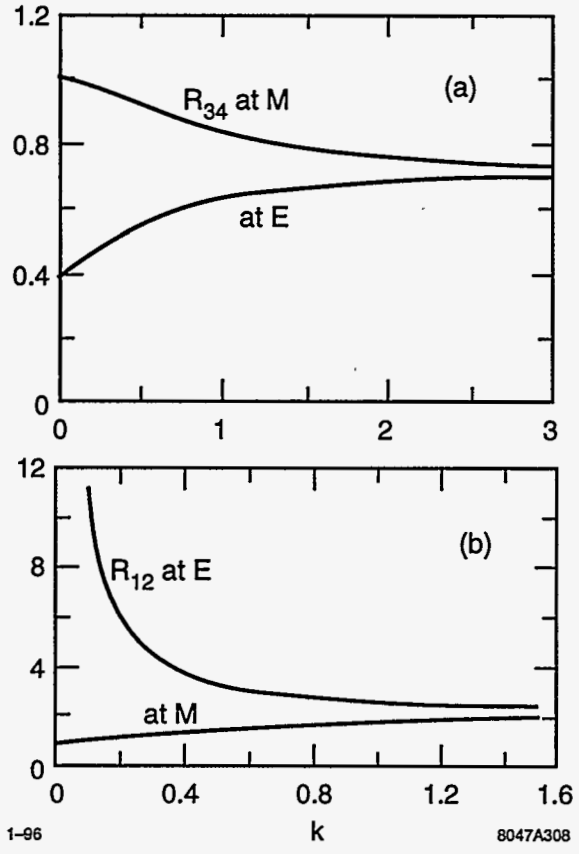


Figure 11-43. *(a) The ray $R_{34}(\tilde{k})$ at M and E. (b) The ray $R_{12}(\tilde{k})$ at M and E.*

The subscript z here can be x or y. $\kappa_z = |\kappa|$ if the quadrupole is focusing in z plane, and $\kappa_z = -|\kappa|$ otherwise.

It is best to consider separately the contributions from the integration of regions specified by ℓ_{11}, ℓ_{12}, and ℓ_2. For the vertical plane, the integral for the regions ℓ_{11}, ℓ_{12} is given by

$$\frac{(\ell^{*2}+1)}{2}\left(\ell + \frac{\cot\ell}{\cot^2\ell + 1}\right) \tag{11.87}$$

where $\ell = \ell_{11}$ or $\ell = \ell_{12}$. The contribution of region ℓ_{11} is independent of $\tilde{k}$ and given by

$$\ell_{11}^{cy} = \frac{\ell^{*2}+1}{2}\cot^{-1}\ell^* + \frac{\ell^*}{2} \quad . \tag{11.88}$$

The contribution from regions ℓ_{12} and ℓ_{l2} is given by

$$\ell_{12}^{cy} + \ell_2^{cy} = \frac{(\ell^{*2}+1)}{2}\left(\ell_{12} + \tilde{k}^2\ell_2\frac{c^2(\tilde{k})-1}{c^2(\tilde{k})+\tilde{k}^2}\right) \quad . \tag{11.89}$$

$\ell_{12}^{cy} + \ell_2^{cy} = 0.13(\ell^{*2}+1)$ at $\tilde{k} = 1$ and approaches 0 as $\tilde{k} \to 0$. These two contributions to the vertical chromatic length add up to a function of ℓ^* plus $(\ell^{*2}+1)/2$ times a function of $\tilde{k}$. At $\tilde{k} = 1$

$$\ell^{cy} = \frac{\ell^{*2}+1}{2}(\cot^{-1}\ell^* + 0.27) + \frac{\ell^*}{2} \quad . \tag{11.90}$$

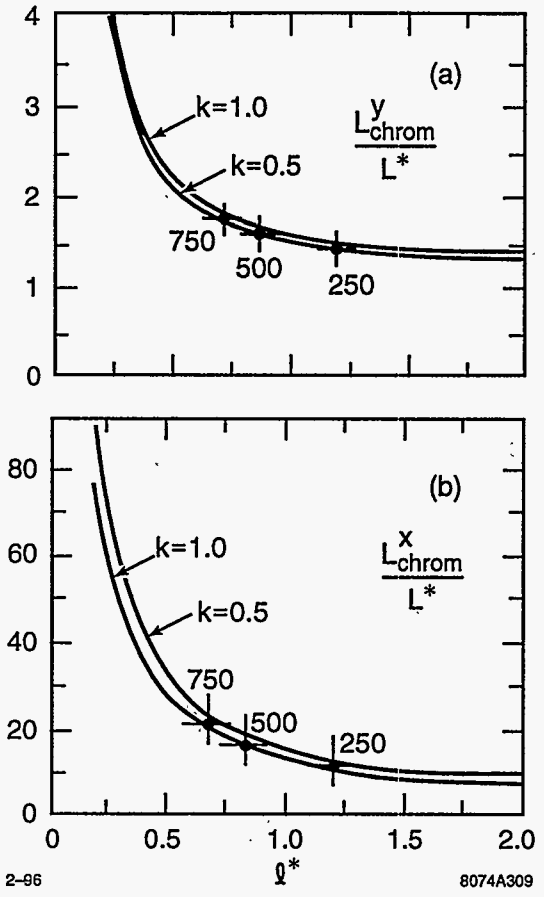


Figure 11-44. *(a) The vertical chromatic length for $\tilde{k} = 1$ and $\tilde{k} = 0.5$ as a function of $k_1 L^* = \ell^*$. (b) The horizontal chromatic length for $\tilde{k} = 1$ and $\tilde{k} = 0.5$ as a function of $k_1 L^* = \ell^*$.*

The chromatic length $L^{cy} = \ell^{cy}/\tilde{k}$. Multiplying and dividing by L^* gives

$$L^{cy} = \frac{L^*}{\ell^*}\left[\frac{(\ell^{*2}+1)}{2}(\cot^{-1}\ell^* + 0.27) + \frac{\ell^*}{2}\right] \tag{11.91}$$

This expression diverges to $+\infty$ for $\ell^* \to 0$ and as $\ell^* \to \infty$ it goes to $L^*(1.0 + 0.135\ell^*)$. There should be a minimum value somewhere. For $\tilde{k} = 1$, the minimum occurs at $\ell^* = 1.85$; the chromaticity at $\ell^* = 1.0$ is 12% larger and at $\ell^* = 0.8$ is 24% larger than this minimum. In Figure 11-44a, we show a plot of L^{cy}/L^* for $\tilde{k} = 1$ and $\tilde{k} = 0.5$. The large crosses marked 250, 500, and 750 show the parameters for doublets at these three beam energies with $d = 0.3\,\mathrm{m}$ (Section 11.6.5). These solutions were done by numerical interaction on the value of x/x' at M, and do not make the assumption that $x/x' = 1$ as in the analytic case. Note how well they fall in the analytic approximative curves!

For the horizontal plane, the chromatic contribution from Q1 is given by

$$\ell_1^{cx} = -\frac{1}{8}(L^*+1)^2 e^{2\ell_1} - \frac{1}{2}(L^{*2}-1) + \frac{1}{2}\ell^* \tag{11.92}$$

and the contribution from Q2 by

$$\ell_1^{cx} = \frac{1}{8}(L^* + 1)^2 e^{2\ell_1} \left[\frac{\tilde{k}^2 + 1}{\tilde{k}} \cot^{-1} \tilde{k} + \tilde{k} \right] \quad . \tag{11.93}$$

The horizontal chromatic length can also be written as the sum of two terms: one which is a function of ℓ^*, and another which is a function of ℓ^* times a function of $\tilde{k}$. We show L^{cx}/L^* for $\tilde{k} = 1$ and $\tilde{k} = 0.5$ in Figure 11-44b. It also has a minimum (at $\ell^* = 6.5$) equal to 5.3. At $\ell^* = 2$, it is 1.24 times this minimum, at $\ell^* = 1.5$, 1.53 times, and at $\ell^* = 1$, a factor of 2.3 times the minimum. Since the β_x^* is a 100 times larger than β_y^*, the horizontal chromaticity is less critical than the vertical.

At this point we can draw the following conclusions:

- The analytic solution based on the R_{12} value-to-slope ratio (x/x') being one between quadrupoles gives the same doublet parameters as the numerical solution (within a few percent).
- There is a minimum horizontal and vertical chromatic length that can be achieved by a final doublet.
- Since the chromaticity is the chromatic length divided by the appropriate β^*, and the desired β_x^*/β_y^* ratio is about 100, the minimum horizontal chromaticity is twenty times smaller than the minimum vertical chromaticity.
- For an L^* of 2 m, field strengths based on a 1.2-T pole-tip field strength, and a beam energy of 250 GeV, the minimum is achieved for the vertical chromatic length and the horizontal chromatic length is about 40% larger than the minimum.
- The worst situation occurs at 750-GeV beam energy where the vertical chromatic length is 25% larger than the minimum and the horizontal chromatic length is three times the minimum.
- These solutions are surprisingly insensitive to changes in the strength of Q2 (Figure 11-44 shows chromatic lengths for equal field strengths and with the Q2 field strength reduced by a factor of 4).

We now turn our attention to the impact of increasing the separation between the magnetic elements Q1 and Q2. We will show that the chromatic lengths are also surprisingly insensitive to this parameter.

Doublets with Large Q1-Q2 Separations

In Table 11-32, we show the chromatic lengths for three cases with parameters appropriate for the 250-GeV beam energy. The pole-tip field was set at 1.35 T, the aperture of Q1 was chosen to be 5 mm, and L^* remains equal to 2 m. The separations were $d = 0.3$ m, 1.5 m, and 3.0 m.

Figure 11-45 shows the doublet solutions corresponding to the smallest and largest separation, $d = 0.3$ m and $d = 3.0$ m. The vertical chromatic length changes very little: in fact, it decreases slightly with increasing separation. The major effect is on the horizontal chromaticity which increases by about 33%, and on the R_{12}^E which increases in Q2 by about 25%. It is interesting to note that the length of both Q1 and Q2 decreases as d increases.

A large separation between Q1 and Q2 can be advantageous: i) the region of the detector fringe can be avoided, and ii) space is available for an adjustable magnet to allow for energy variability. For example, a superconducting magnet could be placed in the fringe region, since there the beam separations are sufficient that the entrance beam-line quadrupole not interfere with the exit beam line.

Q1-Q2 separation (d)	Vertical chromatic length (L_y^c/L^*)	Horizontal chromatic length (L_x^c/L^*)
0.3 m	1.17	14
1.5 m	1.14	16
3.0 m	1.11	19

Table 11-32. *Chromatic lengths for several Q1-Q2 separations.*

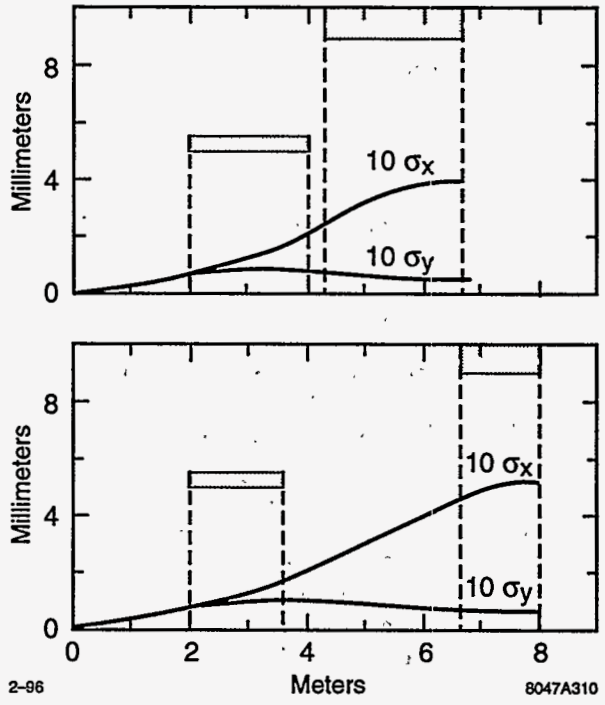


Figure 11-45. *Doublets for separations $d = 0.3$ m and $d = 3.0$ m with field strengths and L^* held constant.*

Skew Quadrupole from Tilt

The vertical displacement of an incoming ray due to the tilt of one or both of the quadrupoles is given by

$$\begin{aligned} \Delta y^* &= 2\int R_{34}(s)\,\theta(s)\,\kappa(s)\,x(s)\,ds \\ &= 2x'^*\int R_{34}(s)\,\theta(s)\,\kappa(s)\,R_{12}(s)\,ds \end{aligned} \tag{11.94}$$

where $\theta(s)$ is the quadrupole tilt specified as a function of s. If we suppose that $\theta(s)$ is constant over some range of s, we can substitute $\kappa(s)\,R_{12}(s) = -R_{12}(s)''$, integrate twice by parts and use $\kappa(s)\,R_{34}(s) = R_{34}(s)''$ to get

$$\Delta y^* = \theta x'^*\,[R_{34}(s)\,R'_{12}(s) - R'_{34}(s)\,R_{12}(s)]_{s_1}^{s_2} \quad . \tag{11.95}$$

We note that because of the boundary conditions $R_{12} = R_{34}$ and $R_{12'} = R_{34'}$ at $s = L^*$, and $R_{12'} = R_{34'} = 0$ at the doublet exit, that $\Delta y^* = 0$ if the whole doublet is rotated. If only parts of the doublet are rotated this

term is quite large. For example, by requiring $\Delta y^*/y^* < 1/5$ (for 2% luminosity loss), we get the condition

$$\theta < \frac{1}{5}\sqrt{\frac{\varepsilon_y}{\varepsilon_x}}\,\frac{\sqrt{\beta_x^*\beta_y^*}}{\left[R_{34}(s)\,R'_{12} - R'_{34}(s)\,R_{12}(s)\right]_{s_1}^{s_2}} \quad . \tag{11.96}$$

Inserting a typical emittance ratio of 1/100 and a $\sqrt{(\beta_x^*\beta_y^*)} \approx 1\,\mathrm{mm}$, and assuming the bracket can be on the order of 1 m, we find the requirement that $\theta < 20\,\mu\mathrm{r}$. Of course, the skew quadrupole aberration can be corrected globally, and this should not be interpreted as an absolute requirement. It does indicate that extreme care should be taken in controlling the tilt of the doublet elements, to the order of 0.1 mr if possible.

11.6.3 Final Doublet Wake Effects

Resistive-wall Wake

We use the standard resistive-wall wake formula (described in Chapter 9, Section 11.8.10).

$$\Delta y_R'^T = 2\,\frac{N r_e}{\gamma \sigma_z}\,\frac{L}{g^3} < f_R > \sqrt{\lambda \sigma_z}\,\Delta y \tag{11.97}$$

where $< f_R >= 0.82$ is the average of the longitudinal shape function of the wake distribution, r_e is the classical electron radius, $\lambda = \rho/(120\pi)$ is called the skin depth (ρ being the resistivity), N is the number of particles per bunch, σ_z is the bunch length, L is the length of the section, g is the beam-pipe radius, and Δy is the offset of the beam from the center. It is the jitter we are primarily concerned with, because a constant steering can be corrected, as long as it is not so large as to effect the beam emittance. In estimating the importance of this term, we use a beam jitter of 1 sigma and integrate the kick along the trajectory of the particle. Since the absolute alignment of the doublet can not be much worse than the maximum of this 1-sigma trajectory (from Oide and second-order dispersion effects), and this term must give a kick that displaces the beam at the IP by less than $0.2\,\sigma_y^*$, emittance growth will not be a concern (see Section 9.2.4 for a comparison of emittance growth with centroid kicks).

The contribution of the resistive-wall wake can be compensated quite easily by increasing the aperture of the quadrupole Q1. The $1/g^3$ is a strong fall-off and apertures of 4 mm are adequate for the 1-TeV-c.m. parameter set. In fact we may take an aperture somewhat larger than this to minimize background effects.

Geometric Wakes

To minimize geometric wakes we suppose that the walls will be tapered anywhere the beam pipe diameter changes. This could presumably occur at the entrance to the final doublet, at the transition between Q1 and Q2, and at the IP end of Q1. In the collimation section, Section 9.2.4, it was shown that there is a broad optimum taper angle at

$$\Theta_T^{opt} \approx 1.1\left(\frac{\lambda\sigma_z}{g^2}\right)^{\frac{1}{4}} \tag{11.98}$$

which equals about 10 mr for $g = 5\,\mathrm{mm}$. Hence the 0.3-m free space between Q1 and Q2 optimally accommodates a 3-mm change of beam-pipe radius. Since radii changes between quadrupoles are less than 4 mm, the total wake from the between-quadrupole region will not be much more than twice the resistive wall part, hence negligible.

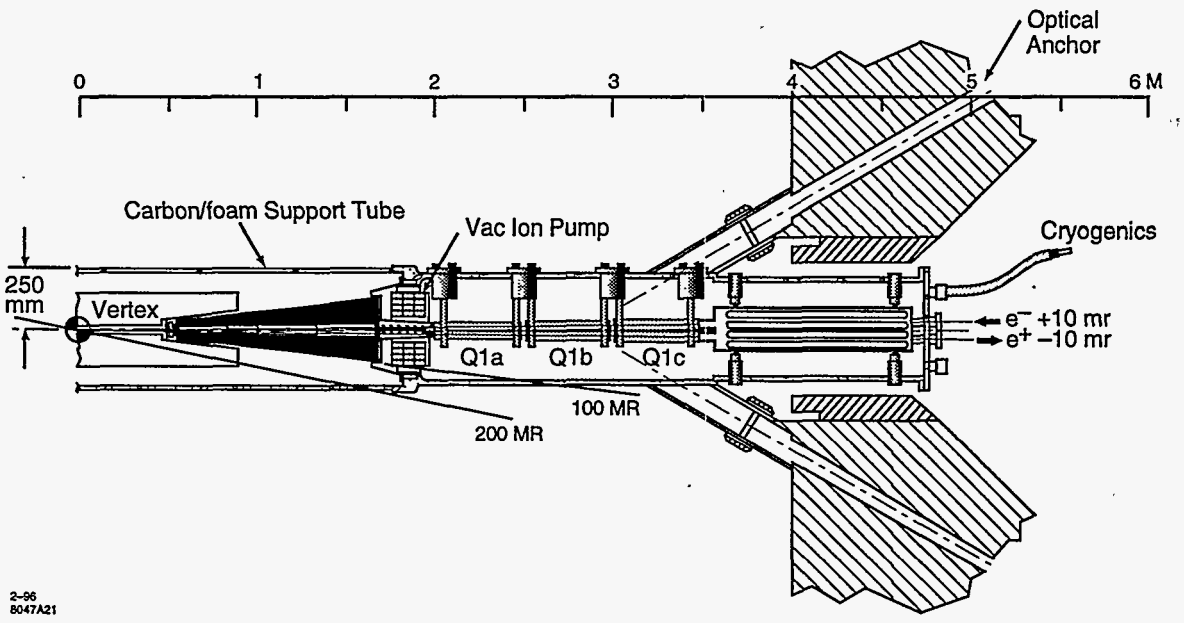


Figure 11-46. *Current final doublet layout with 20-mr crossing angle.*

At the IP end of Q1, the beam pipe will presumably taper up to the tungsten mask and proceed across the IP along the inner radius of the vertex chamber. This section of beam pipe is marked HOM shield in Figure 11-46.

The surface of the beam pipe at the end of Q1 will be complex since there are two holes here, one from the exit beam line and one from the entrance beam line. Though not explicitly considered in the calculations, we imagine that the beam-pipe surface is defined by two cones, each centered on one of the beam-line quadrupole bores, and that these cones terminate either at their intersection with one another or at the inner surface of the tungsten mask. (We have not considered any multibunch effects which might result from resonances of this IP region cavity.) The geometric wake is given by (see Eq. 9.27)

$$\Delta y'^{T}_{G} = 2\frac{Nr_e}{\gamma\sigma_x}\frac{(b-g)^2}{gbL_T} < f_G > \Delta y \quad . \tag{11.99}$$

Using $< f_G >= 0.28$, and $Nr_e/(\gamma\sigma_z) \approx 2\times 10^{-7}$ we have

$$\Delta y^* = L^*\Delta y'^{T}_{G} \approx 10^{-7}L^*\frac{b}{L_T}\frac{\Delta y}{g} \quad . \tag{11.100}$$

Δy at the face of the quadrupole (the smallest part of the cone) is given by $L^* y'_{\rm IP}$. Assuming we have about 1-sigma jitter, this becomes

$$\frac{\Delta y^*_{\rm rms}}{y^*_{\rm rms}} \approx 10^{-7}L^*\frac{b}{L_T}\frac{L^* y'^*_{\rm rms}}{g y^*_{\rm rms}} = 10^{-7}\frac{b}{L_T}\frac{L^{*2}}{g\beta^*_y} \quad . \tag{11.101}$$

Taking $\beta^*_y = 100\,\mu$m, $L^* = 2$ m and $g = 4$ mm, we obtain $\Delta y^*_{\rm rms}/y^*_{\rm rms} \approx b/L_T$. Parameters foreseen for b (2 cm) and L_T (35 cm) would be adequate.

11.6.4 Synchrotron Radiation Effects

The Oide Effect Including Horizontal Motion

The Oide effect [Oide 1988] begins with a calculation of the change in the IP position due to the change of energy of a particle at point s in the beam line. This is determined by the integral

$$R_p^{z6}(s) = \int_0^s ds' \frac{d\theta^z}{ds'}(s') R_{12}^z(s') \tag{11.102}$$

where the subscript p indicates the integral is to be taken along the path followed by the particle from the point s to the origin. For a particle following a betatron trajectory in an aligned quadrupole

$$\frac{d\theta^z}{ds} = -\kappa^z(s) z(s) = -\kappa^z(s) R_{12}^z(s) z'_{\rm IP} \tag{11.103}$$

hence

$$R_p^{z6}(s) = -z'_{\rm IP} \int_0^s ds' \kappa^z(s') R_{12}^2(s') = -z'_{\rm IP} L^{cz}(s) \quad . \tag{11.104}$$

If at the position s a photon of energy u is emitted, the change of the IP position for this particle will be

$$\Delta z_{\rm IP}(s) = R_p^{z6}(s) \frac{u}{E} \quad . \tag{11.105}$$

With a correct accounting of statistics [Sands 1985], the spread in the spot size for an ensemble of particles following this path will then be given by

$$(\Delta z_{\rm IP})^2 = \int ds\, \Delta z_{\rm IP}(s)^2 n(u,s) du = \int ds\, R_p^{z6}(s)^2 \int du \left(\frac{u}{E}\right)^2 n(u,s) \tag{11.106}$$

where $n(u,s)$ is the probability of emitting a photon with energy u (per unit length and energy.) The last integral can be performed and equals

$$\int du \left(\frac{u}{E}\right)^2 n(u,s) = c_s r_e \lambda_e \frac{\gamma^5}{|\rho(s)^3|} \quad . \tag{11.107}$$

Since $\frac{1}{|\rho|} = \frac{|B|}{p/q} = \frac{\sqrt{B_x^2+B_y^2}}{p/q} = \left[\frac{1}{\rho_x^2} + \frac{1}{\rho_y^2}\right]^{\frac{1}{2}}$, this integral can be written

$$(\Delta z_{\rm IP})^2 = c_u r_e \lambda_e \gamma^5 z'^2_{\rm IP} \int ds L^{cz}(s)^2 |\kappa(s)^3| \left[x'^2_{\rm IP} R_{12}(s)^2 + y'^2_{\rm IP} R_{34}(s)^2\right]^{\frac{3}{2}} \quad . \tag{11.108}$$

The integral in Eq. 11.101 becomes a triple integral where one integrates over a Gaussian distribution in $x'_{\rm IP}$ and $y'_{\rm IP}$. We change variables to $x'_{\rm IP} = t\,\sigma^*_{x'}$ and $y'_{\rm IP} = v\,\sigma^*_{y'}$, then in the (t,v) space introduce polar coordinates r and ψ. The radial integral in this space is $\int_0^\infty dr\, r^6\, e^{-\frac{r^2}{2}} = 15\sqrt{\frac{\pi}{2}}$. The angular integral for the vertical direction is

$$\begin{aligned} I(s) &= \frac{1}{2\pi}\int_0^{2\kappa} d\psi \sin^2\psi \left[\cos^2\psi\, \sigma^{*\,2}_{x'} R_{12}(s)^2 + \sin^2\psi\, \sigma^{*\,2}_{y'} R_{34}(s)^2\right]^{\frac{3}{2}} \\ &= \left[\frac{\sigma^{*\,2}_{x'} R_{12}(s)^2 + \sigma^{*\,2}_{y'} R_{34}(s)^2}{2}\right]^{\frac{3}{2}} \frac{1}{2\pi}\int_0^{2\kappa} d\psi \sin^2\psi \left[1 + \lambda(s)\cos(2\psi)\right]^{\frac{3}{2}} \end{aligned}$$

$$\text{where} \qquad \lambda(s) = \frac{\sigma^{*\,2}_{x'} R_{12}(s)^2 - \sigma^{*\,2}_{y'} R_{34}(s)^2}{\sigma^{*\,2}_{x'} R_{12}(s)^2 + \sigma^{*\,2}_{y'} R_{34}(s)^2} \leq 1. \tag{11.109}$$

It is possible to find a good approximation to the final integral by expanding the bracket in a power series. When the integral is carried out on the series, the terms converge very rapidly. Keeping the leading terms, the resulting expression for $I(s)$ is

$$I(s) \approx \frac{1}{16\sqrt{2}} \left[\sigma_{x'}^{*\,2} R_{12}^2(s) + \sigma_{y'}^{*\,2} R_{34}^2(s)\right]^{\frac{1}{2}} \left[\sigma_{x'}^{*\,2} R_{12}^2(s) + 7\sigma_{y'}^{*\,2} R_{34}^2(s)\right] \quad . \tag{11.110}$$

The integral for the horizontal moves the coefficient 7 to the first term in the final bracket. Inserting the result of Eq. 11.110 into Eq. 11.108 we have, for the vertical direction

$$(\Delta y_{\rm IP})^2 \approx \frac{15\sqrt{\pi}\, c_u r_e \lambda_e \gamma^5 \sigma_{y'}^{*\,2}}{32} \int ds\, L^{cz}(s)^2 |\kappa(s)|^3 \left[\sigma_{x'}^{*\,2} R_{12}^2(s) + \sigma_{y'}^{*\,2} R_{34}^2(s)\right]^{\frac{1}{2}} \left[\sigma_{x'}^{*\,2} R_{12}^2(s) + 7\sigma_{y'}^{*\,2} R_{34}^2(s)\right] \quad . \tag{11.111}$$

Thus a good indicator for the importance of the Oide effect, taking into account both the horizontal and vertical motion in the final doublet, can be obtained from a single quadrature.

It is appropriate to first look at the situation described by Oide [Oide 1988] when $\sigma_{x'}^* = 0$. Then

$$I(s) = \sigma_{y'}^{*\,3} R_{34}(s)^3 \frac{1}{\pi} \int_0^{\pi} d\psi \sin^3\psi = \frac{8}{15\pi} \sigma_{y'}^{*\,3} R_{34}^3(s) \tag{11.112}$$

and

$$(\Delta y_{\rm IP})^2 = 4.2 r_e\, \lambda_e\, \gamma^5 \sigma_{y'}^{*\,5} \int ds\, L^{cy}(s)^2\, |\kappa(s)|^3\, R_{34}(s)^3 \quad . \tag{11.113}$$

For a vertical IP divergence angle of 22 mr and $E = 500\,\text{GeV}$, the constant in front of the integral in Eq. (11.113) evaluates to $0.02\,\text{nm}^2$. Hence for a 2% contribution to spot size we need the (dimensionless) integral to remain below 8. The integral over Q2 can be made small by choosing a small value of $\tilde{k}$. The value of $\Delta y_{\rm IP}^2/\sigma_y^2$ is given for each of the doublets in the doublet designs of Section 11.6.5.

Oide Effect with Jitter

In the integral over the Gaussian distribution of $y'_{\rm IP}$, it is possible to offset the distribution assuming a nonzero centroid $y'_{0,\,IP}$. With no centroid offset, the average value of $< y_{\rm IP}'^5 >= 6.8\sigma_y'^5$. With a centroid offset, the average values are given in Table 11-33.

At an orbit displacement corresponding to 1 sigma, the rms of the vertical displacement due to synchrotron radiation effects in the final doublet are doubled. The conclusions that one would derive from looking at Table 11-33 are considerably tempered when one takes into account the non-Gaussian nature of the bunch distribution. Table 11-33 indicates that if we limit the rms beam size increase as an on-axis beam to 2%, then when missteered by 1 σ, the rms would be 4%. However, according to Table 11-34, an rms increase of 4% in reality implies a 2% luminosity loss. (Table 11-34 is based on information taken from Table 1 in [Hirata 1989]. In that study, β^* was varied, keeping the doublet configuration fixed. This is a bit unrealistic. The appropriate studies must be carried out to correctly evaluate the effects of large synchrotron radiation for our parameter configuration.)

Centroid offset $y'_{0\ IP}/\sigma_{y'}$	$\langle(y'_{IP}+y'_{0\ IP})^5\rangle / \langle y'^5_{IP}\rangle$
0.00	1.00
0.25	1.16
0.50	1.70
0.75	2.60
1.00	4.00
1.50	10.00
2.00	22.00
3.00	87.00

Table 11-33. *Growth of synchrotron radiation with centroid offsets.*

Photons radiated per electron (in Q1)	σ/σ_0	L_0/L	Luminosity loss was overestimated by:
0.22	1.012	1.006	2.0
0.25	1.036	1.014	2.6
0.29	1.090	1.025	3.6
0.35	1.330	1.050	6.6
0.46	3.170	1.140	15.6
0.71	10.620	1.380	25.3
1.00	37.700	1.860	42.7

Table 11-34. *Beam-shape modifications due to synchrotron radiation in Q1 are very non-Gaussian. Here we compare luminosity loss that would be estimated from rms beam size increase with proper luminosity integrals (based on [Hirata 1989]).*

11.6.5 Nominal Final Doublet Designs

A Doublet for 250 GeV

A possible doublet for the 250-GeV beam energy is shown in Figure 11-47. The IP parameters are: $E = 250\,\text{GeV}$, $\varepsilon_x = 10^{-11}\,\text{m-rad}$, $\varepsilon_y = 10^{-13}\,\text{m-rad}$, $\beta_x^* = 10^{-2}\,\text{m}$, $\beta_y^* = 10^{-4}\,\text{m}$. The pole-tip field of Q1 equaling 1.5 T was calculated for a SmCo magnet with an outer radius of 20 mm. The pole-tip field in Q2 is 1.2 T. The steering from the end of Q2 is given by $R_{12} = 10.6\,\text{m}$ and $R_{34} = 1.5\,\text{m}$. The stay-clears are 27 σ_x and 60 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute about 0.1% to the vertical beam size. The synchrotron radiation, assuming a final Gaussian shape from the photon emission, would contribute 1% to the beam size. In fact, the shape is not Gaussian, and the luminosity degradation will be less than this number for an aligned doublet. Two-thirds of the beam size growth from synchrotron radiation comes from the vertical motion alone. By choosing a weaker Q2, the horizontal motion contribution was held to 1/2 the vertical contribution.

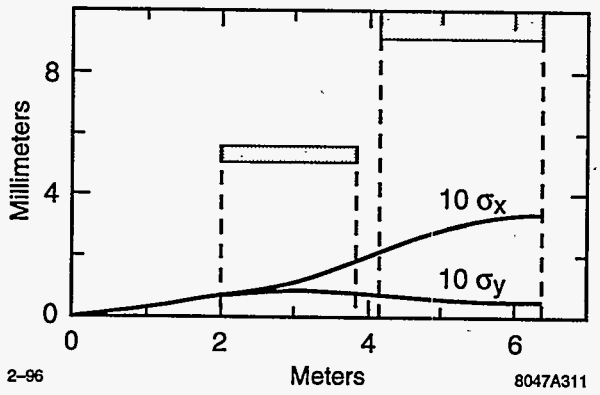


Figure 11-47. *A final doublet diagram for a 250-GeV beam energy showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2\,m$, $d = 0.3\,m$, $B_{T1} = 1.5\,T$, $B_{T2} = 1.2\,T$, $a_1 = 5\,mm$, $a_2 = 9\,mm$. The chromaticities are $\xi_x = 2200$ and $\xi_y = 29,000$.*

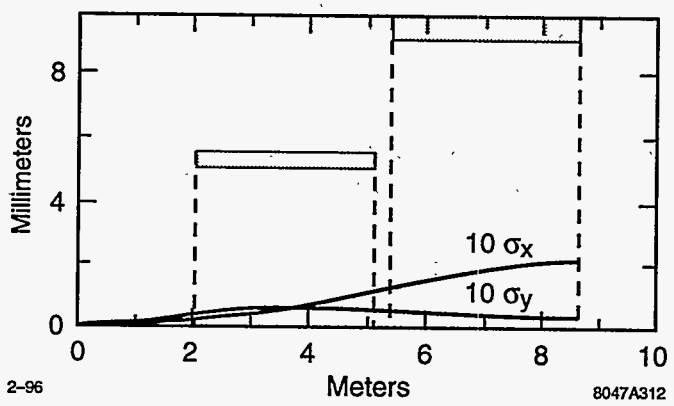


Figure 11-48. *A final doublet diagram for a 1-TeV-c.m. energy showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2\,m$, $d = 0.3\,m$, $B_{T1} = 1.5\,T$, $B_{T2} = 1.2\,T$, $a_1 = 5\,mm$, $a_2 = 9\,mm$. The chromaticities are $\xi_x = 1300$ and $\xi_y = 31,000$.*

A Doublet for 500 GeV

A final doublet for a 500-GeV beam energy is shown in Figure 11-48. The IP parameters are: $E = 500\,GeV$, $\varepsilon_x = 510^{-12}$m-rad, $\varepsilon_y = 5\,10^{-14}$mr, $\beta_x^* = 2.5\,10^{-2}$m, $\beta_y^* = 10^{-4}$m. Note that the beta function ratio is now enlarged to 240 from 100 for the 250-GeV design. This was to keep the parameter n_γ, the number of beamstrahlung photons per electron, to about 1.0. It is also a help for us in reducing the horizontal chromaticity and the contribution of the horizontal motion to synchrotron radiation.

The steering from the end of Q2 is given by $R_{12} = 15\,m$ and $R_{34} = 1.8\,m$. The stay-clears are 42 σ_x and 73 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute 0.2% to the vertical beam size. The synchrotron radiation, assuming a Gaussian shape for the final beam after photon emission, contributes 2% to the beam size. The horizontal contribution is now less than 1/4 of the vertical.

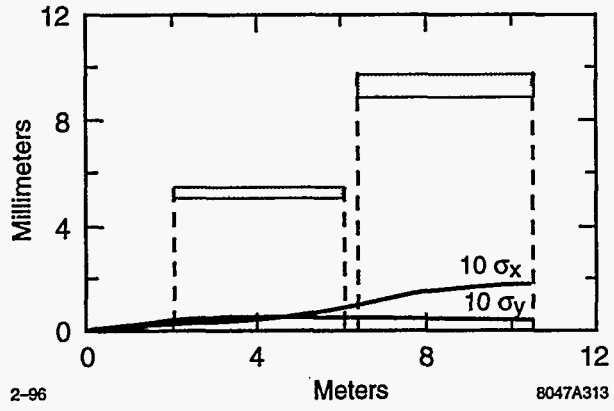


Figure 11-49. *A final doublet diagram for a beam energy of 750 GeV showing 10-σ_x and 10-σ_y ray trajectories. The doublet parameters are: $L^* = 2\,m$, $d = 0.3\,m$, $B_{T1} = 1.5\,T$, $B_{T2} = 1.2\,T$, $a_1 = 5\,mm$, $a_2 = 9\,mm$. The chromaticities are $\xi_x = 1100$ and $\xi_y = 23,000$.*

A Doublet for 750 GeV

Figure 11-49 shows a final doublet for a beam energy of 750 GeV. The assumed IP parameters are: $E =$ 750 GeV, $\varepsilon_x = 3.410^{-12}$ m-rad, $\varepsilon_y = 3.4 \times 10^{-14}$ mr, $\beta_x^* = 3.8 \times 10^{-2}$ m, $\beta_y^* = 1.6 \times 10^{-4}$ m. The steering from the end of Q2 is given by $R_{12} = 19$ m and $R_{34} = 2.0$ m. The stay-clears are 51 σ_x and 96 σ_y. For 1-σ_y jitter, the resistive-wall wake would contribute 0.1% to the vertical beam size. The synchrotron radiation, assuming a Gaussian-shape beam after photon emission, contributes 2% to the beam size.

Final Quartets

Since the doublet performance is not a sensitive function of the distance between Q1 and Q2 (in fact separations of a couple of meters appear satisfactory: the vertical chromaticity decreases slightly, while the horizontal chromaticity increases about 20%) we may place a variable element between Q1 and Q2 to obtain energy adjustability.

We propose to place a variable superconducting quadrupole (Q1.5) between Q1 and Q2 in the region of the detector fringe. This quadrupole provides adjustability, and avoids permanent magnet material in the fringe region where the perpendicular components of the solenoidal field become large, and can cause demagnetization.

The quadrupole Q2 is now outside the solenoid and easy to support and adjust. Because of its length, it may be practical to divide it into two pieces. Thus we arrive at the "final quartet."

Figure 11-50 shows two possible final quartets for 250 and 500 GeV beam energy, respectively. Q2 was not divided into two parts so only three elements appear in this figure.

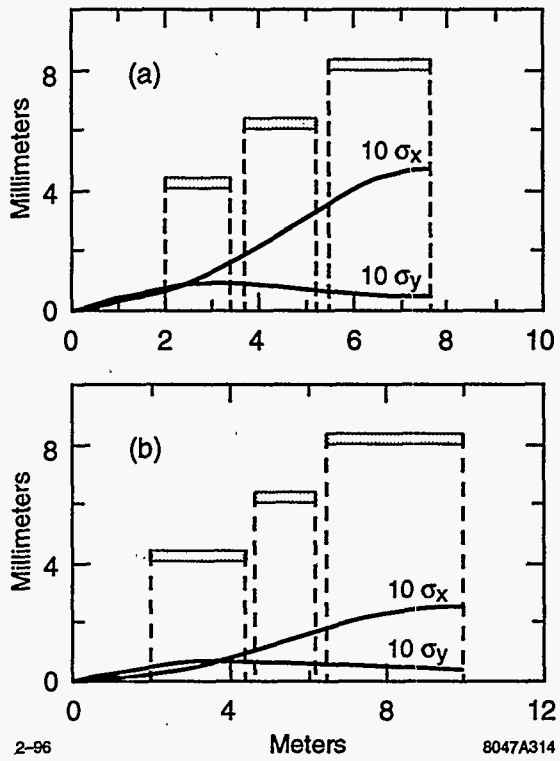


Figure 11-50. *Final quartets at 250- and 500-GeV beam energy.*

11.6.6 Tolerances

Steering Jitter Tolerances

Luminosity Loss

The luminosity for missteered beams, without disruption, goes as

$$L \propto \frac{1}{\sqrt{\sigma_R^2 + \sigma_L^2}} \exp\left(-\frac{\Delta y^2}{2(\sigma_R^2 + \sigma_L^2)}\right) \tag{11.114}$$

where Δy is the distance between the two beams. The beam separation Δy may arise from random uncorrelated missteering of the two beams, in which case the average displacements satisfy $\Delta y^2 = \Delta y_R^2 + \Delta y_L^2$. Assuming the right and left σ are equal, and the right and left error budget is the same, 2% luminosity loss occurs at rms values $\Delta y_L = \Delta y_R = (1/5)\sigma$.

If the offset is arising from correlated motion of elements to the right and left of the IP, then a 2% loss occurs at an rms $\Delta y = \sqrt{2}/5\sigma$. If the motion is correlated and sinusoidal in nature, then $< \Delta y^2 >= 0.5\ \Delta y_{\max}^2$, hence a 2% loss occurs at $\Delta y_{\max} = (2/5)\sigma$.

There is considerable vertical disruption for all NLC parameters, which means that the separated beams are steered toward one another during the collision. The luminosity formula follows the same relation with σ

Machine Section	Steering Budget	Accumulated Jitter	Luminosity Loss
To end of linac	$\sqrt{2}/5\sigma$	$\sqrt{2}/5\sigma = 0.28\sigma$	1%
Collimation system:	$1/5\sigma$	$\sqrt{3}/5\sigma = 0.35\sigma$	1.5%
w/ amplification:		$\sqrt{3}/4\sigma = 0.43\sigma$	2.3%
Big bend	$1/10\sigma$	0.44σ	2.4%
Final focus	$1/5\sigma$	0.48σ	2.9%
Final doublet	$1/5\sigma$	0.52σ	3.4%

Table 11-35. *Proposed NLC jitter budget.*

Aberration	Displacement y_{Q1}	y_{Q2}	y_{Q1T}	y_{Q2T}	y_E	$R_{Ey'E}$
y_{IP}	1.63	−0.63	0.04	0.04	0	−1.00
η_{IP}	−1.33	−0.55	0.11	−0.23	1.88	−2.29
$\eta_{2,IP}$	1.35	0.47	−0.09	0.25	−1.07	2.20
$\eta_{IP\|S}$	2.40	−1.99	0.21	−0.15		
$\eta_{2,IP\|S}$	−2.32	1.85	0.19	0.17		
$\eta_{IP\|SD}$	−0.85	0.72	−0.07	0.08		

Table 11-36. *The strength of aberrations for a variety of doublet displacements and incoming beam errors. This table was compiled for the 250-GeV final doublet of Figure 11-47.*

replaced by 3 σ in the exponent of Eq. 11.114 at design intensities. If the intensity is one-half the design, it is appropriate to replace σ by 2 σ in the exponent. To be conservative, we make the latter assumption.

11.6.7 Steering Jitter Budget

So that 2% luminosity losses do not proliferate to an unacceptable sum, we have assigned the budgets indicated in Table 11-35. A budget is also necessary because with a large beam jitter it becomes difficult to carry out beam-based alignment, wakefield effects are enhanced, and synchrotron radiation in the final doublet is enhanced (Section 11.6.4).

Steering Coefficients. With this section, we begin a discussion of aberrations which arise from the displacement of the doublet elements or from an incoming beam that has a nonzero position or slope. The aberration strength can be related to the displacement by a set of dimensionless coefficients, such as are given for the 250-GeV beam energy in Table 11-36. We will discuss the meaning of each of these coefficients, beginning with the steering coefficients.

There are potentially two coefficients relating IP motion to incoming beam centroid parameters. Since the final doublet focuses parallel rays to the IP, one of them is zero. The minus sign in Eq. 11.115 comes from

the convention we use that the longitudinal distance is increasing, moving away from the IP.

$$\frac{\partial y_{\rm IP}}{\partial y_E} = 0 \qquad \text{and} \qquad \frac{\partial y_{\rm IP}}{R_{34}^{E\to {\rm IP}} \partial y'_E} = -1 \tag{11.115}$$

We include the R_{34}^E ($1.5\,\mathrm{m} < R_{34}^E < 1.8\,\mathrm{m}$) with the y'_E so that steering and dispersion coefficients are dimensionless.

There are two coefficients relating IP beam-position change to quadrupole displacements,

$$\frac{\partial y_{\rm IP}}{\partial y_{Q1}} \approx 1.6 \qquad \text{and} \qquad \frac{\partial y_{IP}}{\partial y_{Q2}} \approx -0.6 \quad , \tag{11.116}$$

which must add to one because a displacement of both elements shifts the focal point by the same distance. There are two coefficients relating IP motion to quadrupole pitch angle. The subscript T designates "tilt." $y_{Q1T} \equiv 0.5\, L_{Q1}\, y'_{Q1}$ is defined as the displacement of the end of the quadrupole farthest from the IP.

$$\frac{\partial y_{\rm IP}}{0.5 L_1 \partial y'_{Q1}} \approx 0.04 \qquad \text{and} \qquad \frac{\partial y_{\rm IP}}{0.5 L_2 \partial y'_{Q2}} \approx -0.04 \quad . \tag{11.117}$$

These are quite small because the R_{34} function is symmetric in Q1 and rather flat in Q2.

We can draw the following conclusions from these steering coefficients:

- If each quad is supported as a rigid unit, the tilt of this unit is less important by more than an order of magnitude than its position.
- Mounting both quadrupoles in a single long barrel should be avoided. In such a case, for the Π-mode vibration of the barrel, Q1 would be moving upward while Q2 was moving downward, and both would move the IP position in the same direction.
- Mounting the Q1 quadrupoles in a single barrel would be better. In such a case, the support of the barrel should be directly beneath Q1. Then the Π-mode would consist of a rotation of the Q1 quadrupole as a whole, and as noted above, this has a very small effect on the IP position.

Steering Jitter Tolerances. We assume, that at about 5 Hz, the steering feedback can reduce the impact of jitter. At frequencies above 5 Hz, element changes in the IP position could be measured and compensated by steering.

With the Q2 quadrupoles in separate barrels, it is appropriate to consider the four degrees (four is enough since we can neglect tilt) to be specified by the four linear combinations: i) $y_{Q1R} + y_{Q1L}$, ii) $y_{Q1R} - y_{Q1L}$, iii) $y_{Q2R} + y_{Q2L}$, and iv) $y_{Q2R} - y_{Q2L}$. The symmetric combinations move the IP symmetrically, so the beams remain in collision. Only the two antisymmetric motions are significant. The Q2 coordinates are less important than the Q1 by a factor of 2.7, so the jitter tolerance on the system is determined by the combination $y_{Q1R} - y_{Q1L}$. If it is 1 nm, the beams will miss by 1.6 nm. Assuming a doublet jitter budget of 1/2% luminosity loss, and assuming a factor of 2 from disruption (valid at 1/2 the design intensity), the tolerance on $y_{Q1R} - y_{Q1L}$ is $\sigma_y^*/5$.

We note that the motion of Q1R and Q1L should be highly correlated because they are separated by only 6 m. This is a fraction of a wavelength for motions up to 15 Hz. At these frequencies, seismic motion is very small (<0.1 nm) at most locations. Thus it is the cultural noise that must be carefully controlled at these frequencies.

Machine Section	Dispersion Budget	Luminosity Loss
To end of linac	$1/5\ \sigma$	3% tent.
Collimation system	$1/10\ \sigma$	0.5%
IP switch/big bend	$1/10\ \sigma$	0.5%
Final focus/final doublet	$1/5\ \sigma$	2%

Table 11-37. *Proposed NLC dispersion budget.*

Stability Tolerances from Dispersion

Dispersion Budget. It is assumed that dispersion is being tuned at regular intervals, every hour or less. Since the dispersion aberration is the most sensitive to beam position, it will likely require the most tuning. Ideally, the dispersion tuning would be automated. Dispersion at the IP implies a centroid off-set that is correlated with pulse-to-pulse energy. This correlation could be detected in the steering feedback data, and continuously minimized.

Dispersion at the IP can arise from several sources; the largest source is from the beam position in the final doublet and at the CCY sextupoles. However, this beam position in the final doublet is with respect to the beam position in the final-focus system, because incoming jitter will not give rise to dispersion. It is the final-focus system, especially the CCY and the final telescope that must be stabilized. A proposed dispersion budget is shown in Table 11-37.

Allowing a 2% loss in luminosity from residual dispersion, the rms value of the dispersion aberration at the IP must be less than 1/5 the beam size, $< \eta_1\delta >_{\rm rms} \leq \sigma y^*/5$. Using a square distribution of width $\pm\Delta$, the rms of $\eta_1\delta$ is $< \eta_1\delta >_{\rm rms} = \eta_1\Delta/\sqrt{3}$. Thus, for $\Delta = 510^{-3}$ ($\delta_{\rm rms} \approx 3\times10^{-3}$) the first-order dispersion must satisfy $\eta_1 < 67\sigma_y^*$. At 1 TeV and $\sigma_y^* = 3\,{\rm nm}$, we must have $\eta_1 \leq 200\,{\rm nm}$. This sets the first-order dispersion scale.

11.6.8 Sources of First-order Dispersion in the Doublet

Dispersion arises from a change of beam position at the doublet entrance or a change of beam position at the IP. These are given in Table 11-36 as

$$\frac{\partial\eta_1}{\partial y_E} \approx 1.9 \qquad \text{and} \qquad \frac{\partial\eta_1}{R_{34}^{E\to IP}\partial y'_E} \approx -2.3 \quad . \tag{11.118}$$

The first term, corresponding to a displacement of the incoming beam or a vertical motion of the doublet system as a whole, indicates a tolerance of $200/1.9 = 105\,{\rm nm}$ for the 1-TeV machine. The vertical beam size at this point is about $33\,\mu{\rm m}$, so this tolerance is about 1/300 of the beam size. This deserves some explanation.

The tolerance arises from the fact that passing off-axis through large chromaticity gives rise to dispersion. Since the chromaticity is compensated by the sextupoles in the CCY section of the final-focus system, an incoming jitter of the beam will pass off-axis through three large chromaticity points (the two sextupoles and the final doublet). Since the chromaticity has been adjusted to cancel, the dispersion will cancel as well. However, if the final doublet moves with respect to the final-focus system, or a quadrupole in the final

telescope moves, or if the strength of one of the dipoles in the final telescope changes, then the dispersion will not cancel. So this problem can be approached in three ways:

- Control all systems to meet the stability requirements for time intervals equal to the time between dispersion tunings (this amounts to 20 nm for a few quadrupoles in the final telescope). For the FFTB, the observed motion (measured relative to a wire alignment system) is less than a micron per hour. This is a ground surface installation, and one might expect to do better underground.
- Have very sensitive BPMs ("rf BPMs") located at crucial positions, such as at the main CCY sextupoles and final doublet quadrupoles, and use all BPMs in the final-focus system to determine the incoming jitter of any pulse. Then with a minute's worth of pulses (10,800), select 1% (108 pulses) with the smallest jitter amplitude for further analysis. The amplitudes of these pulses are small enough so that the higher sensitivity rf BPMs can be used, since for these selected pulses the distribution width at the final doublet entrance will be 330 nm. Now, if the rf BPM had the required sensitivity (<100 nm), any one of these pulses would determine if something in the beam line had moved, because it is a change in relative readings in the system that is of consequence. 100 pulses is an abundance.
- Use the fact that the presence of dispersion at the IP implies energy dependent steering. A correlation of IP position (as determined by beam-beam pulse-to-pulse deflections) with pulse-to-pulse energy from both right and left beams can be used to continuously correct the dispersion.

The second equation in Eq. 11.118 is of importance if the IP collision point wanders. A proper analysis requires consideration of the two doublets and the optimal method for their operation as a pair. In addition to holding the entry position (or the position in Q2) to each doublet constant, some other measurement must be used. One might imagine a good choice would be to minimize the change in the absolute sums of BPM readings in the two exit quadrupoles. This is not a good choice, because the ray that begins at the corrector in front of Q2 has a small value at the exit quadrupole in comparison to the size of this ray in Q1, where the dispersion is created. A better choice is to minimize the sum of the absolute changes of BPM readings in the two Q1s.

So let us imagine that this absolute sum is minimized and the beams are held in collision. To begin with, we could assume that the IP position does not change. The dispersion which arises from a small change in Q1, subject to this constraint, is:

$$\left.\frac{\partial \eta_1}{\partial y_{Q1}}\right|_s = \frac{\partial \eta_1}{\partial y_{Q1}} - \frac{\partial \eta_1}{R^{E\to IP}\partial y'_E}\frac{\partial y_{IP}}{\partial y_{Q1}} \quad . \tag{11.119}$$

The values for these quantities are also shown in Table 11-36

$$\left.\frac{\partial \eta_1}{\partial y_{Q1}}\right|_s = 2.4 \qquad \text{and} \qquad \left.\frac{\partial \eta_1}{\partial y_{Q2}}\right|_s = -2.0 \tag{11.120}$$

The size of these coefficients emphasizes the importance of keeping the beam passing through the centers of these quadrupoles. In other words, misalignments of $200/2.4 = 83$ nm is the maximum displacement of Q1 before the dispersion becomes unacceptable. As mentioned, this must be diagnosed and continually tuned or compensated. The beams must be held in collision, and the dispersion of each beam must remain unchanged. The latter can be arranged by changing the position at E (the IP position does not change, but the dispersion does.)

Stability Tolerances from Waist Motion

The strength aberration is also given by the chromaticity function. If the strength of both magnets in the doublet changed together, the tolerance on the strength would be:

$$\frac{\Delta k}{k} \leq \frac{1}{10\xi} \approx 3.3 \times 10^{-6} \tag{11.121}$$

where we have budgeted a 1/2% luminosity loss for this aberration to the final doublet, and used 30,000 for the vertical chromaticity. If only the strength of Q1 changed, this tolerance is about 20% smaller.

Thermal Stability for Permanent Magnets. The material SmCo has a strength coefficient of $-4 \times 10^{-4}/°\mathrm{C}$, which would lead to the conclusion that the magnets should be stable to 10^{-2} °C in times that are on the order of waist tuning times (which we would like to have at one-hour intervals). This could conceivably be accomplished by providing a controlled thermal environment for the magnets. Temperature deviations of this order could be measured and small waist corrections introduced to compensate thermal drift.

Another possibility is to use a permanent magnet material that is more stable. There exists permanent magnet materials, $Sm_xEr_{1-x}Co$, containing erbium, that have very flat temperature coefficients, down to $10^{-6}/°\mathrm{C}$. In fact there are even materials which increase in strength with increasing temperature, at $10^{-5}/°\mathrm{C}$ which could be chosen so as to compensate for the thermal expansion of the doublet. The remnant field of these materials is about 0.9 T, whereas the straight $SmCo$ remnant field can be as high as 1.06 T. However, 0.9 T seems adequate, and can give a pole-tip field of 1.35 T.

Temporal Stability of Permanent Magnets. The temporal stability of permanent magnet materials has been measured to be about $10^{-6}/hr$. Our experience with permanent magnets in accelerator environments also confirms this number [Spencer 1995]. Thus the temporal stability of permanent magnets is acceptable. The temporal stability of the $Sm_xEr_{1-x}Co$ needs to be checked.

Static Alignment Tolerances

Static Tolerances from Second-Order Dispersion. If we suppose that there is no knob in the final-focus system to adjust second-order dispersion, then this aberration will establish limits on static alignment tolerances. Synchrotron radiation also sets a limit on static alignment. The lowest of these two limits must be taken as the static alignment specification.

In calculating this tolerance, we assume that steering errors arising from static misalignments are being corrected by steering elements at E (the entrance end of Q2), and the dispersion is being corrected by adjusting the beam position at E. The appropriate coefficients are of the form:

$$\frac{\partial \eta_2}{\partial y_{Q1}}\Big|_{sd} = \frac{\partial \eta_2}{\partial y_{Q1}}\Big|_s - \frac{\partial \eta_2}{\partial y_E}\frac{\frac{\partial \eta_1}{\partial y_{Q1}}\Big|_s}{\frac{\partial \eta_1}{\partial y_E}\Big|_s} \quad . \tag{11.122}$$

The aberration term $\eta_2\delta^2$ contains a steering aberration, $\eta_2 < \delta^2 >$. After subtracting this steering, the rms spot size growth is $\eta_2\sqrt{\left(\langle\delta^4\rangle - \langle\delta^2\rangle^2\right)}$. For a square distribution of width $\pm\Delta$, this quantity equals $0.3\eta_2\Delta^2$. If the limit on luminosity loss from the second-order dispersion is taken to be 2% and $\Delta = 5 \times 10^{-3}$, then

the second-order dispersion must satisfy $\eta_2 \leq 3 \times 10^4 \sigma_y^*$. At 1 TeV for $\sigma_y^* = 3$ nm, $\eta_2 \leq 90\,\mu$m. This distance sets the scale of the static alignment tolerances arising from second-order dispersion effects.

Referring to Table 11-36, we see

$$\left.\frac{\partial \eta_2}{\partial y_{Q1}}\right|_{SD} = -0.85 \qquad \left.\frac{\partial \eta_2}{\partial y_{Q1}}\right|_{SD} = -0.72 \qquad \left.\frac{\partial \eta_2}{\partial y_E}\right|_{SD} = -1.0p7 \quad . \tag{11.123}$$

The fact that all of these coefficients are close to unity, indicates the absolute alignment tolerances arising from second-order dispersion are on the order of 90 μm. A smaller absolute alignment tolerance comes from synchrotron radiation considerations.

Static Tolerances from Synchrotron Radiation. Table 11-33 indicates that for a 0.5-sigma vertically-missteered beam, the synchrotron radiation integral is larger by a factor of 1.7. That means that a displacement of the incoming beam by about 16 μm results in an additional 1.4% spot size increase. (This estimate may be a factor of two too high. See Table 11-34). Though we have not computed the synchrotron radiation increase from element misalignments, this tolerance arises from the integral in Q1, hence it is the orbit in Q1 which sets this number. Since the missteered beam has an orbit maximum which is somewhat larger in Q1, the misalignment tolerance will be somewhat larger than the 16-μm number deduced above. This number is a factor of five smaller than the static tolerance limits arising from second-order dispersion.

Meeting Static Alignment Tolerances. To meet absolute (static) alignment tolerances, it will be necessary to locate the beam with respect to the magnetic centers in the final doublet system. One must either provide for beam alignment based on varying quadrupole strengths, or in the case of permanent magnets, provide BPMs with appropriate absolute alignment accuracy. For permanent magnet systems, we will assume that BPMs are located at all element centers and determine the position of the magnetic centers with an absolute accuracy denoted by $\Delta y_{\rm BPM}$. For example, rf BPMs may be integrated into the permanent magnet quadrupole structure.

We begin by assuming the BPM centers are coincident with the magnetic centers. Thus they establish the two doublet center lines (say right and left) to which the incoming beam may be steered. Let us begin by neglecting the steering that may be required to arrange that each incident beam is traveling along this center line. The IP separation of these beams will be given by

$$\Delta y = \frac{s_2}{s_2 - s_1}(y_{1R} - y_{1L}) - \frac{s_1}{s_2 - s_1}(y_{2R} - y_{2L}) \quad . \tag{11.124}$$

As steering is introduced to collide the beams, it should be done with equal corrector strength on each side. Then magnets Q1 or dipole coils around Q1 should be equally adjusted so that the required steering from correctors is zero on each side. If the BPM centers were equal to the magnetic centers, the doublet would now be perfectly aligned. However, there will be BPM error. Synchrotron radiation, as noted in the previous paragraph, sets the limit on the absolute accuracy of these BPMs to be $\Delta y_{\rm BPM} \leq 16\,\mu$m.

11.6.9 Superconducting Quadrupole Q1.5

We propose to place a superconducting quadrupole, Q1.5, upstream of the permanent magnet Q1. This quadrupole serves three purposes: It, first, can provide a high gradient up to about 300 Tm^{-1}; it, second,

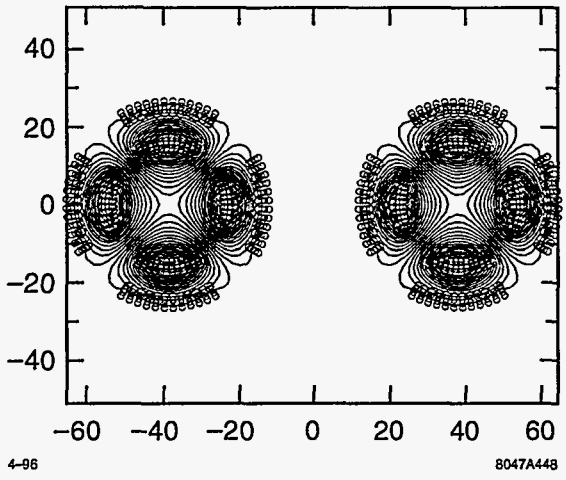


Figure 11-51. *Current and field distributions for the superconducting quadrupole pair Q1.5.*

allows operating the same doublet over a wide range of beam energies; and, third, it avoids permanent-magnet material in the solenoid-fringe region. The exit face of Q1.5 is located 3.85 m from the IP, where the incoming and outgoing beam lines are separated by about 7.7 cm. The inner bore is 11–14. mm, and the pole-tip field at 1-TeV-c.m. energy is 3–4 T.

Using a regular superconducting quadrupole would lead to a considerable field in the region where the second beam will pass, which is not desirable. We, therefore, intend to use a special magnet configuration: the current distribution will be arranged in an inner and outer shell so as to approximate two concentric $\cos 2\theta$ distributions with opposite polarity. While such a magnet can generate about the same field gradient as a conventional superconducting magnet, it has the additional advantage that the stray field experienced by the other beam is essentially zero (it is about 2–3 G).

Figure 11-51 illustrates the proposed current distribution and magnetic field pattern at the front face of the superconducting quadrupole pair viewed from the IP. Figure 11-52 shows the field-strength variation along the centerline of the two magnets on a linear scale, assuming a pole-tip field of 2 T and an inner-bore radius of 14 mm. The field falls off rapidly outside of and between the two quadrupoles, and it is very linear inside of each magnet. Finally, in Figure 11-53 the absolute value of the field is depicted on a logarithmic scale. Note that the field should go through zero at the center of the magnet, and that the finite value shown is an artifact of the drawing.

11.6.10 Summary

We have presented doublet designs for beam energies of 250, 500, and 750 GeV, whose chromaticities are close to the theoretical minimum and additionally

- have acceptable Oide effect synchrotron radiation,
- have acceptable resistive-wall and geometric wakes,
- have sufficient beam stay-clear,

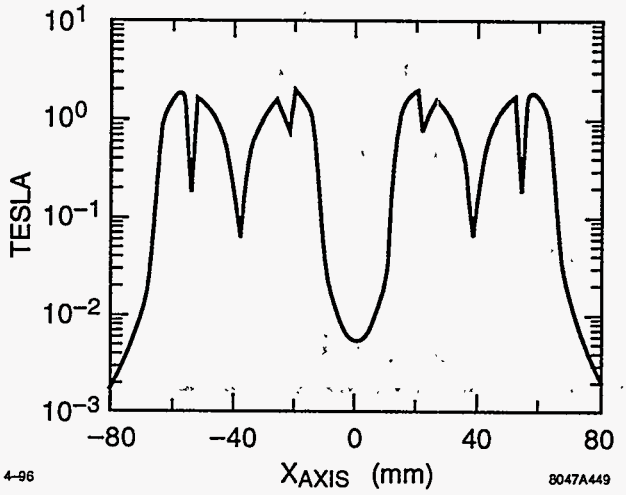


Figure 11-52. *Field strength along the centerline of the superconducting quadrupole pair Q1.5 on a linear scale.*

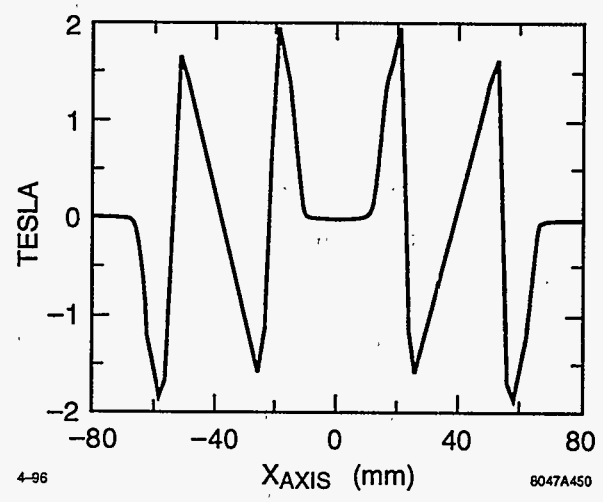


Figure 11-53. *Absolute field value along the centerline of the superconducting quadrupole pair Q1.5 on a logarithmic scale.*

- provide for a large range of adjustment to changing beam energy (about a factor of 2),
- are sufficiently stable thermally and temporally,
- have no permanent-magnet material in solenoid fringe region,
- provide sufficient space for an exit beam path, exit beam-line elements and instrumentation at the 20mr crossing angle, and
- provide a mechanism (dipole coils) to achieve the required absolute alignment and to compensate solenoid field changes.

Final-doublet support is discussed in Section 12.4.2.

11.7 Crossing Angle, Crab Cavity, and Solenoid

Section 11.7 begins with a discussion of the factors involved in the choice of 20 mr as the IP crossing angle. Such a relatively large crossing angle implies the need for a crab cavity. The crab-cavity parameters and tolerances are discussed in Section 11.7.2. The crab cavity is placed either at the entrance of the final doublet, or if a quartet is used to obtain energy variability, then the crab cavity will be placed between the second and third element of the quartet so that the aperture of the cavity better conforms to the apertures of the quadrupoles. The crab cavity can be six $\lambda/3$ X-band cells with a total length of 5.2 cm.

This section concludes with a discussion of the effects of the detector solenoidal field including the skew aberrations, steering and dispersion effects, and synchrotron radiation.

11.7.1 Determination of Crossing Angle

The crossing angle must be large enough so that:

- There is no multibunch instability from parasitic beam-beam interactions.
- There is physical room for the final doublet and the exit beam port.
- The big bend angle provides sufficient muon protection for the detectors.
- The IPs are sufficiently separated to provide radiation and ground noise isolation.

The last two items are necessary only for beam-line geometries that interrelate the big bend angle with the crossing angle, as in the geometry we have chosen.

The crossing angle should be chosen as small as possible to minimize:

- The impact of the solenoidal field on steering, dispersion, and synchrotron radiation.
- The power requirements of the crab cavity.
- The voltage and phase stability tolerances of the crab cavity.

- The length of the big bend (especially at 750-GeV beam energy).

The NLC crossing angle parameter has been chosen to be $\theta_c = 20\,\text{mr}$ and the corresponding big bend angle is 10 mr (Figure 11-46).

11.7.2 Crab Cavity

The purpose of the crab cavity is to provide a transverse displacement within the bunch that varies with longitudinal distance according to

$$\frac{\partial x^*}{\partial z} = \frac{\theta_c}{2} \quad . \tag{11.125}$$

Using $\Delta x^* = R_{12}^E \Delta x'_E$ and $\Delta x'_E = \text{eV}/E_0$, we find

$$\frac{\theta_c}{2} = \frac{\partial x^*}{\partial z} = R_{12}^E \frac{e\,\partial V(z)}{E\,\partial z} = R_{12}^E \frac{\text{eV}_{\text{max}}}{E} k\cos(kz) \approx R_{12}^E \frac{\text{eV}_{\text{max}}}{E}\frac{2\pi}{\lambda} \tag{11.126}$$

where in the last step we assumed the bunch was traveling through the cavity at the zero of the rf wave. R_{12}^E, which is the R_{12} matrix element from the entrance of the doublet to the IP, can vary somewhat with doublet design and machine energy, but it is in the range of 10 m to 20 m.

Crab Cavity Parameters

We have chosen the rf wavelength at X-band to keep the power requirements low and the structure length short. Table 11-38 compares X-band with S-band [Wilson 1995]. The parameters of the S-band cavity are "LOLA-III" taken from SLAC Report 17 (1963) and SLAC-PUB-135 (1965). The X-band parameters we will refer to as LOLA-X. Lolita-X is six cells of LOLA- X.

The required values for V_{max} and corresponding power are given in Table 11-39. We see that for the LOLA-X structure, the required input power is close to 0.25 MW for all energies. Since this is a very small power, the structure can be substantially decreased in length. The X-band structure with six cells (Lolita-X, 5.23-cm long) would suffice if driven with 7.4 MW of power. Shortening the structure also greatly reduces wakefield effects (almost as the square of the length).

A potential problem with the X-band structures is the iris aperture of radius 5.9 mm. We have chosen a somewhat larger Q2 aperture, to minimize resistive-wall wakes as well as provide a large beam stay-clear. At 250-GeV beam energy, the 23-σ_x aperture at the end of Q2 (see Figure 11-14) is about 11 mm for the parameter set with an 8-mm β_x^*. However, it is also possible to place the crab cavity between Q1 and Q2. For this parameter set R_{12} is still about 8 m, due to the space occupied by the superconducting Q1.

Crab Cavity Tolerances

We discuss five tolerance issues:

- Crab cavity voltage stability.
- Crab cavity pulse-to-pulse phase difference jitter.

Parameter	Symbol	Unit	Scaling	S-Band LOLA-III	X-band LOLA-X	X-band Lolita-X
Transverse shunt impedance	$r_\perp$	$\mathrm{M}\Omega/\mathrm{m}$	$\omega^{1/2}$	11.7	23.4	23.4
Structure length	L	m	$\omega^{-3/2}$	3	0.375	0.0523
Attenuation	τ	none	1	1.04	1.04	0.15
Iris diameter	$2a$	mm	ω^{-1}	47	11.8	11.8
Q value	Q	1	$\omega^{-1/2}$	11,000	5,500	5,500
Group velocity	v_g/c	1	1	−0.0078	−0.0078	−0.0078
Fill time	TF	ns	$\omega^{-3/2}$	1,280	160	22
Number of $\lambda/3$ cells	N_c	1	$\omega^{-1/2}$	86	43	6
Deflection volts/$\sqrt{\mathrm{MW}}$	$p_\perp c/\sqrt{P_0}$	$\mathrm{MeV}/\sqrt{\mathrm{MW}}$	$\omega^{-1/2}$	5.3	2.65	0.55

Table 11-38. *Scaling of crab-cavity parameters with wavelength.*

Beam Energy (GeV)	R_{12}^E (m)	θ_c (mr)	V (MV) S-band	V (MV) X-band	Power (MW) LOLA-III	Power (MW) LOLA-X	Power (MW) Lolita-X
250	8.9	20	4.8	1.2	0.8	0.20	4.8
500	17	20	5.0	1.3	0.9	0.24	5.6
750	21	20	6.0	1.5	1.3	0.32	7.4

Table 11-39. *Required values of maximum crab-cavity voltage for three energies.*

- Crab cavity phase difference short-term stability (1 hour).
- Beam phase jitter.
- Crab cavity alignment.

Crab Cavity Voltage Stability. If the crab-cavity voltage changes, the "crabbed" angle changes. A 2% luminosity loss would occur for a crab angle error of $\theta_d/5 \approx 0.2\sigma_x/\sigma_z \approx 0.6\,\mathrm{mr}$. For $\theta_c/2 = 10\,\mathrm{mr}$, the permissible voltage error would be $\Delta\mathrm{V}/\mathrm{V} = (\theta_d/5)/(\theta_c/2) \approx 0.06$. This estimate yields a permissible 6% voltage error.

Crab Cavity Phase Difference Jitter. If the phase of the crab cavity on the left jitters with respect to the crab cavity on the right, the beams will not collide head-on. A 2% luminosity loss occurs at $\Delta x = (2/5)\,\sigma_x$ (right beam minus left beam). Since $\Delta x = (\theta_c/2)\Delta z$, the allowed $\Delta z = (4/5)\,(\sigma_x/\theta_c)$, which for $\theta_c = 20\,\mathrm{mr}$ has the value $\Delta z = 12.8\,\mu\mathrm{m}$, corresponding at X-band to about 0.2°. At S-band, the number of degrees would be an intimidating factor of four smaller.

It should be noted that this tolerance is a phase difference tolerance between the two crab cavities. It is a jitter tolerance, since steering correctors would compensate for any error in crab cavity phase difference

which occurs for times greater than 1/5s. To minimize phase difference jitter, both cavities should be driven by the same klystron. The phase jitter for an X-band system in such a circumstance has yet to be experimentally determined, but a simple experiment which uses a phase mixer to measure the relative phase in two rf transport lines coming from a splitter, driven by one klystron, can provide the required information. (Note: It is routine in the SLC klystron gallery to compare the S-band phase of the output from a SLED-I cavity with the rf reference phase driving that klystron. There the phase jitter is known to be less than or the order of 0.1°.)

It is interesting to estimate the thermal change on the transmission line on one side that would be result in a 0.2° phase change. The phase change is given by $\Delta\phi = \Delta(k_g L) = k_g L(\Delta k_g/k_g + \Delta L/L)$. Using $k_g = 2\pi/\lambda_g = 2\pi/\lambda(1-y^2)^{1/2}$ where $y = 0.61\lambda/a$ we have $\Delta k_g/k_g = (y^2/(1-y^2))\Delta a/a$. And $\Delta a/a = \Delta L/L = \alpha\Delta T$ with $\alpha = 1.8\,10^{-5}$, leads to $\Delta\phi = 2\pi L/\lambda\alpha\Delta T(1-y^2)^{-1/2} = 2.7°$ per °C (taking $L = 10\,\mathrm{m}$, $\lambda = 0.026$). In other words, the temperature must be stable to 1/15°C in times of 1/5s, or drifts should be smaller than 20°C per minute. It is expected that the temperature will be more stable than this.

A concern is that the phase difference may creep, so that the steering corrector kicks become large. Phases should be monitored for this, and a system installed to maintain the phase of the rf reaching each cavity to be the same, and at a null at the beam phase. 0.5 degrees corresponds to a horizontal steering at the IP of 1 σ_x. Probably 40-σ_x steering, corresponding to a crab-cavity-to-beam phase difference of 20°, is possible before unacceptable dispersion is created. The phase of the cavity with respect to the beam can be monitored by observing steering of the beam at a post-IP BPM as the power of the cavity is ramped. There will be no steering if the beam-pulse arrives at the rf null.

Incoming Beam Phase Jitter. Although the crab-cavity jitter tolerance is only 0.2°, this does not imply a need for a corresponding phase stability of bunches in the incoming beam. The function of the crab cavity is to align the bunch for a complete collision no matter how long. Hence the limits on beam phase stability arise from considerations of luminosity loss due to hourglass effects at the IP. If we can have $\Delta z_{\mathrm{collision}} = \frac{1}{5}\beta_y^*$, and if the arrival of each beam is taken to be statistically unrelated, then we can have $\Delta z_{\mathrm{beam}} = \sqrt{2}(100)\,\mu\mathrm{m}$, which is about 1.9 degrees of X-band.

Crab Cavity Alignment for Multi-bunch Effects. If the beam passes through the cavity on-axis, no dipole mode is induced, even if the bunches are not at the zero phase. This is because the cavity does not put energy into the beam when steering it. So beam loading is entirely associated with misaligned cavities. We will assume a misalignment of 100 μm for numerical estimates.

Even if the cavity is misaligned, the induced fundamental field has a null at the longitudinal phase where the charge passes. As a result, if all particles are passing through the cavity at the same phase (modulo 2π), the induced field will change the slope of the field, but not its magnitude at the location of the bunches. We need an estimate for the change of the fundamental (dipole) mode and for the higher-order dipole modes which could steer the beam.

An estimate for the induced transverse field of 90 bunches traveling through the Lolita-X cavity off-axis by 100 μm can be found from the formula (P. Wilson)

$$V_\perp^{\mathrm{ind}}(n) = \frac{1}{2} I\, r_\perp\, L_s\, kx \left(\frac{1 - e^{-2\tau'} - 2\tau' e^{-2\tau}}{1 - e^{-2\tau}}\right) \quad \text{where} \quad \tau' = \tau\, Min\left[1, \frac{n\,\tau_{\mathrm{sep}}}{T_F}\right] \quad , \tag{11.127}$$

where n equals the bunch number and τ_{sep} is the separation time between bunches. The expression in brackets reaches its maximum value of 0.15 at $n = 16$. The corresponding maximum induced voltage is

2.5 kV. Dividing by 1.2 MV, the crab-cavity voltage at 250 GeV, gives $V_{\rm ind}/V_{\rm crab} = 0.002$, or 0.2% change. This is comfortably smaller than the allowed 6% calculated in Section 11.7.2.

A rough estimate for higher-order dipole modes is that they might be about 10% of the induced fundamental, or about 0.25 kV for Lolita-X. This voltage could cause a missteering at the IP of $R_{12}(\text{eV})/E = 17(9\ 10^3)/(500\ 10^9) = 8.3\,\text{nm}$ which is just about equal to $1/40\ \sigma_x$. This is also negligible.

Stability Tolerance from Energy Change. The Hamiltonian that gives the desired crab-cavity kick is of the form

$$H = \frac{\theta_c}{2R_{12}}\, x\, z \quad , \tag{11.128}$$

where the kick $\Delta x'$ is given by the derivative with respect to x, and has the linear z-dependence desired. However, there is also a change in energy implied by this Hamiltonian given by the derivative with respect to z. It is

$$\Delta\delta = \frac{\theta_c}{2R_{12}}\, x \quad . \tag{11.129}$$

The change in energy could be $\delta \approx 6 \times 10^{-6}$, which is also the stability tolerance of the final doublet. Inserting values for the crossing angle and R_{12}, we find the condition $x < 6\,\text{mm}$. This is of no concern as an alignment stability requirement. This number also places a limit on the beam size at the crab cavity. This limit is met since the beam size is on the order of $R_{12}\sigma_{x'} < 0.3\,\text{mm}$.

Feedback Systems

The physical quantities to be monitored and stabilized are i) the crab-cavity phases with respect to beam arrival times, ii) the crab-cavity voltage, iii) the crab-cavity phase difference between the two crab cavities, and iv) the crab-cavity alignment. The phase with respect to the beam, as pointed out above, can be checked by changing the crab voltage and looking for steering in the outgoing beam. The voltage itself must be set to minimize the horizontal beam size, which one should be able to monitor by vertical and horizontal beam deflection scans. It is equivalent to a beam size measurement.

By installing an insulated return line from each crab cavity to the tee (following the klystron) and comparing the phase of the two return lines, a phase correction can be introduced to keep the return phases stable. A schematic of this stabilizing system is depicted in Figure 11-54. The system is designed to keep large drifts from occurring in the phase difference. In addition, phase jitter must be small, and short-term phase stability will be taken care of by the IP steering feedback system.

11.7.3 Solenoid Effects

We will assume that the maximum solenoidal field strength will be 3 T, and that the maximum crossing angle to be considered is 30 mr. Under these assumptions, the maximum field perpendicular to the beam trajectory is $B_\perp = 450\,\text{Gauss}$. For this field, $B_\perp/B\rho = \{0.5 \times 10^{-4}, 0.25 \times 10^{-4}, 0.2 \times 10^{-4}\} m^{-1}$ at $E = \{250, 500, 750\}\,\text{GeV}$ respectively.

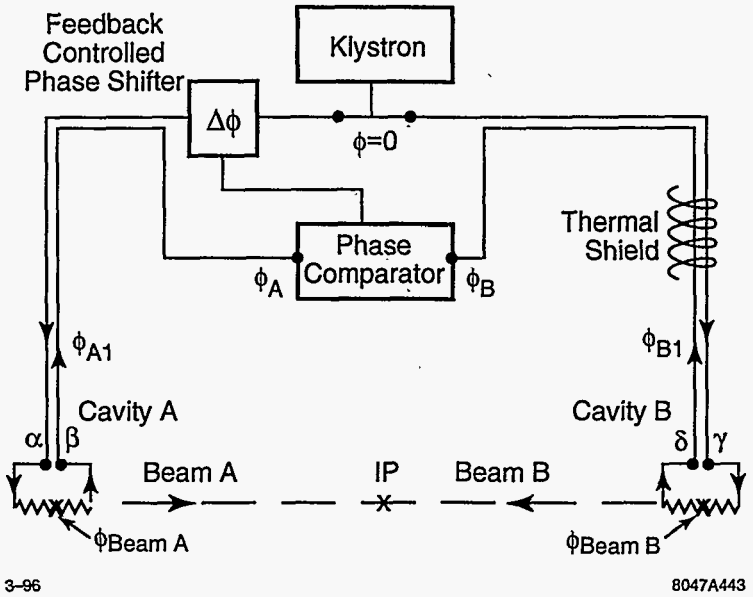


Figure 11-54. *Schematic of a system which stabilizes the phase difference between the two crab cavities.*

Skew Quadrupole Aberrations

The skew quadrupole term in the solenoid Hamiltonian is $1/2\,k_s(xp_y - yp_x)$ with $k = B_s/B\rho$. If there were no quadrupole within the solenoid, then this term is a constant (by rotational symmetry) and can be evaluated at the IP, to give an integrated beam-line generator of

$$G = LH = \frac{1}{2}\frac{B_s L}{B\rho}\left(x^* p_y^* - y^* p_y^*\right) \tag{11.130}$$

giving $\Delta y^* = (B_s L_s/B\rho)\,x^*$. Taking $B_x = 3\,\mathrm{T}$ and $L_s = 4\,\mathrm{m}$ leads to $B_s L_s/(B\rho) = \{0.014,\ 0.007,\ \text{and}\ 0.005\}$ for $E = \{250,\ 500,\ \text{and}\ 750\}\,\mathrm{GeV}$. Since $\sigma_x^* = \{100,\ 160,\ 160\}\ \sigma_y^*$, $\Delta y_{\mathrm{rms}}^*/\sigma_y^* = \{0.7,\ 0.6,\ 0.4\}$. So this term of the generator is small, but still too large by a factor of about five, and must be compensated with a skew quadrupole. We note that it is not the generator $p_x p_y$, which is the main term generated and corrected in the final-focus system. This skew quadrupole may have to precede the final-focus system, or be obtained from two skew quadrupoles on the opposite sides of an IP image in the final-focus system.

The interaction of Q1 and the solenoid will also produce the skew quadrupole generator $p_x p_y$, which may readily be compensated by the final-focus system skew quadrupole.

Steering and Dispersion

The solenoidal field in the region between Q1 and the IP will have a component $B_s\theta_c/2$, and will steer the beam by $\Delta y^* = B_s L^*/B\rho\,\theta_c/2$. For $\theta_c = 20\,\mathrm{mr}$, this has the value $\Delta y^* = \{70,\ 35,\ \text{and}\ 25\}\,\mu\mathrm{m}$ at $E = \{250,\ 500,\ \text{and}\ 750\}\,\mathrm{GeV}$. This will also be the magnitude of the dispersion generated.

The two beams are deflected apart so steering must be introduced to correct this displacement. If this is done with a corrector at the Q2 end of the doublet, then because (see Table 11-36)

$$\frac{\partial \eta_1}{R^E_{34}\partial y'_E} = -2.3 \quad , \tag{11.131}$$

the dispersion generated by the doublet adds to the dispersion generated by the solenoidal field, giving a net dispersion 3.3 times the amount from the solenoidal field alone. This is a huge dispersion, and attempts to correct it will create unacceptable second-order dispersion terms and synchrotron radiation.

The only available alternative is to displace Q1 by $\Delta y_{Q1} = -\Delta y^*/1.6$ or to introduce an equivalent dipole field in the region of Q1. This will create dispersion which has the opposite sign but does not totally cancel the original. Theoretically, a small section at the IP end of Q1 could be displaced to re-steer the beam, in which case the dispersion would almost totally cancel. This does not work because the synchrotron radiation in such a corrector would be unacceptable. The remaining dispersion is $(1.63 - 1.33)/1.63 \approx 1/5$ of the original (from Table 11-36: $\Delta\eta = -1.33\Delta y_{Q1}$).

The remaining residual dispersion of $\{21, 11, \text{and } 9\}\,\mu$m at $E = \{250, 500, \text{and } 750\}$ GeV can be corrected by displacing the incoming beam by $\{11, 5, \text{and } 4\}\,\mu$m. This is at the edge of acceptability vis-à-vis the generation of synchrotron radiation. For 20-mr crossing angles, these numbers are reduced by a factor of 1.5, and are acceptable.

Synchrotron Radiation

We first calculate the additional synchrotron radiation created by the solenoid for the on-axis particles. Then we calculate the effects of the increased dispersion function on the final doublet synchrotron radiation for focused particles (Oide effect + solenoid).

Solenoid Field Alone. The $R_{36}(s)$ function generated by the solenoid in the L^* region is given by

$$R_{36}(s) = \int R_{34}(s)d\theta = \int R_{34}(s)\frac{B_s\frac{\theta_c}{2}}{B\rho}ds = \frac{B_s\theta_c}{4B\rho}s^2 \quad . \tag{11.132}$$

The synchrotron radiation from this region results in

$$\frac{\Delta y^2_{\rm IP}}{\sigma^{y2}_{\rm IP}} = \frac{c_u r_e \lambda_e \gamma^5}{\sigma^{y2}_{\rm IP}}\int ds\, R_{36}(s)^2\left|\frac{1}{\rho(s)^3}\right| = \frac{c_u r_e \lambda_e \gamma^5}{\sigma^{y2}_{\rm IP}}\int ds\left(\frac{B_s\theta_c}{4B\rho}s^2\right)^2\left|\frac{B_s\theta_c}{2B\rho}\right|^3 = \frac{1}{20}\frac{c_u r_e \lambda_e}{\sigma^{y2}_{\rm IP}}\left(\frac{B_s\theta_c L^*\gamma}{2B\rho}\right)^5 \tag{11.133}$$

Assuming $L^* = 2$m, $B_s = 3$T, and $\theta_c = 40$mr, $[\Delta y_{\rm IP}/\sigma_{y^*}]^2 = \{0.04, 0.06, 0.06\}$. Fortunately, this term goes like $(B_s\theta_c)^5$, and can be reduced by reducing the crossing angle or the solenoidal field. For $\theta_c = 30$mr, the coefficients are reduced by a factor of 7.6, and $[\Delta y_{\rm IP}/\sigma_{y^*}]^2 = \{0.005, 0.008, 0.008\}$. These are already acceptable values.

The contribution from the compensation of the solenoidal steering that lies within Q1 must be added. Here the value of $1/\rho$ is smaller and the R_{36} function from the solenoid plus quadrupole is becoming smaller. This term adds an additional 15% to the above coefficients (at $E = 250$ GeV).

Solenoid Field Plus Final Doublet. The effect of the synchrotron radiation (Oide effect) in the quadrupole is enhanced by the presence of the solenoidal field because of the increased dispersion function to the IP. This will be negligible at a 20-mr crossing angle.

Tolerance on Solenoidal Field Strength Stability

If the solenoid field were to change, the orbit to the IP would change and the beams would not collide. With the beam separation proportional to the solenoidal field and between 200 and 80 μm, and a tolerance on the beam separation for a 1/2% luminosity loss of about 2.4 nm, the field must be stable to one part in 10^5. Because the effect is a steering effect, this is a jitter tolerance and should be met quite easily.

11.7.4 Summary

From the point of view of the requirements on the crab cavity and the solenoidal field, the considerations of this section suggest that the 20-mr crossing angle is acceptable. This allows us to couple the big bend angle (used to obtain IP separation and reduce muon backgrounds) with the crossing angle, and to have the space to use permanent magnets for the first entrance and exit quadrupoles. The permanent magnet exit quadrupole is helpful by providing focusing close to the IP (which is necessary because of the large exit beam divergence) securing the high-quality beam removal optical path crucial for post-IP diagnostics. Permanent magnet fields can be quite stable under thermal change.

It is an important result that the first quadrupole must be displaced by up to 45 μm when the solenoidal field is turned on to correct the dispersion. This is a large movement compared to the nanometer stability required for this quadrupole under colliding conditions.

An experiment with an X-band phase mixer should be undertaken to confirm the possibility of maintaining a 0.2° relative phase tolerance between the two crab cavities.

11.8 The Beam Extraction and Diagnostic System (The Dump Line)

The main functions of this system are to optimize luminosity, to characterize beam properties at the interaction point, and to transport beams from that point to a dump with minimal backgrounds in the detector from these or any secondary functions that may occur. Beam characterization includes the measurement of current, position, profile, polarization, energy, and low-order correlations on a bunch-to-bunch basis for feedback and stabilization. A magneto-optical and diagnostic layout is described that provides such functions. We also consider the possibilities for e, μ, n, and γ secondary lines and dump experiments as well as energy recovery and local reuse of an assumed 10 MW of power in each 500-GeV beam. An earlier condensed version of this section was given in [Spencer 1995b].

To accomplish all of this, we need to know the detailed composition and characteristics of the outgoing beam(s) under different, possible circumstances. Clearly, these characteristics depend on those assumed for the incoming beam. Based on an assumed set of incoming beam parameters and transport optics we then make predictions for the outgoing beams that are used to guide the design of the outgoing, dump line optics. Typically, these beams have a significant number of pairs and more photons than leptons. Thus, if background simulations for the detector then imply an unacceptable situation, this procedure has to be iterated until a consistent solution is achieved.

Our approach was to develop a "straw man" design, *i.e.,* to assume a set of upstream characteristics that are reasonable but conservative to provide a worst-case scenario that emphasizes the major problems in constructing the transport system. The current design is a reasonably versatile one that takes the primary, outgoing disrupted beam (including photons) from the IP to the dump with minimal loss of particles. We have also included the low-energy pairs that are produced during the strong beam-beam interaction. We have found no basic limitations on the achievable luminosity imposed by the dump line. Furthermore, it provides a range of configurations that allow a variety of options.

While our primary goal was to optimize the luminosity while disposing of the various secondary beam components, it is also important to provide any monitoring and feedback that can optimize the average, usable collision rate at the IP. Thus, beyond simply dumping the beam, there are other functions that run from the necessary to the desirable that will be considered a part of the dump line:

- Beam control and stabilization,
- Diagnostics and monitoring - including luminosity and polarization,
- e, μ, n, γ secondary beams and parasitic experiments,
- Polarized sources for γ, μ and e^+ beams, and
- Energy recovery.

Due to the high power in the outgoing photon beam and the cost of beam dumps, we decided that the photons and primary leptons should share a common beam dump. This implies an available distance for beam studies of $\approx$150 m before the outgoing, high-energy photons begin to cause problems for the present dump design.

Because the SLC was a prototype for the NLC, we begin by reviewing the SLC and the FFTB experience relative to the NLC design. Next, we summarize the more important working assumptions in terms of a table of parameters characterizing the input and the output optics and beam characteristics at the IP. Next we describe the design procedure and give some intermediate results that were used to arrive at the final design. A horizontal chicane is used to separate the outgoing photon and electron beams, which are comparable in number and density, to allow separate experiments, before both beams are recombined into a common dump. Some rather sophisticated measurements and capabilities are possible over the available distance that are compatible with a low-background dump. We have also tried to accommodate the possibility of beam dump experiments. However, various additional simulations need to be done to confirm this and certify specific designs. Certain hardware is discussed where relevant. Finally, we summarize our conclusions and discuss the work that still needs to be done.

11.8.1 Comparison with the SLC

One advantage for the NLC dump design relative to that for the SLC is a horizontal crossing angle ($\theta_{c,x}$=2×10 mr) at the IP that allows us to avoid kickers and septa for separating the counterpropagating beams. A related but more subtle advantage of this is that it allows us to reverse the polarities between the ingoing and outgoing IP doublets for better control of the larger outgoing horizontal disruption angle and emittance.

The much smaller vertical spot sizes assumed for the NLC overcome other differences to make its disruption larger. Although this complicates the disposal problem, it still allows all of the SLC measurements to be done

as well as others that may not be practical there. For example, we can measure the energy loss distribution and use precise rf BPMs in the beginning of the line to measure position, angle, and timing of individual bunches where this may be impractical for the SLC due to its larger peak deflection angles that produce rms outgoing angular spreads that are four times larger than our worst case.

However, the outgoing, disrupted energy spreads and energy offsets calculated for the NLC are more than a factor of twenty larger. The outgoing beam also includes a significant number of synchrotron radiation photons that can produce low-energy pairs quite apart from the beamstrahlung. Because the emittance and disruption angle is largest in the horizontal, this implies that we want our analyzing magnets to bend horizontally which should also simplify any beam dump experiments.

Rather than 30 kW in each SLC beam, one has to deal with nearly 10 MW in each NLC beam. There is a factor of ten in energy for 500-GeV beams in the NLC, the same rf pulse repetition rate of 120/s, and a factor of twenty or so in beam current per rf pulse from accelerating a multibunch train in each pulse. This increased beam power poses certain problems for intercepting detectors and implies significantly higher operating costs. For 10¢/kWh, this represents a potential refund of as much as $45K/day if energy is restored to the grid or otherwise recycled before it is dissipated in the dump.

11.8.2 Basic Design Procedures, Constraints, and Assumptions

The optics can only be realized after we know the characteristics of the outgoing beam from the IP. The procedure was to take the upstream final-focus design in the form of TRANSPORT [Brown 1977] and convert this to DIMAD [Servranckx 1990] for predicting the spot characteristics at the IP due to emittance growth from synchrotron radiation in the dipoles and quadrupoles. One assumption was that the effect of the energy loss on the optics, especially in the quadrupoles, can be neglected. The predicted beam parameters resulting from DIMAD were then used in ABEL91 [Yokoya 1986] to predict the composition and characteristics of the outgoing beam as well as to produce ray sets for all particle types for tracking. The ABEL results were checked against analytic calculations where possible and used to simulate the dump line with TRANSPORT (for design) and TURTLE (for tracking) [Carey 1982]. Table 11-40 gives several configurations that were used with those labelled 'A' the current ones for 250- and 500-GeV incident beams.

In the Table, $\mathcal{L}_{e^\pm}$ and $\mathcal{L}_{e^-e^-}$ are the predicted incident channel luminosities for these parameters whereas $\mathcal{L}_{\text{Compton}}$ and $\mathcal{L}_{\gamma\gamma}$ are secondary to the e^+e^- channel. $\mathcal{L}_{\gamma\gamma}$ is rather large because it includes contributions from real and virtual photons and the beamstrahlung parameter Υ is fairly large. Υ is an important measure of the QED and QCD backgrounds expressed in terms of $\gamma\gamma \rightarrow e^+e^-$ pairs and minijets ($\gamma\gamma \rightarrow$X + anything). Such parameters are described in more detail below. Because there is still debate on how to calculate the hadronic backgrounds in terms of the various possible quark and gluon contributions at these energies, they are not listed. There is interesting physics available here if it could be measured but an upgraded SLC would be a better place to find new hadrons, quark molecules or glueballs.

Because electrons and positrons are generally interchangeable, we will often talk about electrons but all parameters such as N_B, the number of electrons (or positrons) per bunch apply equally to both beams unless otherwise stated. In fact, N_B can also be the number of photons per bunch [Spencer 1995c] in a general linear collider (GLC). The last column for 500-GeV beams was our primary case for the optics design. In the absence of other input, we used only the **disrupted** beam from this worst case having the lowest $\beta^*_{x,y}$ and highest Υ. The numbers in the table and the ray sets that were used were predicted using a modified version of ABEL that was made operational on the SLAC VM system. Calculations were done for all of the columns shown and the output was archived for general use.

$\sqrt{s}_{\rm NLC}$	(GeV)	500	500A	1000	1000A	'Worst'
$f_{\rm rep}$	(Hz)	180	180	120	120	120
n_B		90	90	75	90	75
$\gamma\epsilon_y$	(10^{-8} m)	5	8	5	10	5
N_B	(10^{10})	0.65	0.65	1.1	0.95	1.1
σ_x^*/σ_y^*	(nm)	319.7/3.2	285.9/4.52	360/2.3	226.1/3.57	245/2.5
σ_z	(μm)	100	100	100	125	125
β_x^*/β_y^*	(mm)	10/0.1	8/0.125	25/0.1	10/0.125	10/0.1
$\mathcal{L}_G$	(10^{34}cm^{-2}sec^{-1})	0.53	0.42	1.05	0.96	1.42
D_x/D_y		0.073/7.3	0.090/5.70	0.049/7.60	0.132/8.33	0.104/10.2
$\theta_D = \theta_{x,\max}$	(μrad)	233	257	175	238	256
Υ		0.09	0.10	0.27	0.30	0.32
$\mathrm{H}_D \equiv \mathcal{L}/\mathcal{L}_G$		1.37	1.42	1.30	1.36	1.56
$\mathcal{L}_{e\pm}$	(10^{34}cm^{-2}sec^{-1})	0.73	0.60	1.37	1.30	2.20
$\langle s\rangle/s_{\rm NLC}$			0.972		0.913	
$\langle \mathrm{E}_e - \mathrm{E}_o\rangle/\mathrm{E}_o$		0.019	0.032	0.0711	0.103	0.127
$\delta_B \equiv \delta\mathrm{E}_e^{\rm rms}/\mathrm{E}_o$		0.042	0.065	0.120	0.143	0.167
$\mathcal{L}_{100}/\mathcal{L}_{e\pm}$		0.441	0.376(0.405)	0.305(0.323)	0.195(0.218)	0.182(0.206)
$\mathcal{L}_{90}/\mathcal{L}_{e\pm}$		(0.83)	(0.80)	(0.84)	(0.75)	(0.69)
N_γ/N_B		0.88	0.98	1.21	1.67	1.77
$\langle \mathrm{E}_\gamma\rangle/\mathrm{E}_o$		0.027	0.033	0.059	0.062	0.072
N_p/N_B	(10^{-7})[a]		3.95		8.69	
$\langle \mathrm{E}_p\rangle$	(MeV)[a]		23.4		25.8	
$\mathcal{L}_{\rm Compton}$			0.23		0.80	
$\mathcal{L}_{\gamma\gamma}$	(10^{34}cm^{-2}sec^{-1})		0.10		0.56	
$\mathcal{L}_{e^-e^-}$			0.23		0.45	

[a] N_p/N_B is the ratio of outgoing positrons per incident electron for energies $\mathrm{E}_p > 2$ MeV and angles $\theta \gtrsim 10$ mr.

Table 11-40. *Beam-beam effects for e^+e^- at $E_{\rm cm} = 0.5$, 1.0 TeV and $\gamma\epsilon_x = 5\times10^{-6}$ m. Quantities enclosed in (...) are calculated analytically rather than from the simulations.*

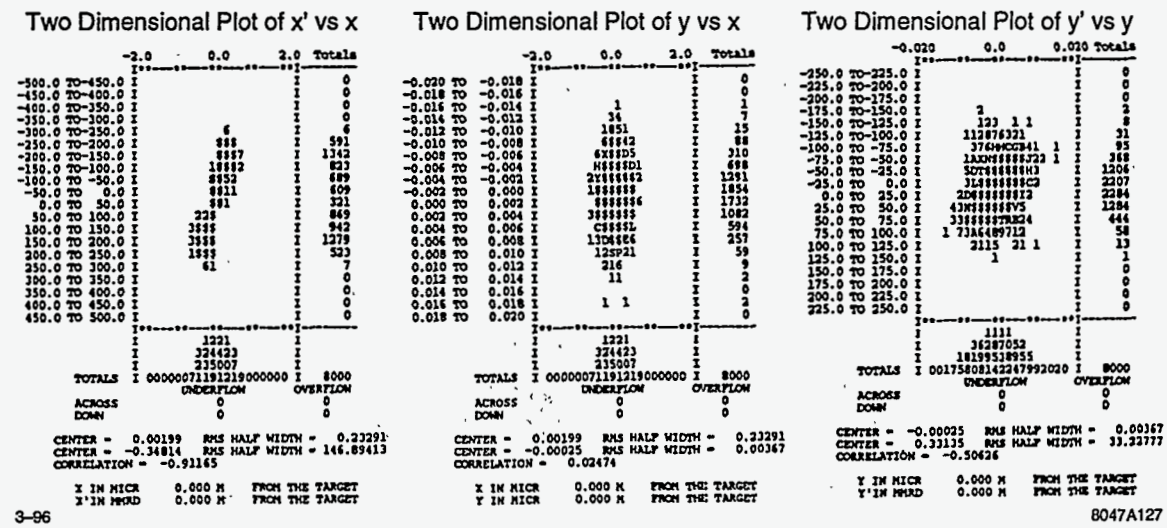


Figure 11-55. *Disrupted beam distribution at the IP for the 'Worst' case in Table 11-40*

The length of the extraction line was set by the outgoing beamstrahlung photons and their predicted size at the dump:

$$12\sigma_{x'}^{\gamma} \times L_D \leq 2R_D \ , \tag{11.134}$$

where L_D is the distance from the IP to the dump face and R_D is its radius. For a dump window of 20 cm in diameter and outgoing angular spreads of $\sigma_{x',y'}$=92,43 μr, taken from Figure 11-58, we have $\approx$150 m of usable length available. The apertures of the outgoing magnets were sized to pass these photons (and pairs) outside the detector so that the outgoing photons pass through the system into the dump without serious loss.

11.8.3 The Beam-Beam Calculations

Histograms of our initial working set of 8000 disrupted particles are shown in Figures 11-55 and 11-56. From these figures we see the shearing of the upright ellipses from the focusing by the other beam just after the interaction as well as emittance growth in both transverse and longitudinal directions. However, the energy loss and spreading with a very long tail extending nearly down to the Compton edge (50 GeV for 500-GeV beams and a Nd:YAG laser) is probably the most significant effect. These and the subsequent histograms based on them were taken directly from third-order TURTLE. The beamstrahlung distributions that are shown next were also done with TURTLE for uniformity.

Figures 11-57 and 11-58 show the outgoing photon ray set for the same input beam. From the Table there are actually more outgoing photons than electrons but a thinning factor was used. When the configurations are frozen we will enlarge the ray sets and also characterize them analytically. The average photon energy and distribution looks very much like a bremsstrahlung spectrum and relates directly to that for the outgoing disrupted electrons.

The effective, transverse photon emittances are typically smaller than the disrupted, charged particle beam because of the differing effects of disruption. Although this is discussed later, we note that the photon spot size provides a better, realtime measure of the incoming, unperturbed beam profile than the disrupted

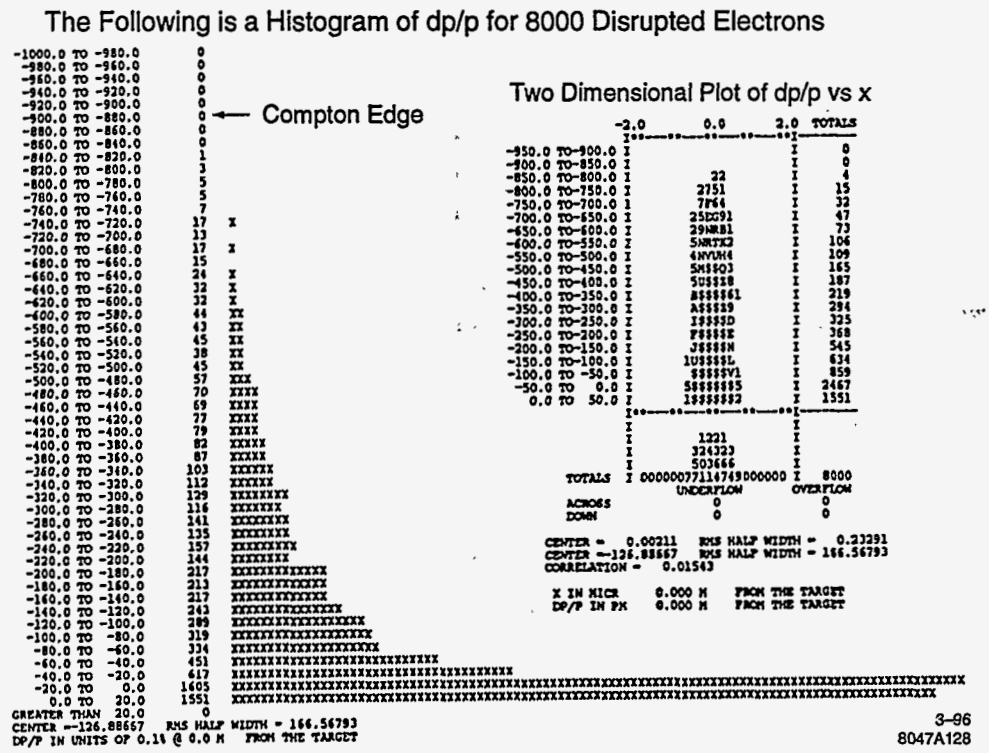


Figure 11-56. *Disrupted energy distribution at the IP in units of 0.1% for the 'Worst' case in Table 11-40. The insert shows the correlation with the horizontal position x.*

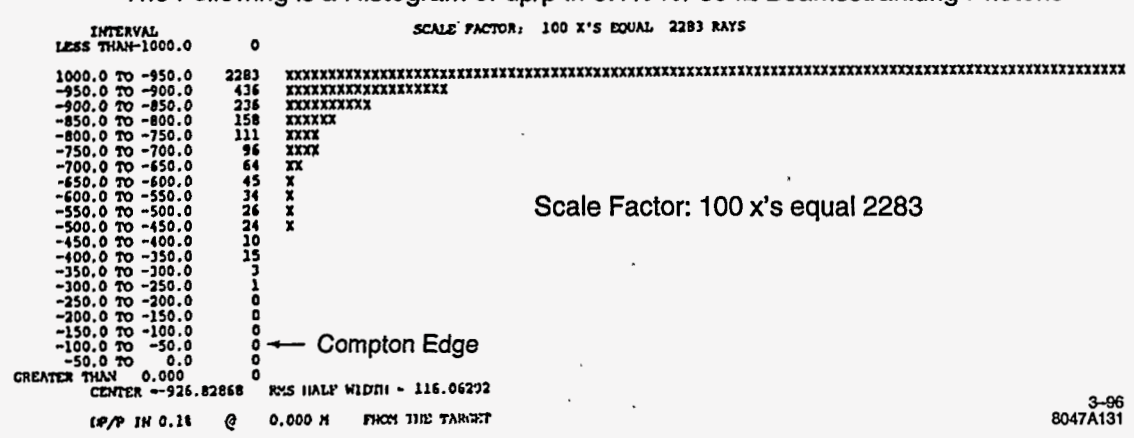


Figure 11-57. *"Worst" case outgoing photon energy distribution from Table 11-40.*

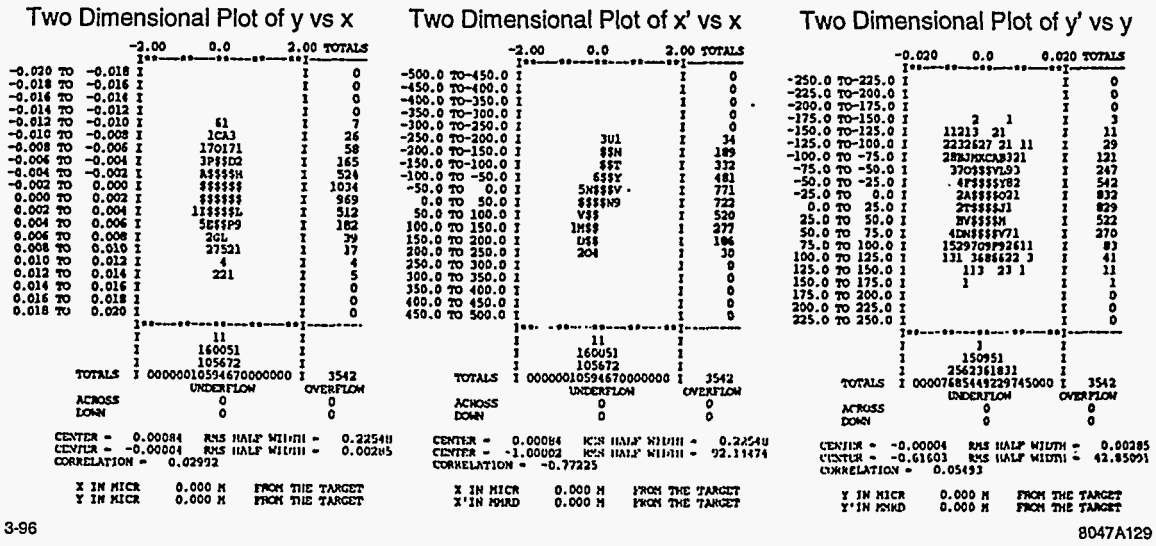


Figure 11-58. *'Worst' case transverse, outgoing photon distribution from Table 11-40.*

electrons. The outgoing photons were important in determining the length and acceptance of the extraction line because they can't be controlled in the same way as charged beams and because they have enough high energy photons to allow a good, nonintercepting profile monitor for bunch-to-bunch measurements outside of the solenoid. Figures 11-59 and 11-60 show the TURTLE output at various distances from the IP after tracking the beams through the quadrupoles inside the detector. From these figures and Figure 11-65 below, we see from the rms beam sizes and the numbers of particles outside the picture frame, that the disrupted primary beams and their associated beamstrahlung should clear the quadrupole apertures and solenoid with negligible loss.

From these figures, the primary disrupted beam is still the determinant of the magnet apertures but neither beam imposes any severe restrictions—at least while they are in the solenoid. The most damaging effect is then expected to come from the low-energy pairs. We have calculated their spectrum and detector simulations are now being done. This is complicated by the low energies, the crossing angles, and the strong solenoidal fields at the IP that necessarily couple the transverse coordinates.

Such beams necessarily experience and are influenced by nonlinear fields in both the transverse and longitudinal directions before exiting the solenoid. Of equal importance, the response of an outgoing pair to the collective beam field after they are produced depends drastically on the charge and energy of each component.

Figures 11-61–11-64 show some characteristic predictions for the Case A columns in Table 11-40 corresponding to cuts $10 \leq E \leq 500\,\text{MeV}$ and $\theta \geq 10\,\text{mr}$. The figures illustrate how the different species of outgoing particles complicate the design of the line—especially when it is desired to pass them without intercepting any (except via the required magneto-optical fields). While there are comparatively large angular and energy spreads between the incoming and outgoing electron and photon beams as shown above, the pairs are much worse and also include two opposite charge states. Figure 11-61 and 11-64 show this when we allow the pairs to drift 10 m after leaving the solenoid. In these figures, we assumed two different solenoid lengths for comparison. The results can be compared to the previous estimate of the length of the extraction line based on the outgoing photon beam characteristics. Fortunately, the pairs do not change our previous design in any essential way as we will discuss.

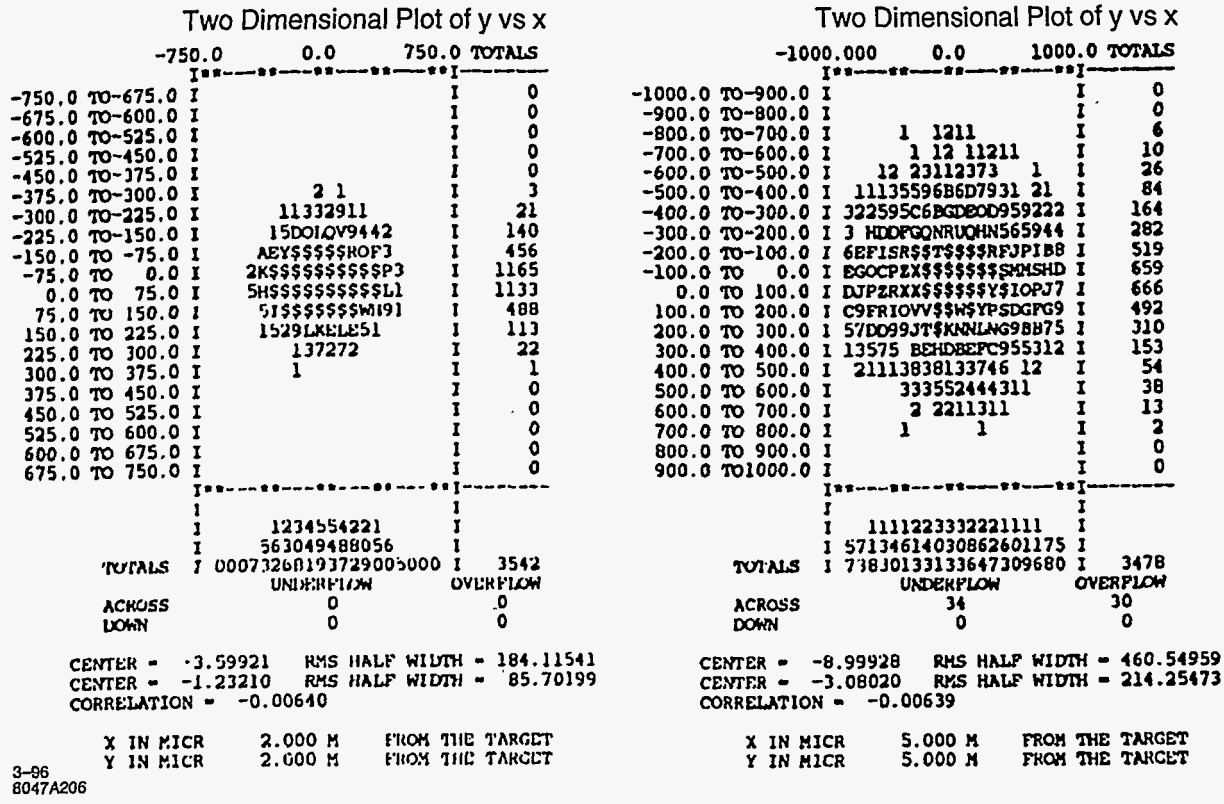


Figure 11-59. *Transverse, outgoing photon distribution at various distances from the IP in the solenoid (L_{sol}=5 m). The picture frame is ≥1/8 of the relevant quadrupole aperture.*

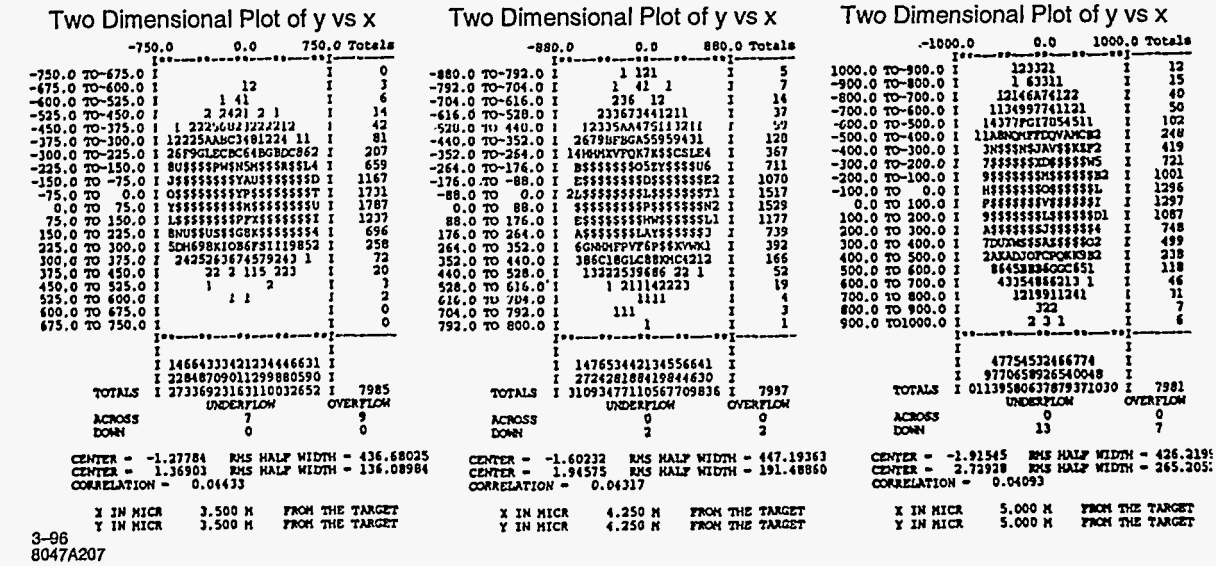


Figure 11-60. *Transverse, disrupted beam distribution at various distances from the IP in the solenoid. The picture frames are always ≥1/8 of the nearest quadrupole aperture.*

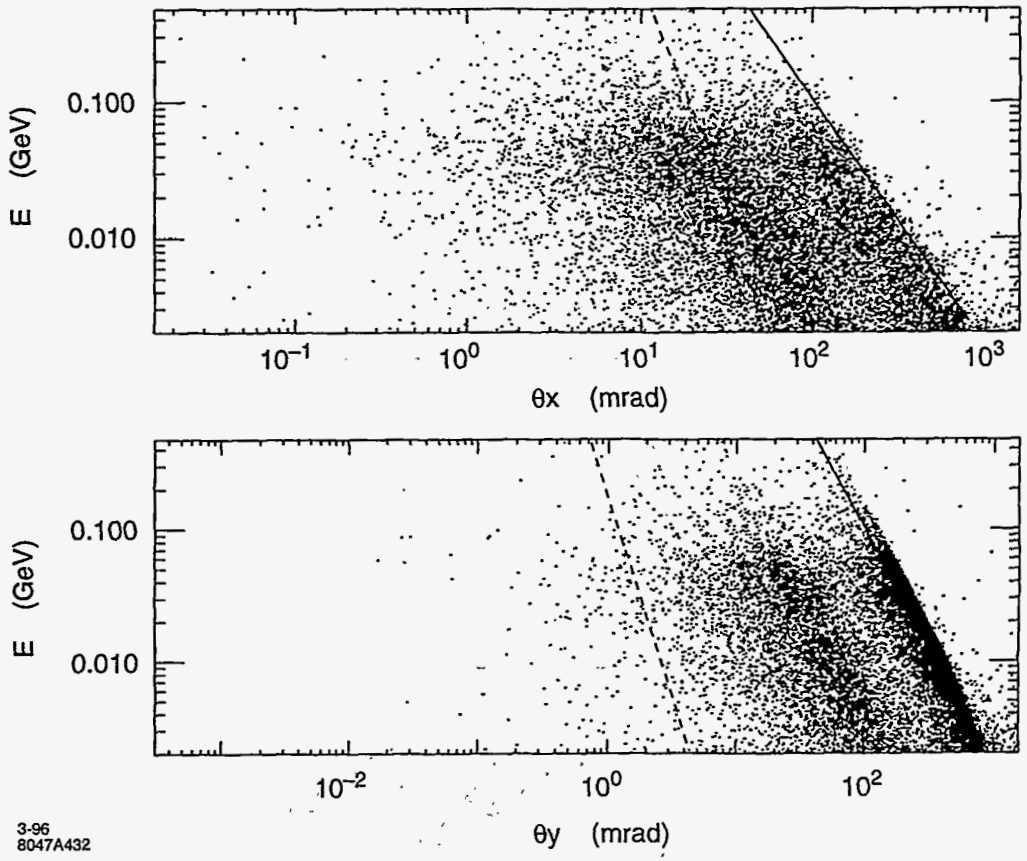


Figure 11-61. *Outgoing charged pairs at 1 Tev for Case 1000A in Table 11-40.*

Besides constraining our optical and hardware design options, this forced us to continually constrain the apertures as well as the field integrals during the design optimization to ensure viable magnets that allow all of these beams to pass. Other constraints were related to the requirements imposed by the different possible measurements and experiments that might be required. For example, the electron spin rotation and depolarization from the bends and the beam-beam interaction as well as the energy resolution necessary to resolve the low-energy electrons near the Compton edge (the most sensitive ones for monitoring beam polarization) constrained the strength and disposition of the dipoles. A related constraint was the need to capture off-energy bunches from the linac or ones that did not collide and therefore lose any energy from the strong beam-beam interaction. Lastly, because the pairs also carry information about the luminosity in terms of the beam sizes at the IP, we wanted to pass them outside the solenoid to make measurements on them before dumping them in separate, reentrant dumps.

The simplest procedure is to sample the higher energy part of the pair spectrum because this is the most accessible and easiest to measure with the necessary resolution. The pair dumps are modest in both size and cost as indicated in Table 11-40. The relative proportion of Breit-Wheeler, Bethe-Heitler and Landau-Lifshitz pairs depends on the cuts, but Breit-Wheeler production from two real photons is generally negligible at those energies.

We have laid out the essential assumptions and presented the results of calculations that provided the various input beams required for the detector simulations and the optics design. On the assumption that the detector

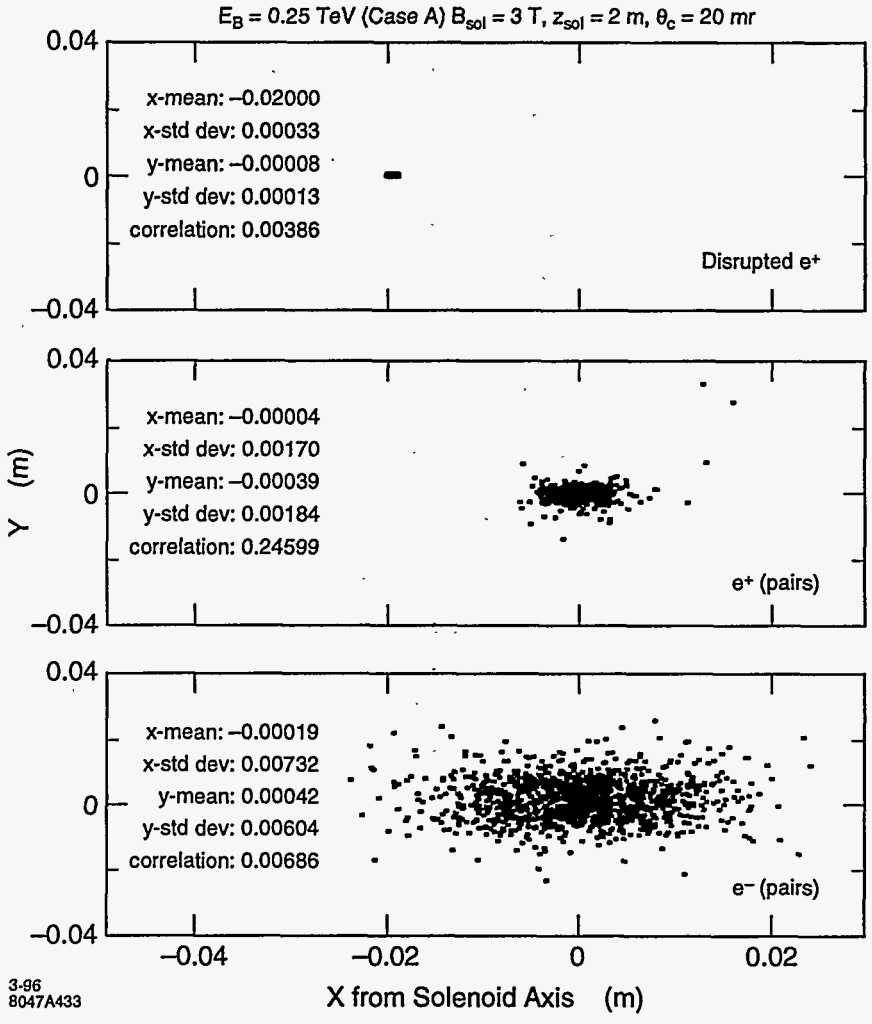


Figure 11-62. *Outgoing charged beams at 2 m from the IP in a strong solenoidal field.*

studies show no irreconcilable conflicts with the ingoing or outgoing IR optics, the problems to be solved become reasonably well defined.

For example, given that l^*=2 m, we showed what the distributions of the outgoing beam components look like along the solenoid axis, *i.e.*, what the masking, the magnets, and the diagnostics within the detector must look like to accommodate such beams and provide the required beam stay-clear. For these outgoing beams, we anticipated the optics design and showed that we can get all the beam components out of the solenoid with virtually no loss if we can design rather strong quadrupoles with an open structure that can achieve their assumed characteristics in a strong solenoidal field (≈3 T). A stronger solenoid for containment of the lower energy pairs appears desirable but this depends on the topology of the aperture, the crossing angles, etc. Ultimately, it is a question of the cost of the ideal structure *vs.* the practical question of what the detector design can accommodate. The next section gives the complete optics design from the IP to the dump that allows a variety of configurations and hardware for the options discussed in the introduction. Subsequent sections discuss these options and their hardware.

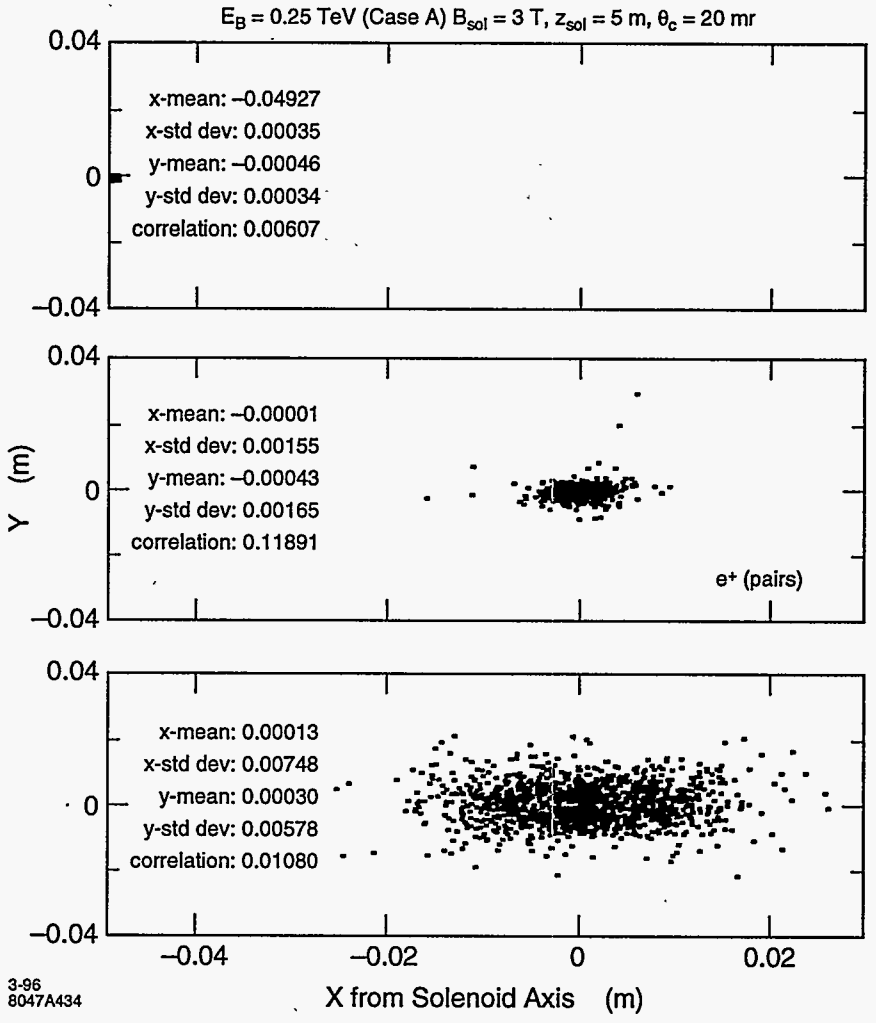


Figure 11-63. *Outgoing charged beams at 5 m from the IP in a strong solenoidal field.*

11.8.4 Beam Line Optics

In the SLC, the angular distribution predicted for the outgoing disrupted beam due to the strong focusing from the counterpropagating beam at the IP was a major factor in sizing the apertures of the strong IR quadrupoles. This was $\approx 10\ \sigma_x^*$, for round beams which translates to $\approx 10\ \sigma_x$ at the exit of the last quadrupole adjacent to the IP. For the NLC, a closely analogous situation exists concerning the outgoing angles but perhaps the most important characteristic that dictates the optics is the energy spread in the outgoing beam. This also complicates the energy recovery for various methods such as rf deceleration where we might want the beam to pass through an accelerating structure where the energy-phase structure of the bunch at the cavities becomes important as does the transverse size. In the NLC, there will also be a significant number of lower-energy particles from beamstrahlung as well as pair production and the effective disruption parameter for such particles, that scales inversely with their energy, can become very large.

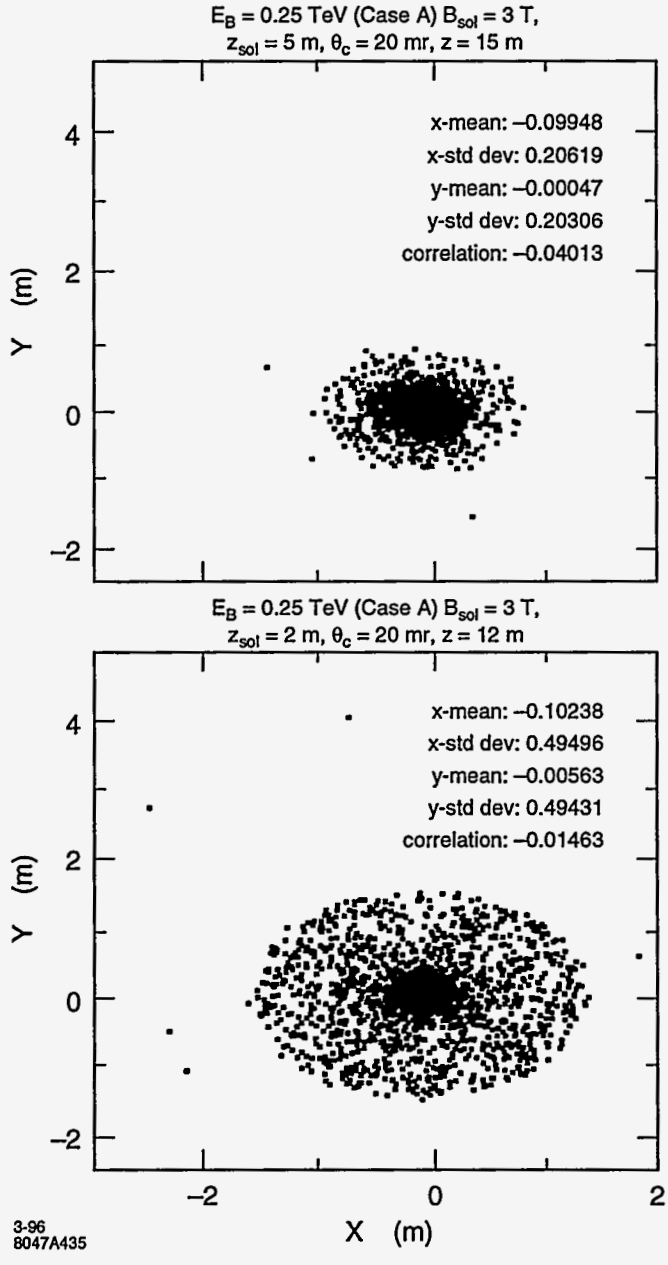


Figure 11-64. *Outgoing beams 10 m after a strong solenoidal field with $L_{sol} = 2$ and 5 m.*

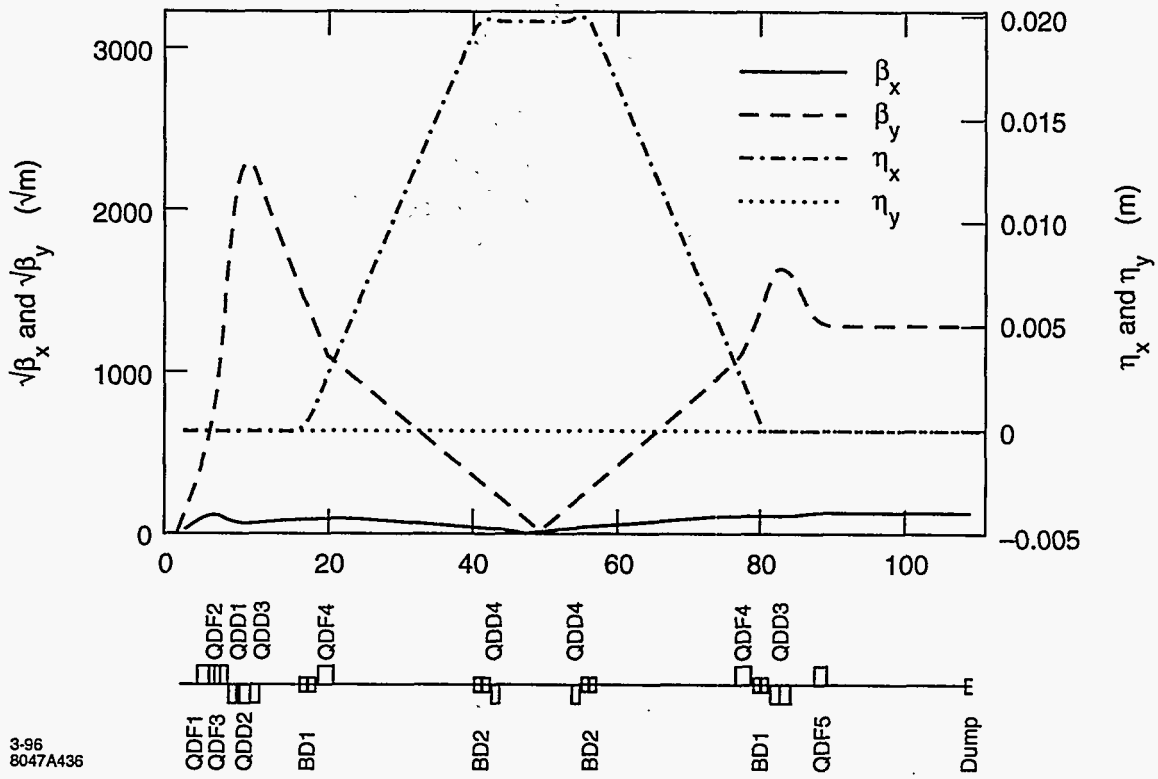


Figure 11-65. *Optical layout for the NLC dump line for the 'worst' case in Table 11-40.*

We note that the primary problem of transporting the outgoing beam to the dump is much simpler after it exits the detector. Figures 11-59 to 11-64 show the outgoing electrons, positrons and photons after various quadrupoles for various assumptions about the detector—typically assumed to end at 5.0 m from the IP. These results were computed through third-order with TURTLE and are dominated by the low order, chromo-geometric terms. In each Figure, the picture frame is ≤1/8 of the magnet aperture and the beam σ's are given in μm and μr. In all cases, the pole-tip fields and apertures of the magnets (including the beam dump) were sized to give $R_M \geq 12\sigma_{x,y}$ for the outgoing photons and disrupted electrons. In the initial doublet, that extends beyond 10 m from the IP, the apertures are $\geq 15\sigma_{x,y}$.

A number of configurations were tested and shown to be adequate under the stated assumption that we must take all photons and electrons to the beam dump with the stated size limit. Figure 11-65 shows one possible configuration that has space available for experiments, monitoring, and higher-order multipoles. It satisfies our criteria using only dipoles and quadrupoles. Further, some of the quadrupoles that are shown here were not excited for this particular optics. They act as potential tuning elements, *e.g.*, if QDD3 is turned off, the full energy, undisrupted beam will propagate to the dump in a point-to-parallel configuration. Lower energy particles will be overfocused so that the envelope of electrons tracked with TURTLE intersects that for the photons near 19 m from the IP. Thus, our first space available for beam size and polarization measurements comes between 10.5–16.2 m.

The first dipole of the horizontal chicane, used to separate the outgoing photon and electron beams, begins at 16 m. It allows separate e and γ measurements or experiments before recombining both beams into a common dump at 110 m. The various bends in the chicane can serve as spectrometers as well as the chicane

itself. Figure 11-65 shows the Twiss functions when the four bends are sized to separate the two beams by 12 sigma. Their maximum separation is

$$\Delta x = 2\rho_B(1-\cos\theta_B) + L_1\tan\theta_B \ , \tag{11.135}$$

where θ_B is the bend angle for the full energy of any one rectangular dipole of length $L_B{=}\rho_B sin\theta_B$ and L_1 is the separation between bends BD1 and BD2. Notice that this is just the dispersion η_x in the center of the chicane. This separation requires a minimum distance of

$$L_{min} = 2\rho_B \sin\theta_B + L_1 \ . \tag{11.136}$$

The change in the bunch separation, due to the chicane, after this point in the central region is

$$R_{56} = \frac{\delta l_z}{\delta p/p} = 2\rho_B(\tan\theta_B - \theta_B) + L_1\tan^2\theta_B \ . \tag{11.137}$$

This is proportional to the rf phase shift [Spencer 1995a]. Thus we have a common beam pipe and sufficient dispersion to measure the energy and spread of the undisrupted beam.

For example, if we want to use the first bend for analyzing low-energy particles from the IP or from a laser interaction before this bend, then the first-order resolving power for some downstream location L is

$$R_1(\rho,\theta,L) = \frac{\rho(1-\cos\theta) + L\cdot\tan\theta}{[x_i\cos\theta + x_i'(\rho\sin\theta + L/\cos\theta)]} \ . \tag{11.138}$$

In the middle of the chicane R depends on the optics we impose. $R_{12}{\to}0$ for point-to-point so $R{=}\Delta x/10\sigma_x^*{=}8000$ for a magnification of 10. This region of the chicane can then resolve a single-beam, undisrupted energy spread of δp/p=0.0125% while the region directly in front of BD2 gives $R{\approx}800$ or a 0.13% capability.

Notice that there were several factors that constrained the bends: electron spin rotation as well as the energy resolution necessary to resolve low-energy electrons near the Compton edge (required for monitoring beam polarization). Further, dipoles drive many higher-order aberrations that act to blow the beam up that require higher multipoles to correct. These were not needed to get the beams into the dump with the 12-sigma constraint through the line.

Characteristic Angles and Some Related Disruption Effects

There are several processes that influence the beam and various characteristic angles that deserve discussion. The typical emission angle for most high-energy radiative processes such as incoherent bremsstrahlung is $\theta_r \approx 1/\gamma \approx 1\,\mu$r at 500 GeV. As the photon energy decreases, the angles begin to grow compared to θ_r as does the interaction volume of the electron. When the electron interacts with the collective field of the other bunch, it sees transverse electric and magnetic fields that are nearly equal. It is easy to show that these fields can bend an electron by angles significantly greater than θ_r so this angle is physically interesting. The resulting radiation, similar to synchrotron radiation, is called beamstrahlung.

A characteristic angle for the "full energy" primary, disrupted electrons is [Hollebeek 1981]:

$$\theta_d \equiv \frac{2Nr_e}{\gamma(\sigma_x+\sigma_y)} = \frac{D_{x,y}\sigma_{x,y}}{\sigma_z} \ , \tag{11.139}$$

where $\sigma_{x,y}/\sigma_z$ is sometimes called the diagonal angle and equals θ_d for $D_{x,y}{=}1$. Typically, in the e^+e^- channel, the maximum disruption angle is $\theta_{x,max}{=}\theta_d$ because the disruption parameter in the vertical is so

large that one gets overfocusing or a thick lens effect whereas the focusing over the length of the beam in x is weaker but cumulative or more like a thin lens. The rms angles are $<\sigma_{x,y}>= 0.550\theta_d$. For , the situation reverses and the vertical disruption angle dominates as expected.

Beyond $\theta^*=\sigma_{x',y'}\ll\theta_d$, another angle of relevance here is the spin precession angle θ_s. This can be expressed in terms of the spin tune

$$\nu_s = \frac{E[GeV]}{0.44065}\cdot\frac{\Theta}{2\pi} \quad , \tag{11.140}$$

where Θ is some deflection angle in radians. For the bends used here, this is typically twice the maximum disruption angle θ_d=256 μr according to Table 11-40. Further, the bends rotate the polarization in a correlated way in the bend plane between the input and output rather than depolarize the beam assuming that radiative effects (spin flip) are small and that the emittance is small and that there are no significant orbit distortions. The effective polarization after such a bend is $P_{\rm eff} = P_{\rm inc}\cos(2\pi\nu_s)$=0.42–0.84.

There are still several depolarizing effects to be considered. The rms disruption angles cause a net depolarization. A previous section gives another due to an rms gaussian energy spread δ. Clearly, an energy spread will give rise to a spread in precession frequencies that can cause some depolarization but this appears small for the disruption angles and bend angles used here which is about 0.1 % for θ_d=256 μr and <2 % for a 1 mr of bend even at 500 GeV. This classical effect can be avoided insofar as the beam polarization measurement is concerned because the beam is dispersed with sufficient resolution to emphasize the undisrupted portion of the beam. A worse effect, pointed out by Yokoya and Chen, relates to spin-flip during the beamstrahlung process. This can be related to n_γ in Table 11-40. and was one constraint on our bend angles of 1 mr or so but this is not an intrinsic limitation and could prove useful to monitor beam overlap.

The final angle of relevance is the crossing angle θ_c at the IR. For multibunch trains, we may need to introduce a crossing angle and design the FF quadrupoles accordingly. This can decrease the collision efficiency ζ so that we have to introduce variable, crab-crossing cavities [Palmer 1990] that rotate the beams to the appropriate orientation at the IP or CP (the $e\gamma$ conversion point) to restore ζ. We have shown that such cavities will be required for all incident channels [Spencer 1995c] so we will assume that they exist.

Optical Considerations on Energy Recovery

For energy recovery, we want the purely transverse dispersive as well as the **mixed** transverse-longitudinal terms in the optics to be zero to all orders in $\delta\equiv(E-E_o)/E_o$. However, for energy recovery, we want to add another constraint that would not be important otherwise—the synchronicity or time (phase) variation with energy that makes different energy electrons arrive with the correct but differing phases required to lose (or gain) the correct energy. Until otherwise noted, this implies that a high-order achromat is desired (that is not otherwise required), with the corollary condition that we want a first-order or linear R_{56} term that we would like to make variable while otherwise maintaining the high-order achromatic behavior.

In practice, it is the energy-angle correlations as well as the purely angular and chromatic terms that are of most importance because of the small spots and the strong disruption at the IP. The most relevant terms without bending magnets, in standard notation, are:

Second Order:	(1|26), (3|46)	(1|16), (3|36)
Third Order:	(1|266), (3|466)	(1|166), (3|366)

where the second set of terms in each row is less important. The (1|26) and (3|46) terms are dominant but would not require correction if we only wanted to take the beam to the dump. In fact, they would help the

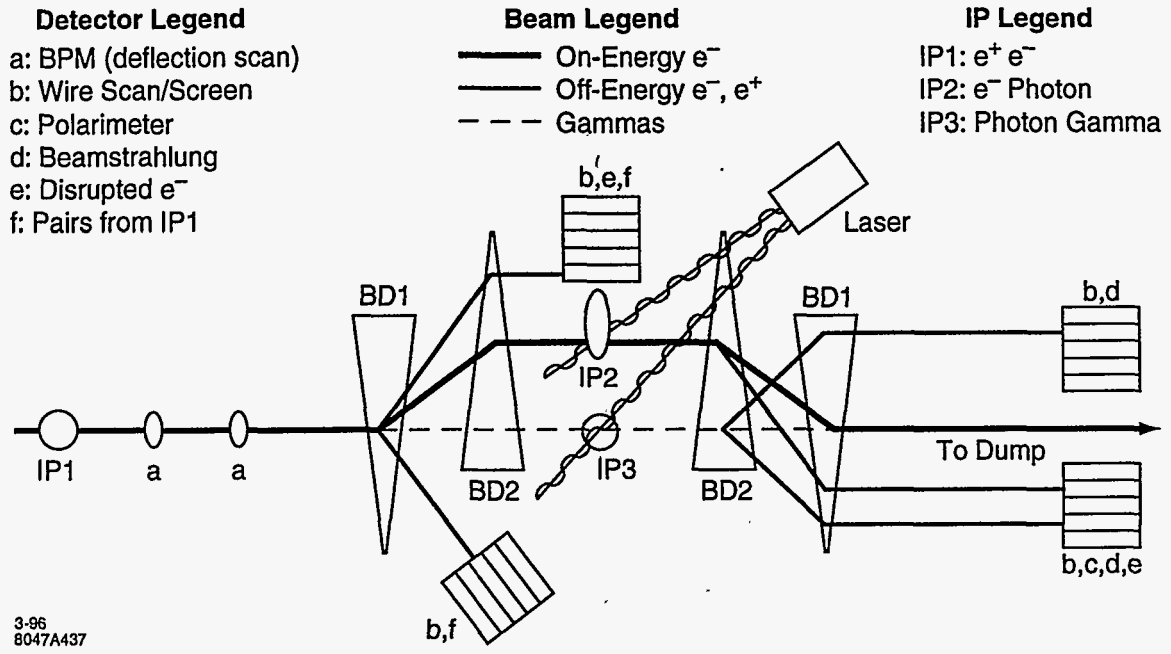


Figure 11-66. *NLC extraction line and diagnostic layout.*

loading on the dump window. The bunch length terms (5|22) and (5|44) enter at the level of 0.1 μm or 0.1 % of σ_z. If we add bending magnets, the only additional terms of importance are (1|66) and (1|666).

Although not essential for most of the other applications, reducing such terms improves things—especially the spectrometer resolving power that is important for the Compton measurements as well as measuring the energy distribution of the outgoing disrupted beam which in turn measures the overlap and distribution of the incoming beams. We should note that it appears more practical to use the 10 MW of power than to recover it to the grid.

Configurations for Beam Monitoring and Experiments

We have already discussed the point-to-parallel configuration for the primary beam. Of course, the undisrupted portion of the beam goes to the dump anyway but this has several different single-beam applications with differing focal points along the way to the dump. Some of these possibilities are discussed below and shown schematically in Figure 11-66.

11.8.5 Beam Control Hardware

The initial quadrupoles are assumed to be permanent magnets (PMs) made from SmCo and NdFe. We have actually made and used such multipoles at SLAC since 1982. Prototypes for the SLC final-focus quadrupoles were made and tested under a variety of conditions such as very strong transverse fields [delCorral 1982]. Figure 11-67 shows several PM systems including the SLC prototype, an SLC split-ring sextupole and an

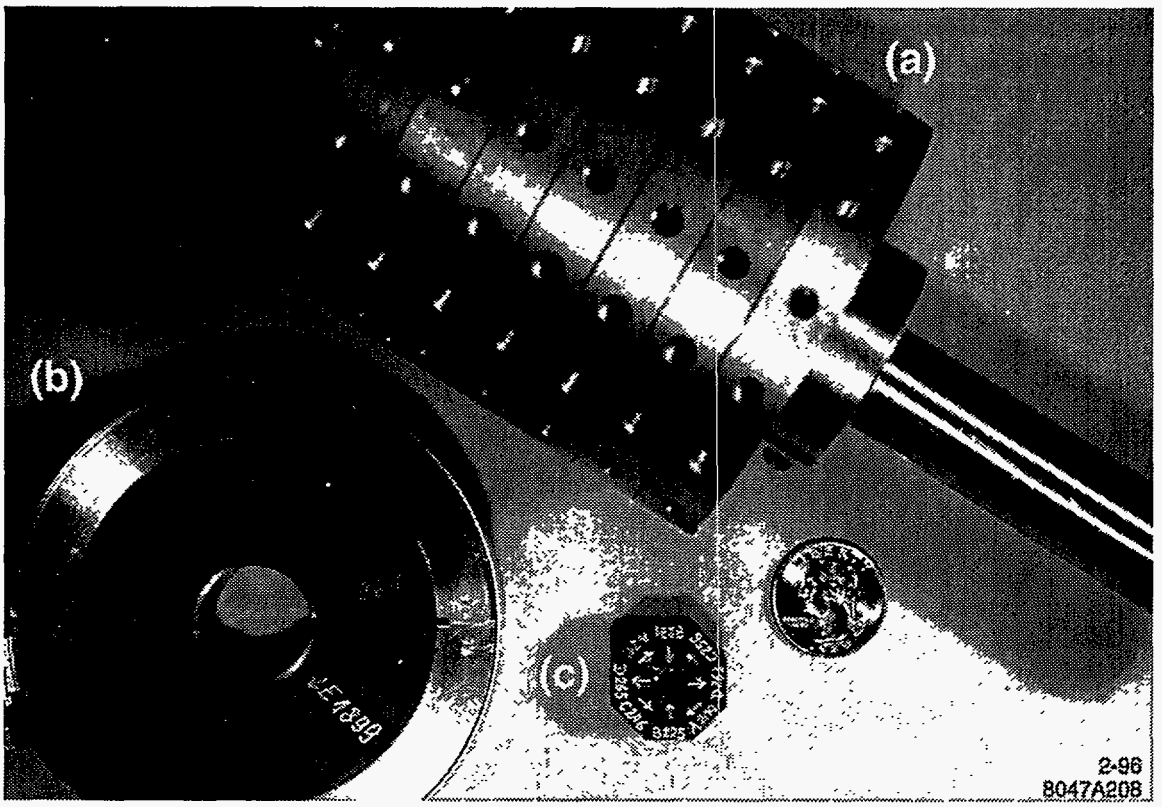


Figure 11-67. *Representative permanent magnet multipoles: (a) a 7-layer quadrupole for the SLC, (b) an SLC damping ring sextupole, and (c) a 500 T/m NLC quadrupole prototype.*

NLC prototype quadrupole with a gradient G>500 T/m that was made in 1989. More recently our experience with the permanent magnet sextupoles used in the SLC damping rings since 1985 was reported [Gross 1994].

An important characteristic demonstrated by these sextupoles, applicable to any PM multipole that has symmetry about the median plane, is that there does not have to be any material in the median plane of the magnet. Thus, the large number of low-energy pairs that spread out from the solenoid axis to the position of the primary outgoing beam can be passed without necessarily intercepting any material. While it appears that conventional PMs can be used in the IR, some sample calculations using a new technique by which magnets with open structures can be designed and still achieve very strong fields also appear to be possible. Additionally, some of the dipoles are combined function magnets (dipoles/quadrupoles) to simplify the transmission of the larger energy spreads and emittances in the outgoing beam from the IP.

Finally, to avoid the nonlinearities in the end-field of the detector solenoid, we propose using a special wiggler in that region that is driven by the strong solenoid field itself. This element serves as a different kind of mirror plate for the detector field that shields the beam from the nonlinear transverse and longitudinal fields that occur in the end region by rerouting and reshaping the flux there. Ultimately that flux must be returned to the detector's conventional endcap by topologically distorting the wiggler's exterior mirror plate. Currently, a superconducting quadrupole (300 T/m) occupies this region and accomplishes a similar result.

11.8.6 Beam Monitors and Diagnostics

One of our primary goals is to provide diagnostics to help put the beams into collision, to maintain this state, and to measure and optimize luminosity. The latter involves diagnostics to characterize the beams: polarization, energy, intensity, profile, positions (both absolute and relative), and the disruption of such quantities from the beam-beam and environmental factors.

We also need to dump the beams cleanly and safely. This involves the possibility of other functions such as secondary beams: e, μ, n and γ as well as other forms of energy recovery before final disposal.

The guiding principle in the instrument layout shown in Figure 11-66 was to minimize the material in the high-power beams. Thus, there is a significant use of lasers to control the production of additional particles. Rf BPMs for both position, angle, and timing information are assumed to begin near 5 m where the outgoing beams are still small and C-band cavities [Hartman priv] can have apertures comparable to the quadrupoles ($\approx \lambda/4$). Nevertheless, since beamstrahlung is unavoidable, there are possibilities [Norem 1995] to use either Compton or beamstrahlung photons that could prove useful for monitoring the position, size, and correlations of the bunches at the IP on a bunch-to-bunch basis. Clearly, beamstrahlung is quite sensitive to any changes in these parameters at the IP. In fact, the photon distribution is a better measure of the bunch profile at the IP than the outgoing, disrupted electrons as shown by Figures 11-55 to 11-60. We also assume BPMs, wire scanners, and screens similar to those used for the SLC [Field NIM].

The usual monitors such as BPMs and wire scanners are deployed at key locations for stabilization and feedback for position and energy control. It is understood that the BPMs include both amplitude and phase sensing. There are also more arcane monitors that look at the beamstrahlung from the IP as well as Compton backscattering that have nanometer resolutions [Norem 1995] and can function on a bunch-to-bunch basis during colliding beam operation or with single beams.

Luminosity Measurements

To estimate the difficulty of this measurement, one can begin by looking at what the Bhabha rate might be for an "annular" detector subtending some angular range $\delta\theta$ near the front face of the first quadrupole in the dump line after the IP. Because this quadrupole (QDF1) is located at l^*=2 m with a radius of 7.5 mm, a reasonable minimum angle θ_1 occurs in the range $\theta_1 \approx$1–3 mr. If the predicted rate is reasonable for a bunch or even a train, we can then consider the backgrounds predicted in the beam-beam simulations to get an estimate of the signal-to-noise ratio. One can then consider the necessary characteristics required for the detector based on the characteristics of the backgrounds.

For unpolarized Bhabha scattering, a good approximation for the angles of interest here (a few hundred mr or less) is

$$\sigma(\delta\theta) = \int_0^{2\pi}\int_{\theta_1}^{\theta_2} \frac{d\sigma}{d\Omega}(e^+e^- \to e^+e^-) \approx \frac{4\pi\alpha^2}{s}\left[\frac{1}{\sin^2(\frac{\theta}{2})} + 4\ln\left(\sin\frac{\theta}{2}\right)\right]_{\theta_2}^{\theta_1} . \qquad (11.141)$$

Using $4\pi\alpha^2 = 261\,\mathrm{GeV}^2\mathrm{nb}$ with 250-GeV beams at very forward angles we have

$$\sigma(\delta\theta)|_{\theta_1}^{\theta_2} \approx 1.05\left(\frac{2}{\theta_1\theta_2}\right)^2(\theta_2-\theta_1)(\theta_2+\theta_1)\ \mathrm{p}b = 4.2\frac{1}{\theta_1(mr)}\frac{1}{\theta_2(mr)}\left[\frac{(\theta_2-\theta_1)(\theta_2+\theta_1)}{\theta_1\theta_2}\right]\mu\mathrm{b} \quad . \quad (11.142)$$

Assuming $\mathcal{L} = 6 \cdot 10^{33}$ at 250 GeV for the 500-A column in Table 11-40 implies a Bhabha rate

$$R_B^{500A} = \mathcal{L}\sigma(\delta\theta)\Big|_{\theta_1=1mr}^{\theta_2=3mr} \approx 22,400/\text{s} = 125/\text{train} = 1.4/bunch \tag{11.143}$$

for a rep-rate of 180 pps in the linac with 90 bunches/train. Increasing θ_2 to 100 mr increases this rate by less than 15%. Again, using the column for 1000 A in the Table for 500-GeV beams, the corresponding numbers are

$$R_B^{1000A} = \mathcal{L}\sigma(\delta\theta)\Big|_{\theta_1=1mr}^{\theta_2=3mr} \approx 12,100/\text{s} = 101/\text{train} = 1.1/\text{bunch} \quad . \tag{11.144}$$

Increasing θ_2 to 100 mr increases the rate to 114/train or 1.3/bunch. θ_2 was made small to be compatible with an rf shield that would extend from the last quadrupole toward the IP with a radius comparable to that of the quadrupole. While such a shield is not as important for the outgoing beam lines, there are no serious penalties imposed on the luminosity monitor for having one.

The value of θ_1=1 mr is reasonable from Figures 11-55, 11-58, and 11-59 for the typical beams of outgoing photons and electrons that we expect. While these two beams have comparably high intensities and are not expected to limit the measurement, the pairs are clearly a problem even though we expect on the order of 10^4/bunch crossing or 10^6 less than for the incident beams (or photons) per crossing. This is demonstrated in Figure 11-62 for the 500-A configuration. However, because these particles are predominantly below 1 GeV, some form of energy discrimination should be sufficient to deal with them.

However, because the pairs that are relevant here are in the higher-energy part of the pair spectrum, two important points can now be made. First, a sampling measurement of these particular pairs is itself a measure of the luminosity and is therefore a useful adjunct with a distinctly different signature. Further, this pair rate, in this location, is a problem that has to be dealt with anyway. Second, we can avoid this problem, at least for the Bhabha luminosity measurement, by going further downstream. This is clearly demonstrated in Figure 11-63 that shows the different beams at 5-m from the IP. However, from Figures 11-59 and 11-60 showing the x and y profiles for both the photons and electrons, this is not the preferred location. Clearly, the optimal location is somewhere in between but a location near 2 m appears acceptable for making both measurements—especially because this allows the possibility of varying the solid angle subtended by the detector via tromboning or moving it longitudinally toward the IP. We expect almost 100 pairs within the quad aperture there.

Because halving the angle θ_1 quadruples the rate, the main questions are then how small we can make this angle, and whether it is practical to move the detector longitudinally to vary the subtended solid angle for different operating conditions such as beam scans. There are a number of interesting questions concerning the detector and its energy and timing discrimination for individual bunches.

Polarization Measurements

We propose to measure the longitudinal electron beam polarization (P_e) with a Compton polarimeter as shown in Figure 11-56 similar to how it is done for the SLC [Woods 1994]. The polarimeter will detect Compton-scattered electrons from the collision of the longitudinally polarized electron beam with a circularly polarized photon beam. The detector will be a segmented threshold Cherenkov counter similar to that in use at the SLC.

The counting rates in each Cherenkov channel are measured for parallel and anti-parallel combinations of the photon and electron beam helicities. The asymmetry formed from these rates is given by

$$A(E) = \frac{R(\rightarrow\rightarrow) - R(\rightarrow\leftarrow)}{R(\rightarrow\rightarrow) + R(\rightarrow\leftarrow)} = P_e P_\gamma A_C(E) \tag{11.145}$$

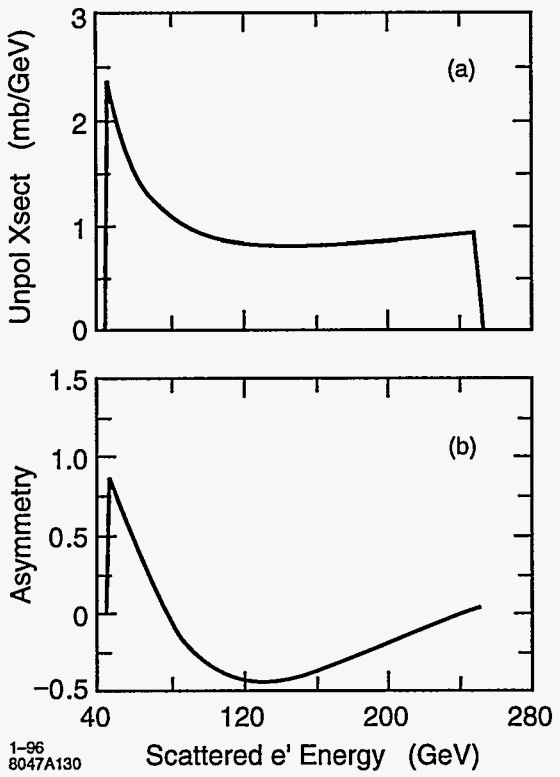


Figure 11-68. *(a) Linear Compton scattering and (b) asymmetry A_c for 250-GeV electrons.*

where P_γ is the circular polarization of the laser beam at the CIP at IP2 in Figure 11-66 and $A_C(E)$ is the Compton asymmetry function. $A_C(E)$ and the unpolarized Compton cross section are shown in Figure 11-68 for an electron beam energy of 250 GeV and a photon energy of 1.17 eV. The linear Compton spectrum is characterized by a kinematic edge at 46 GeV (180°) backscatter in the center of mass frame, and the zero-asymmetry point at 77 GeV (90° scattering in the center of mass frame).

The photon energy of 1.165 eV corresponds to using an Nd:YAG laser operating at 1064 nm. The long wavelength is chosen to avoid background from the reaction $e^-\gamma \rightarrow e^+e^-e^-$ that opens for

$$E_\gamma(\text{eV}) > \frac{522}{E_e(\text{GeV})} \quad . \tag{11.146}$$

However, this channel is mediated by a second photon and lower thresholds are possible when the second photon is real as from a secondary process such as Compton backscattering. Another laser photon can then produce a Breit-Wheeler pair. This cross channel process opens, in lowest order ($\eta \rightarrow 0$), for:

$$E_\gamma(\text{eV}) > \frac{315}{E_e(\text{GeV})} \tag{11.147}$$

and clearly depends on laser intensity and wavelength. Similarly, for beamstrahlung photons, the threshold is related to the Schwinger critical field by the laser's intensity and wavelength for a given incident electron energy.

The kinematics for Compton scattering gives the location of the Compton edge, $E_C(\text{edge})$, and the edge asymmetry, $A_C(\text{edge})$, as

$$\begin{aligned} E_C(\text{edge}) &= E_e \cdot y \\ A_C(\text{edge}) &= \frac{y^2-1}{y^2+1} \end{aligned} \tag{11.148}$$

$$y = \frac{1}{1+\frac{4E_eE_\gamma}{m_e^2}} \tag{11.149}$$

where this is again in the limit that the classical, strong field parameter $\eta \ll 1$. The operant dimensionless, classical, strong field parameter is:

$$\eta = \frac{e\mathcal{E}_{\text{rms}}\lambda\!\!\!^{-}_\gamma}{mc^2} = \frac{e\sqrt{\langle A_\mu A^\mu \rangle}}{mc^2} \quad . \tag{11.150}$$

This is described elsewhere [Bula 1995] where one sees the smearing effect on the edge from the nonlinear and multiple scattering effects due to high laser intensity. We have calculated this for an NLC example at 250 GeV using YACC, a variant of CAIN. A sample result is shown in Figure 11-69 and discussed further in the section on secondary beams.

Disruption of the electron beam during the collision process can have a significant effect on both the effective polarization of the electron beam during the collision process [Yokoya 1988] and on the measurement of beam polarization. The polarization of the undisrupted incoming beam can easily be determined by measuring the electron polarization when the positrons are absent. The effective polarization of the disrupted electron beam can be determined from Compton asymmetry measurements during collisions with studies of this asymmetry for different targeting of the the laser beam on the electrons at IP2. Corrections for the effects of jitter in the offsets between the beams at the IP can be made using multiple BPM measurements near the IP which is one objective of the spectrometer. Good knowledge of the disrupted electron beam energy spectrum will be required and this can be measured with a wire scanner near IP2.

Because these electrons are quite low in energy and fall on the far out tail of the disrupted beam distribution, the only requirement is the ability to resolve their energy to within a few GeV, *i.e.,* we would like a resolving power at 500 GeV of:

$$R = \frac{P}{\Delta P} \geq 500.$$

We have provided $R > 2000$ for the full energy beam so there is no problem. We have pointed out before that the nonlinear effects are quite interesting in this regime but this will be discussed elsewhere.

Beam-Beam Deflection Scans

At the SLC, measurement of the deflection of one beam by the opposing beam at the IP, when they are offset, provides an important diagnostic which allows measuring, optimizing, and maintaining luminosity [Bambade 1989]. At the NLC, deflection angle measurements could prove just as valuable. For undisrupted beams, the luminosity varies exponentially with the square of the separation or offset between beams:

$$\frac{\mathcal{L}(\Delta_{x,y})}{\mathcal{L}(0)} = e^{-\Delta_{x,y}^2/2\Sigma_{x,y}^2} \tag{11.151}$$

where $\Delta_{x,y}$ is the transverse offset of the two beams and Σ_y (Σ_x) is the quadrature sum of the rms vertical (horizontal) beam size for the two colliding beams. A good test of our ABEL code is to compare to this

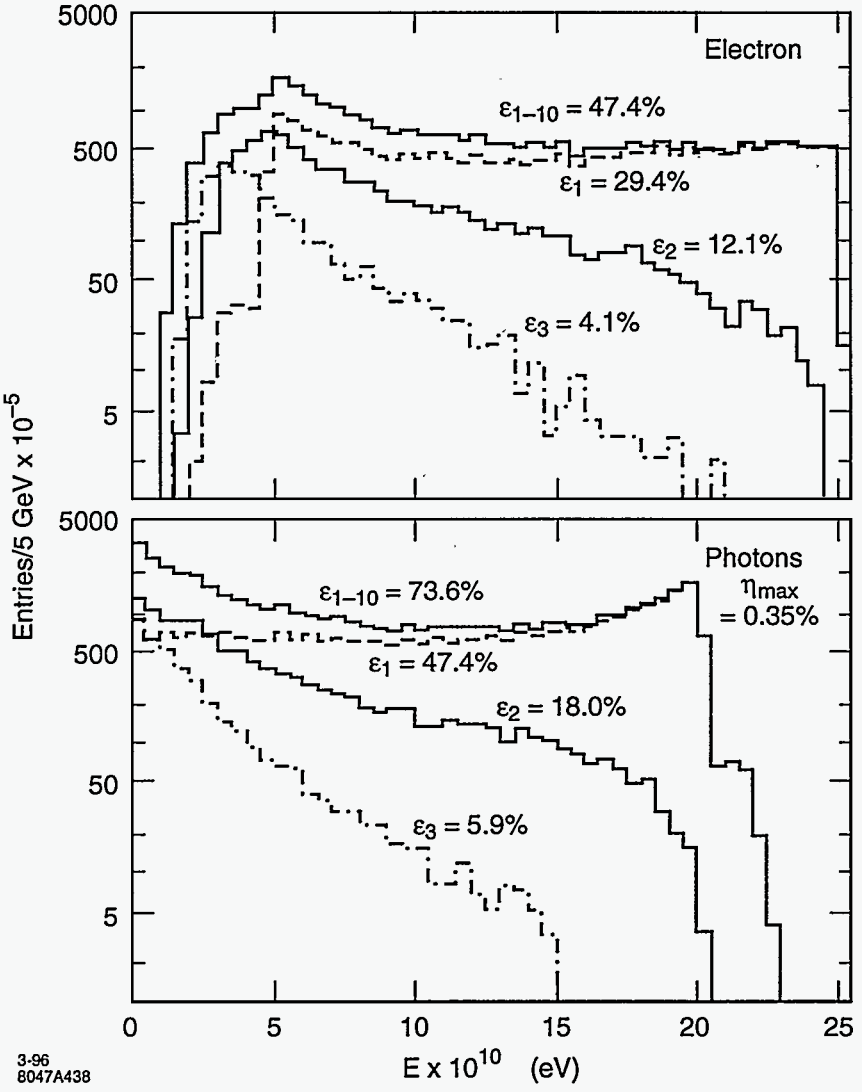


Figure 11-69. *Compton spectra for undisrupted beams at IP2 for high laser intensities.*

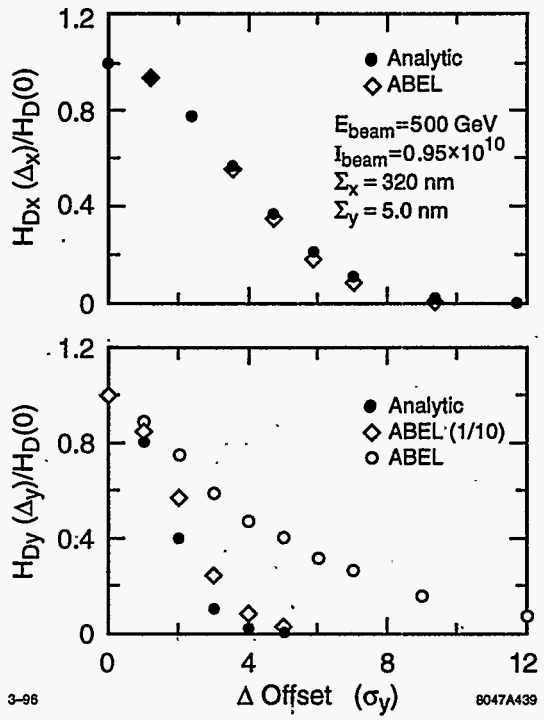


Figure 11-70. *Comparison of beam-beam enhancement factors H_D (or efficiencies ζ).*

prediction for a relevant set of NLC design parameters (Case A) while calculating a deflection scan. The results are shown in Figure 11-70 and are labeled by H_D but are more appropriately termed a collision efficiency factor ζ [Spencer 1995c] especially because the analytic calculations do not include disruption effects.

The deflection angles can be determined from BPM measurements before and after the IP. For flat, upright Gaussian beams which are centered on the x axis, the mean y deflection angle is given by:

$$< \theta_y >= \frac{2r_e N}{\gamma} \Delta_y \int_0^\infty dt \frac{e^{-\Delta_y^2/(t+2\Sigma_y^2)}}{(t+2\Sigma_y^2)^{3/2}(t+2\Sigma_x^2)^{1/2}} \tag{11.152}$$

where r_e is the classical electron radius, N is the number of target particles, and γ is the beam energy divided by the electron mass. Sample deflection scans for the NLC design parameters of Case A corresponding to Figure 11-70 are given in Figure 11-71.

According to Ref. [Raimondi 1995], the effect of beam-beam disruption will cause larger deflection angles than given in the above formula for y deflection angles, but will have a negligible effect on horizontal deflection angles. Our results indicate observable differences in both transverse directions, *e.g.*, 6–7% in the peak deflection angle in x and 9–10% in y. In both cases, the location of the peak deflection is pushed outwards but one sees that the predominant effect of the disruption is to push the location of the y peak outwards beyond 10σ or so for the chosen parameters. In fact, we can summarize the results of a number of calculations as follows: 1) small disruptions ($D_{x,y} \ll 1$) can use the analytic results while large Ds need simulations, 2) for a given set of parameters typical of Case A, the Y deflection angle converges to the same

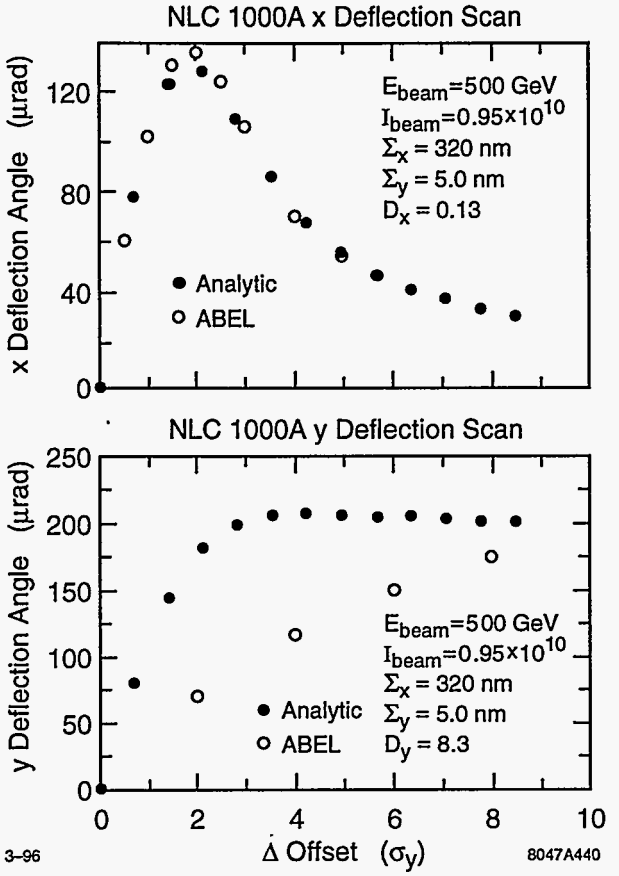


Figure 11-71. *Comparison of beam deflection scans corresponding to Figure 11-70*

value ($\approx 220\,\mu$r) independent of σ_y, 3) the location at which this value is achieved comes at about $10\sigma_y$ in Figure 11-71, 4) this is true for a large range in σ_y, but 5) this value depends on the the value of σ_x assumed.

We can understand these results as due to the significant disruption/focusing of the beam in the collision region. A 200-μr angle over a 200-μm bunch length can give an offset of 40 nm or $\geq 10\ \sigma$. It seems advisable to test the characterizations made above with real experiments on the SLC where practical.

Because the maximum deflection angle will be comparable to or less than θ_d in Table 11-40, we may conclude that a radially uniform aperture in the outgoing quadrupoles should be acceptable for doing this diagnostic test even though the outgoing beamstrahlung now has a pointing angle that can sweep out angular apertures of nearly half a mr. This is not a problem because the apertures were sized to allow such angles. Rather, it provides another reason to place the primary beam dump closer to 100 m rather than the 150–200-m distance that would otherwise be allowed. Pair production is a worse problem for the quadrupole acceptances.

Another signature for disruption could come from measuring the photon moments as well as those of the electrons according to Refs. [Norem 1995, Field NIM]. The drawbacks of the deflection scan are obvious so that other alternatives, involving less beam time overhead and aperture, are being considered. This is one reason for the good energy resolution at IP1 in Figure 11-66. However, regardless of the method of choice, the deflection scan can be expected to provide a well-understood comparison that can be used for the NLC.

11.8.7 Beam Dump

The primary beam dumps have to dispose of essentially all of the power at the highest beam energy (750 GeV). Appendix 11.A argues that water should be the primary absorber in a cylindrical vessel housing a vortex-like flow of water with vortex velocity ≈1–1.5 m/s normal to the beam momentum. The vessel is 1.5-m, diameter and has a 6.5-m-long water section, followed by ≈1 m of water-cooled solids to attenuate a 750-GeV EM cascade shower. The beam enters through a thin window ≈1-mm thick and 20-cm diameter. Production of ≈5 l H_2/16 MW beam power from radiolysis [Walz, 1967] can be mitigated with a catalytic H_2/O_2 recombiner that has a closed-loop system that contains all radioisotopes.

The reader is referred to Appendix 11.A for the complete details. Here we only summarize the main characteristics and point out that a window size of 20 cm appears practical from all standpoints. We also note that there is the possibility to do some classical beam dump experiments after the dump which may be facilitated by dispersing the beams horizontally.

11.8.8 Secondary Beams

Beyond the usual possibilities, there are interesting opportunities to produce intense beams of higher mass leptons and hadrons as well as neutrons that haven't really been explored yet but could provide both a unique and comparatively inexpensive alternative source to nuclear reactors for some applications.

Muon Possibilities

Discussions on muon colliders have been resurrected recently so that the associated possibility of muon beams and colliders based on very high-energy, high-quality electron beams naturally arises. In fact, this is almost unavoidable when one considers the backgrounds from muons that are also unavoidable and other questions of energy recycling and beam disposal.

Most discussions about muons use proton beams to produce the muons but Barletta and Sessler have looked at electroproduction recently. They concluded it was not viable but also did not consider the kinds of beams we will assume here. An important aspect is that we can use all the cooling tricks that have been discussed to cool the very large emittances from proton production although our assumption is that we can make them in a way that avoids many of the steps (and problems) normally required because we start at higher energies where there are several intrinsic advantages. The procedure is to produce high-energy photons as in the preceding section which are then used to produce the muon beams. Photoproduction was proposed previously for producing high-energy polarized positrons and there are a number of variants including the use of crystals and high-power lasers or FELs.

Photon Possibilities

This possibility relates to the second interaction region (IP2) and to the nonlinear QED experiment on the FFTB line at SLAC (E144). Clearly, similar problems are encountered with the Compton polarimeter and profile monitors based on photons. Related questions concern beam stability and synchronization so that much of the hardware has a number of interrelated uses. In Figure 11-69, we gave a worst-case scenario for IP2 assuming we were limited in both laser wavelength and power to the specified quantities. We point out

c.m. Energy (TeV)	0.5	1.0	1.5	TeV
Bunch Current, N_{max}	0.70	1.45	1.45	$\times 10^{10}$
Bunches/Train	100	75	75	
Bunch Separation	42	42	42	cm
Trains/s	180	120	120	s^{-1}
Coupling, $K \equiv \epsilon_y/\epsilon_x$	1.0	1.0	1.0	%
IR Beta, β_x^*/β_y^*	10/0.1	10/0.1	30/0.3	mm
IR Size, σ_x^*/σ_y^*	316/3.16	224/2.24	316/3.16	nm
Sizes at δz=4 cm	1.31/1.27	0.93/0.89	0.53/0.42	μm
Sizes at δz=8 cm	2.56/2.53	1.81/1.79	0.90/0.84	μm
Sizes at δz=12 cm	3.82/3.80	2.70/2.68	1.30/1.26	μm
Sizes at δz=20 cm	6.53/6.33	4.48/4.47	2.13/2.10	μm
Divergences, $\sigma_{x'}^*/\sigma_{y'}^*$	31.6/31.6	22.4/22.4	10.5/10.5	μrad
Laser $\lambda(\mu)$	1.05	2.1	3.2	μm
$< \delta t >_L$	0.1	0.1	0.1	ps
$< \delta l >_L$	0.3	0.3	0.3	ps
σ_C[b]	0.197	0.197	0.197	$\times 10^{-24} cm^2$
Photons/Pulse, N_λ	4.2	4.7	2.9	10^{18}

Table 11-41. *Parameters for 0.5, 1 and 1.5 TeV. Electron emittances are $\gamma\epsilon_x$=5 × 10^{-6} m and $\gamma\epsilon_y$=5 × 10^{-8} m. The energy spread is σ_E/E=0.20 % and the bunch length is σ_z=100μ.*

that ignoring nonlinearities in the laser-electron interaction so that we can use unlimited laser power if we could get it, we could conceive of higher luminosity in this channel than currently appears slated for IP1. Table 11-41 gives some characteristic numbers. Additional results can be found in the section on the second IP and elsewhere [Spencer 1995c].

11.8.9 Energy Recovery and Its Applications

Following the introductory discussion, we note that it costs a significant amount to build and continuously dispose of such high-power beams in an environmentally acceptable way. Depending on the options that are pursued, there are at least four beam dumps required—not counting possible secondary lines. If we suppose that it takes another 4% of the total beam power just to dispose of the beams, this is another $1,000/day in power costs per dump. Clearly, this is worth pursuing if, in fact, one can find some nondisruptive means that are themselves environmentally acceptable. This question has been raised in the past at SLAC but the power involved and its subsequent uses implied it was not economically feasible.

Some ways to look at rf recovery methods are to remember how klystrons produce rf power from an electron beam, how recirculating systems work, and how high-frequency colliders such as CLIC propose to produce its rf power. If we want to store it or use it in some associated storage rings then we should consider superconducting systems. Any subsequent physics application should justify the cost. Ideally, whatever

the sequence, one would end up with a final, dumped beam that has been reduced to an energy below the neutron production threshold to minimize activation of long-lived isotopes.

One obvious synergistic application is to try to drive an FEL with this beam in such a way that the same FEL can be used to produce variable, high-energy photon beams. It is interesting to point out that the Stanford recirculating SCA proposed and tested energy recovery to improve FEL efficiency [Rohatgi 1987]. One can expect such methods to provide up to 25 MV/m.

11.8.10 Other Questions and Problems

There is a question concerning the sufficient separation between the various beams. This is especially relevant for the primary e and γ beams that are to be monitored and that have comparable densities and power. Another question is the e^- depolarization that we have not thought much about except for measuring it. Finally, there are all the alternatives for monitoring the luminosity, beam sizes, and offsets at the IP under normal operating conditions. We have proposed several methods in each case, and that work is still in progress. We have shown that there are significant differences between the SLC and NLC and that the NLC has more possibilities than the SLC because of the greater disruption and its secondary effects. We have also verified that virtually all of the corresponding diagnostic tools that are used for the SLC such as conventional deflection scans can be accommodated by the extraction optics and diagnostic devices even in the presence of pairs. However, this subject is a fertile area for exploration and it is being pursued together with that of strong PMs having open structures that are matched to the outgoing beam distributions and characteristics under differing diagnostic conditions. Finally, we strongly emphasize the importance of testing as many of the various calculations that have been presented here as practically possible. This could benefit the SLC as well as any future generation machine.

11.9 Conclusions and Comments

The NLC final-focus design fulfills all specifications, and achieves the desired IP spot sizes of about 300 nm×4 nm, in the entire c.m. energy range between 350 GeV and 1.5 TeV, while maintaining a reasonable momentum bandwidth of ±0.60% or larger. The proposed design lends itself to an easy and straightforward upgrade to 1.5-TeV-c.m. energy, requiring only minor changes in geometry. The price to be paid for this flexibility is a final-focus length of about 2 km.

All possible sources of spot-size dilution, which are relevant on different timescales, have been budgeted, and are controlled by stabilization, tuning and maintenance systems. These systems also ease the tolerances on element vibrations, position drifts, and field stability. Most of the tolerances are not particularly tight, and have already been achieved at the FFTB or at other places. It is interesting that the luminosity loss caused by ground motion appears to be insignificant and that, therefore, the ground (bedrock) provides an ideal reference for magnet alignment and stabilization. In order not to destroy the coherence of the ground, care will be needed in the mechanical design of components and magnet supports. The primary engineering task will be to contain the effect of cultural noise.

The design of the NLC final focus profited greatly from experience with SLC and FFTB operation, whose influence is clearly exhibited by optical layout, aberration and tolerance budgets, tuning schemes, sensitivity studies, diagnostics, and operational procedures.

In conclusion, the NLC beam delivery system will not only produce small spot sizes at the IP, but it also promises redundant tunability, adjustability over a wide energy range, and, last not least, tolerable detector background.

11.A The Beam Dumps

If the efforts to economically recover a significant fraction of the spent electron beam energy are not successful, we will be faced with the tasks of safely disposing and dissipating these 10-MW beams. The working beam parameters are either:

90 Bunches/Train($\equiv n_B$) at $N_B \leq 0.85 \cdot 10^{10} \equiv 76.5 \cdot 10^{10}$/Train and f_{rep}=180 Hz

or

90 Bunches/Train($\equiv n_B$) at $N_B \leq 1.25 \cdot 10^{10} \equiv 112.5 \cdot 10^{10}$/Train and f_{rep}=120 Hz.

Then the beam energies and average power per beam are:

E_o (GeV)	40	80	150	250	500	750
$P_{av,90/180}$ (MW)	0.88	1.76	3.30	5.51	11.02	16.52
$P_{av,90/120}$ (MW)	0.86	1.73	3.24	5.40	10.80	16.20

To assess the magnitude of the task, let us examine some engineering materials commonly used at accelerators. Specifically, we will look at the energy deposited in such materials at the beginning of an electromagnetic cascade and also in the region of peak energy deposition at shower maximum. At the beginning of the cascade (before any shower multiplicity), the average energy loss is approximately given by the ionization loss and can be written as

$$P'_{av} \sim 1.6 \cdot 10^{-19} (-\rho \frac{dE}{dx}) f_{rep} N_B n_B \tag{11.153}$$

and at shower maximum

$$P'_{av,max} \sim \Pi_{max} P'_{av} \tag{11.154}$$

where P'_{av} and $P'_{av,max}$ are the average power depositions per unit length at the beginning and at shower maximum. ρ is the specific gravity of the material, dE/dx is the minimum ionization loss, and the other quantities are defined in Table 11-40.

For electrons, the location of the shower maximum $T^{e^-}_{max}$, in units of radiation lengths X_o, and the maximum shower multiplicity Π_{max} are given by Rossi [Rossi 1952] as

$$T^{e^-}_{max} = 1.01[\ln \frac{E}{\epsilon_o} - 1] \tag{11.155}$$

and

$$\Pi^{e^-}_{max} = 0.31(\frac{E}{\epsilon_o})[\ln \frac{E}{\epsilon_o} - 0.37]^{-\frac{1}{2}} \tag{11.156}$$

where the critical energy ϵ_o of an absorber material is the unit of measure that makes these expressions independent of material. Typical values for carbon and lead are 76 and 7.6 MeV when one includes the effect of density variation, otherwise ϵ_o goes inversely with atomic number. The following tables give various quantities derived from the equations given above to judge the suitability of various materials for this application.

Table 11-42 gives the location of T_{max} as a function of beam energy. Table 11-43 gives the shower multiplicity Π_{max} at T_{max} as a function of beam energy. Table 11-44 gives the average power deposition per unit length

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	4.9	5.3	5.4	6.0	6.5	6.6	6.7	7.6	7.7
80	5.6	6.0	6.1	6.7	7.2	7.3	7.4	8.3	8.4
150	6.3	6.6	6.7	7.3	7.8	8.0	8.1	8.9	9.0
250	6.8	7.1	7.2	7.8	8.3	8.5	8.6	9.4	9.5
500	7.5	7.8	7.9	8.5	9.0	9.2	9.3	10.1	10.2
750	7.9	8.2	8.3	8.9	9.4	9.6	9.7	10.5	10.6

Table 11-42. *Location of $T^{e^-}_{max}$ for various materials (in units of X_o).*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	48	65	74	121	195	224	244	537	584
80	90	123	132	231	371	428	467	1030	1122
150	162	220	237	415	670	773	843	1867	2034
250	260	350	382	670	1084	1252	1365	3031	3302
500	497	678	732	1287	2087	2413	2631	5793	6390
750	727	993	1072	1889	3067	3546	3868	8630	9407

Table 11-43. *The shower multiplicity $\Pi^{e^-}_{max}$ at $T^{e^-}_{max}$ for various materials.*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
$P_{av-90/180}$	68	80	44	98	158	255	282	500	284
$P_{av-90/120}$	67	79	43	96	154	250	277	490	279

Table 11-44. *Power deposition per unit length $P'_{av}(W/cm)$ for the two bunch train configurations and various materials before shower development.*

at the beginning of the cascade (Π=1) for the two bunch train configurations (to first approximation energy independent). Table 11-45 give the same values for the two bunch train configurations at T_{max} as a function of beam energy. Note that the values in the two tables are very similar to each other.

It should be pointed out that calculations using the Monte Carlo code EGS gave power deposition values P'_{max} which are about 25–50% below those given in the tables using the equation for the shower maximum, whereas at the beginning of the shower, they are in very close agreement with the results obtained using P'_{av}.

As is readily apparent from the values presented in Table 11-45, removal of the vast amounts of thermal energy generated in solid materials in the region of shower maximum looks nearly impossible for the higher energy cases and the higher Z materials. Therefore, we will concentrate on examining the pros and cons of using water as the principal power absorption and dissipation medium in the beam dump.

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	3.3	5.2	3.3	11.9	30.8	57.1	68.8	268	166
80	6.1	9.8	5.8	22.6	58.6	109.0	132.0	515	319
150	11.0	17.6	10.4	40.7	106	197.0	238.0	933	578
250	17.7	28.0	16.8	65.7	171	319.0	385.0	1516	938
500	33.8	54.2	32.2	126	330	615.0	742.0	2896	1815
750	49.4	79.4	47.2	185	485	904.0	1090	4315	2672

Table 11-45. *A. Maximum power deposition per unit length $P'_{av,max}$ (kW/cm) for various materials as a function of beam energy with 90 bunches/180 Hz/0.85·10^{10}e.*

Material	Be	C	H_2O	Al	Ti	Fe	Cu	W	Pb
E_o=40	3.2	5.1	3.2	11.6	30.0	57.1	67.6	263	163
80	6.0	9.7	5.7	22.1	57.1	107.0	129.0	505	313
150	10.9	17.4	10.2	39.8	103.0	193.0	234.0	915	567
250	17.4	27.7	16.4	64.3	167.0	313.0	378.0	1485	921
500	33.3	53.6	31.5	124.0	321.0	603.0	729.0	2839	1783
750	48.7	78.4	46.1	181.0	472.0	886.0	1071.0	4229	2625

Table 11-45. *B. Maximum power deposition per unit length $P'_{av,max}$(kW/cm) for various materials as a function of beam energy with 90 bunches/120 Hz/1.25·10^{10}e.*

11.A.1 The Beam Dump Vessel

For pulse- or bunch-train repetition rates of 180 Hz and 120 Hz, the train spacings are 5.6 ms and 8.3 ms, respectively. For a bunch spacing of 1.4 ns, the length of one train is 125 ns for 90 bunches and 105 ns for 75 bunches. Since energy is principally lost by ionization from the primary electrons and secondary electrons or delta rays which are highly relativistic, the local microvolume exposed to these rays will have reached final temperature in the time frame of about 10^{-15} s after an individual bunch has passed. This is some three orders of magnitude less than the bunch width, some six orders of magnitude less than the bunch spacing, and some eight orders less than the length of one bunch train. Therefore, in the time frames of the bunch and the bunch train final temperature is reached instantaneously. On the other hand, significant thermal relaxation takes place on a timescale which is long compared to the approximately 100-ns length of a bunch train (and, incidentally, the pressure or shock waves generated by the sudden thermal expansion of the beam-heated microvolume will propagate outward at the velocity of sound in the local medium [~1450 m/s], which is also on a timescale very long compared to the bunch spacing and long compared to the length of the bunch train). So, the pertinent question to ask at the onset of this examination is "What is the expected temperature rise due to one bunch train at T_{max}?".

From experience and from simulations with the Monte Carlo code EGS, we can expect a beam size at T_{max} of $2\sigma_b$ ~0.8 cm for an input size of $2\sigma_b$ ~1 mm and the higher beam energies under consideration. Let us define a heat source term for the region of $0 \le r \le \sigma$ at T_{max} and E_o=500 GeV

$$S \stackrel{\text{def}}{=} \frac{P'_{av,max} C}{(2\sigma_b)^2 \pi / 4} \tag{11.157}$$

where $C = 0.393$ is a constant to reflect the assumed transverse Gaussian beam intensity distribution. Numerically, we find

$$S = \frac{32.2 \cdot 0.393}{0.8\pi/4} \sim 25.2 \frac{kW}{cm^3} \tag{11.158}$$

for the 90-bunch/180-Hz case. Then the temperature rise per bunch train is

$$\Delta T_T = \frac{S}{\rho c_P f_{rep}} \tag{11.159}$$

where ρc_P=4.22 J/cm^3°C is the specific heat capacity for water. This is approximately constant over the temperature interval of interest (it only decreases ~1.5 %). We then find $\Delta T_T \approx 33$°C for the 90-bunch train/180-Hz case at 500 GeV and ≈50°C for 750 GeV.

For comparison, the maximum water inlet temperature in mid-summer at the SLC is ≈30°C (≡ 90°F). So, peak bunch train temperatures at T_{max} would be ≈65 and 80°C. We should recall that the heat source term was defined as an average for the body of rotation bounded by $2\sigma_b$, and therefore somewhat higher values can be expected for a Gaussian distribution near r=0. On the other hand, EGS simulations with a Gaussian input often show a double-humped distribution at a depth of several radiation lengths with a small depression at r=0 resulting from the transverse spread of the electromagnetic cascade. Averaging over 2σ is a reasonable approximation. As a matter of fact, peak temperatures which are somewhat higher than the ones calculated at T_{max} are consistently found at a depth short of T_{max}. The reason for this is that in the early part of the electromagnetic cascade, the increase in shower multiplicity dominates over the increase in the transverse size of the beam on account of the higher average energy of the shower particles and consequential smaller scattering angles. At a depth short of T_{max}, this role is reversed and it is at that depth where peak temperatures are found. The discrepancy appears to be larger in low-Z materials.

The next question which needs to be asked is, "What is the significance of a temperature of 80°C, for example, in the beam-heated microvolume of water?" We can make a categorical statement that volume boiling of water needs to be avoided at all cost. The consequences would be a great reduction in the density of the power absorption medium and therefore a downbeam shift of T_{max}, perhaps no longer fully containing the longitudinal cascade shower within the dump itself. The boiling point of water at, for example, 7 atmospheres (≡100 psig) is ≈160°C. So, assuming for the moment that there is no thermal relaxation during the interbunch train period of 6–8 ms, one could accept a second bunch train and not reach the boiling point at the local pressure; with thermal relaxation, it would be at least three trains. But for continuous operation at 120–180 Hz, we must move the water transverse to the beam momentum vector, fast enough, with some element of conservatism, that no two successive bunch trains target on the same microvolume of water. Suppose we use a water velocity of ~1 m/s, then the distance traveled by the lump of water during the interbunch train period is $V_w \geq 10^3$ mm/s×$5.6 \cdot 10^{-3}$s≈ 5.5 mm. This is almost what we need to avoid significant overlap of the microvolume element bounded by 2σ from two successive bunch trains.

The next question which we need to answer is, "What are the longitudinal and transverse dimensions to guarantee complete containment of the electro-magnetic cascade of a 750 GeV/16-MW electron beam in water?" From EGS and analytical work, we find ~40 X_o are needed to sufficiently attenuate the beam longitudinally. This does not address containment of 750-GeV muons. Of this total, ~15–18 X_o should be water. The remainder could be water-cooled solid materials. For water, $1X_o$≡35.7 cm. Thus, the beam would first take a cold bath in ~5.5–6.5 m of water. An additional 1 m of solids would complete the attenuation of the longitudinal cascade. The total dump length would then be ~7.5 m.

The dump diameter needs to be large enough to attenuate the radial shower. Based on EGS and experience, D_o ~1.5 m of water should be sufficient. The next question to be answered is, "How much flow rate is required to keep the bulk water temperature rise from exceeding ~30–35°C?" From simple calorimetry, we find ~6800 l/min (≡1750 gpm). This should not present any great logistics problems.

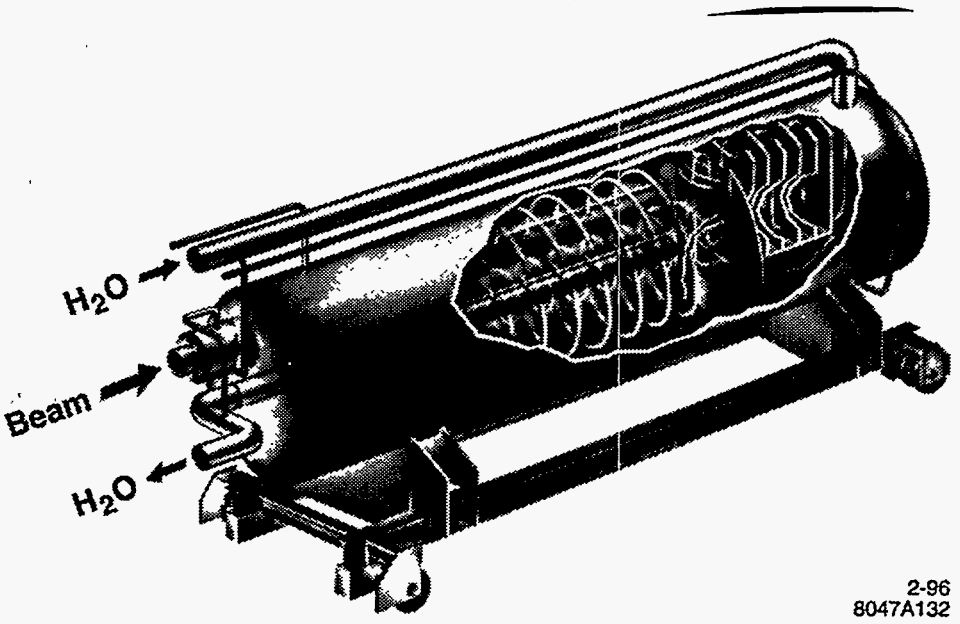


Figure 11-72. *Schematic representation of the primary beam dumps.*

Two beam dumps were originally designed and built at SLAC which use a vortex-like flow arrangement to economically dissipate the power from 2.2-MW, 25-GeV electron beams [Walz 1967a]. A schematic of such a dump is shown in Figure 11-72. Water is injected approximately tangential to the dump wall through a series of evenly-spaced holes from an inlet flow header located at the periphery of the cylindrical shell of the dump vessel. The water then flows spirally toward the center of the dump vessel where the exit manifold is located. In the region of interest, the velocity follows the laws of potential flow theory, namely V×r=const. Introducing the electron beam at a radius of ~0.3 m guarantees a flow velocity transverse to the beam of ~1 m/s for the flow rate of 2000 l/min (≡550 gpm). After traveling one complete revolution in the vortex, the heated microvolume of water has undergone complete mixing and is returning at a lower orbit. No danger exists for volume boiling.

11.A.2 The Window

The electron beam would enter the dump through a thin window separating the water from either vacuum or atmosphere. A number of materials such as aluminum, titanium, or copper appear suitable for this application, the latter being the most attractive. Since beam size and beam excursion from a nominal trajectory are expected to be small for an NLC, the window diameter can be small, say 10–15 cm. A reasonable wall thickness might then be δ=1 mm for a hemispherically-shaped window. The power deposited in the window can then readily be computed by multiplying $P'_{\rm av}$ by the window thickness and using the minimum ionization loss in copper $(-\rho\frac{dE}{dx})$ ~12.8 MeV/cm:

$$\mathrm{P}_w = \mathrm{P}'_{\rm av}\delta = 12.8\cdot 10^6(112.5\cdot 10^{10})120(1.6\cdot 10^{19})0.1 \sim 28\mathrm{W} \qquad (11.160)$$

for the 90-bunch/120-Hz case. An EGS simulation at 50 GeV gave a value ~3% higher. To first approximation, we can neglect heat losses from the window by thermal radiation (and natural convection if there is atmosphere) to the upbeam world. We assume then that all heat is conducted through the window and

removed by forced convection by water on the inside. Let us correct for a double convoluted Gaussian and obtain $P_{w,corr}$=28·0.393=10.9 W for the region $0 \le r \le \sigma_b$. After averaging over $2\sigma_b$ the effective heat flux into the water, neglecting lateral conduction for the moment, becomes $q'' = P_{w,corr}/A_b = 10.86/0.01\pi/4 \equiv 1.38\,\mathrm{kW/cm^2}$. This is in the nucleate boiling heat transfer range and, depending on flow velocity and subcooling, could be close to the transition from nucleate to film boiling. The latter is defined as a burnout condition.

Still neglecting lateral conduction, we find for the maximum temperature difference from the outside to the inside of the window (δ=1-mm thickness and k=3.9 Ws/mm^2°C/cm for copper).

$$\Delta T = \frac{q''\delta}{2}\mathrm{k} = \frac{1.38 \cdot 10^3 \cdot 0.1}{2 \cdot 3.9} \sim 18°\mathrm{C} \quad . \tag{11.161}$$

To get an idea by how much lateral conduction reduces this temperature difference and the heat flux into the water, we use Fourier's Law of conduction

$$k\nabla^2 T + S = \rho c \frac{\partial T}{\partial \tau} \quad . \tag{11.162}$$

For steady state and internal heat generation, this reduces to the familiar Poisson equation

$$\nabla^2 T + \frac{S}{k} = 0 \quad . \tag{11.163}$$

For the axisymmetric case, the Laplacian ∇^2 can be written as

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2} + \frac{S}{k} = 0 \quad . \tag{11.164}$$

For finite cylinder boundary conditions with $0 < z \le 2\delta$, $0 < r \le \sigma_b$ and with heat production at a constant rate S per unit volume per unit time, a solution was given by Carlslaw [Carlslaw 1959] as

$$T = \frac{Sz(2\delta - z)}{2k} - \frac{16\delta^2 S}{k\pi^2} \sum_{n=0}^{\infty} \frac{I_o[(2n+1)\pi r/2\delta]}{(2n+1)^3 I_o\left[(2n+1)\frac{\pi\sigma}{2\delta}\right]} \sin\frac{(2n+1)\pi z}{2\delta} \quad . \tag{11.165}$$

The maximum temperature occurs at r=0, z=δ. Using only the first term (n=0) this reduces (assuming an adiabatic interface on the upbeam face of the window) to

$$\Delta T = \frac{s\delta^2}{2k}\left[1 - \frac{1.032}{I_o\left(\frac{\pi\sigma}{2\delta}\right)}\right] \tag{11.166}$$

where I_o is a modified Bessel function of order zero. S=$0.393 \cdot P'_{av}$=$13.9\,\mathrm{kW/cm^3}$. Then, we find

$$\Delta T_{\max} = \frac{13.86 \times 10^3 \times 0.1^2}{2 \times 3.9}\left[1 - \frac{1.032}{I_0\left(\frac{\pi \times 0.05}{2 \times 0.1}\right)}\right] \sim 2°\mathrm{C} \quad . \tag{11.167}$$

Thus, for this case, lateral conduction is of great help and there are no worries about the heat flux into the water approaching the critical heat flux for the transition from nucleate to film boiling, even if we did not average over $2\sigma_b$ and instead used the real Gaussian distribution.

The last item to be checked for the window is the temperature rise per one bunch train (since we have already concluded that thermal relaxation is too slow to have any effect on individual bunches). Again, we find S=$13.9\,\mathrm{kW/cm^3}$ and $\rho c|_{\mathrm{Cu}}$=3.45 J/cm^3°C

$$\Delta T_{\mathrm{T},W} = \frac{S}{\rho c f_{\mathrm{rep}}} = \frac{13.9 \times 10^3}{3.45 \times 120} \sim 33°\mathrm{C} \quad . \tag{11.168}$$

In a fully restrained body and for $E\alpha|_{Cu}$= 282 psi/°C, we obtain a thermal stress rise

$$\sigma_{th} \approx E\alpha\Delta T = 282 \times 33 \approx 9330 \text{ p}si \quad . \tag{11.169}$$

This would be excessive for fully annealed copper but should not lead to thermal fatigue in semi-hard or hard copper with fatigue strengths ranging from ~13,000 to 18,000 psi for N>10^8 cycles. Of some concern is the effect of cumulative radiation damage in the copper lattice due to the formation of clusters of vacancies and interstices with resultant embrittlement. For the expected current densities every atom in the lattice will "play musical chairs" and change its lattice site several times daily. It may be prudent to schedule a window change on a regular basis and not wait for a perforation to occur.

11.A.3 Isotope Production

The dominant radioisotopes produced in water by these high-energy electron beams are by photospallation on ^{16}O [Neal 1968]. They are in order of decreasing abundance and increasing half life:

$^{16}O \rightarrow {}^{15}O$	$T_{1/2} \sim$ 2 min
$\rightarrow {}^{13}N$(small)	$\sim$ 10 min
$\rightarrow {}^{11}C$	$\sim$ 20 min
$\rightarrow {}^{7}Be$	$\sim$ 53 day
$\rightarrow {}^{3}H$	~12 3 year

For practical purposes, we can neglect ^{15}O since it has mostly disappeared in ~20 min. The isotope which influences tunnel entry and maintenance work the most is ^{11}C. A delay of three hours will allow most maintenance work. ^{7}Be is primarily removed by filtration in the demineralizer loop (not by ion exchange). Tritium builds up very slowly as $^{3}H_2O$ and can be managed by regularly planned disposal of the water as radioactive waste, albeit at a cost. Clearly, compared to solid power absorption materials (metals), water has significant advantages when viewed in light of residual radioactivity.

11.A.4 Radiolysis and Hydrogen Evolution

The deposition of the electron beam's energy in water causes radiolysis and dissociates the water molecule, thereby producing hydrogen, hydrogen peroxide, and oxygen per the following two relationships

$$\begin{aligned} H_2O &\xrightarrow{e^-} H_2^{\uparrow} + H_2O_2 \\ H_2O_2 &\xrightarrow{e^-} O_2^{\uparrow} + H_2O \quad . \end{aligned} \tag{11.170}$$

After saturation and in the absence of oxygen removal processes such as corrosion, hydrogen and oxygen will evolve in gaseous form in places like the surge tank at the stoichiometric mixture ratio.

The measured rate of evolution of free hydrogen [Walz, 1967] is ~0.3l H_2/MW-s [G(H_2)~0.14 molecules H_2/100 eV]. Thus, for P_{av} ~16 MW, the expected rate of evolution would be ~5l H_2, less an allowance for power not directly deposited and dissipated in water or lost from the dump boundaries. The lower explosive limit (LEL) of hydrogen in air is 4%. A conservative maximum allowable concentration would be ~1%.

Thus, some 500l of the 1% H_2 in-air mixture has to be disposed of or processed per second. Since this gas contains most of the radioisotopes given above, release into the atmosphere is not an option.

One could build a delay line including holding tanks which would take care of the ^{15}O isotope. But ^{11}C with its $T_{1/2}$=20.3 min which evolves in the form of $^{11}CO_2$ is not manageable in a cost-effective way by this method. We have also studied the absorption of $^{11}CO_2$ in regenerative lime columns, but that method also proved to be cumbersome. Therefore, we developed at SLAC a catalytic H_2-O_2 recombiner with a capacity to process the radiolytically-evolved hydrogen for average beam powers up to 2 MW [Neal 1968, Walz 1969]. Four such recombiner facilities have been in use for more than 25 years with very low maintenance. Such a recombiner can be readily scaled up to process the gas from a 16-MW beam.

The recombiner uses a platinum-palladium catalyst through which the gases are pumped. The water vapor generated in the recombination of H_2 and O_2 is condensed out in the water spray which is generated by an ejector pump that provides the motive power to recirculate the gases. It is a closed-loop system and all gases and radioisotopes are always fully contained.

11.A.5 Summary

We showed that a beam dump to absorb and dissipate the power of 750 GeV/16-MW NLC beams is feasible using water as the primary power absorption medium. Such a dump could be a cylindrical vessel housing a potential vortex-like flow of water with the beam impinging at a radius of ~0.3 m and a vortex flow velocity ~1–1.5 m/s normal to the direction of the beam momentum vector. The vessel dimensions might be on the order of 1.5-m diameter and ~6.5 m (~18 X_o)-long for the water section followed by ~1 m of water-cooled solid materials of higher Z to adequately attenuate the electro-magnetic cascade. The beam would enter the dump through a thin window (~1-mm-thick copper would do). Isotope production in the water appears manageable relative to tunnel access requirements for maintenance and seems preferable to what can be expected for solid materials. The only drawback is production of ~5l H_2/16 MW from radiolysis in water, but solutions exist to deal with this, notably the catalytic recombination of H_2 and O_2 into water, allowing for a closed-loop system containing all the radioisotopes.

References

[Burke 1995] D. Burke, "Proposal for New NLC Operating Plane", SLAC memorandum, unpublished (1995).

[Adolphsen 1995] C. Adolphsen, private communication (1995).

[Bernadini 1963] C. Bernadini *et al., Phys. Rev. Lett.* **10**, 407 (1963).

[Bambade 1989] P. Bambade *et al.*, "Observation of Beam-Beam Deflections at the SLAC Linear Collider", *Phys. Rev. Lett.* **62**, 2949 (1989); F. Zimmermann *et al.*, "Performance of the 1994/95 SLC Final Focus System", SLAC–PUB–95–6790, in *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Bane 1995] K. Bane, C. Adolphsen, F.-J. Decker, P. Emma, P. Krejcik, F. Zimmermann, "Measurement of the Effect of Collimator Generated Wakes on the Beams in the SLC", SLAC–PUB–95–6865, in *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Berz 1989] M. Berz, "Differential-algebraic description of beam dynamics to very high orders", *Particle Acclerators* **24**, 109 (1989).

[Bini 1991] C. Bini *et al., Phys. Lett.* **B** 135 (1991).

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

[Brown 1977] K. Brown, F. Rothacker, D. Carey, C. Iselin, "TRANSPORT - A Computer Program for Designing Charged Particle Beam Transport Systems", SLAC–91 (1977).

[Brown 1979] K. Brown and R. Servranckx, "Chromatic Correction for Large Storage Rings," SLAC–PUB–2270, in *IEEE Trans. Nucl. Sci.* **26**, 3598 (1979).

[Bula 1995] C. Bula, K.T. McDonald, E.J. Prebys (Princeton U.), C. Bamber, S. Boege, T. Kotseroglou, A.C. Melissinos, D.D. Meyerhofer and W. Ragg (Rochester U.), D.L. Burke, R.C. Field, G. Horton-Smith, A.C. Odian, J.E. Spencer, D. Walz (SLAC), S.C. Berridge, W.M. Bugg, K. Smakov and A.W. Wiedemann (Tennessee), "Observation of Nonlinear Effects in Compton Scattering", Submitted to *Phys. Rev. Lett.* , (Nov. 1995).

[Carlslaw 1959] Carlslaw, H.S. and Jaeger, J.C., "Conduction of Heat in Solids", (Oxford University Press, London, 1959).

[Carey 1982] D.C. Carey *et al.*, "DECAY TURTLE", SLAC Report 246 (March 1982).

[Chen 1993] P. Chen, "Disruption Effects from the Collision of Quasi-Flat Beams", SLAC–PUB–6215 (1993).

[delCorral 1982] J. del Corral, J.E. Spencer, S. St. Lorant, "Effects of Strong Transverse Fields on Permanent Magnets", SLC-CN-189 (1982).

[Dragt 1976] A.J. Dragt and J.M. Finn, "Lie Series and Invariant Functions for Analytic Symplectic Maps", *J. Math. Phys.* **17**, 2215–2227 (1976).

[Emma 1992] P. Emma, "Beam-Based Alignment of Sector-1 of the SLC Linac", *Proc. of 3rd EPAC*, Berlin (1992)

[Field NIM] Clive Field, "Wire Scanner System for the FFTB", NIM, To Be Published. See also SLAC–PUB–6717.

[Forest 1989] É. Forest, M. Berz and J. Irwin, "Normal form methods for complicated periodic systems: a complete solution using differential algebra and Lie operators", *Particle Acclerators* **24**, 91 (1989).

[Forest] The code DESPOT was written by É. Forest; the code TRACY by J. Bengtsson.

[Gross 1994] G. Gross, J. Spencer and SLAC's MMG, "Experience with the SLC Permanent Magnet Multipoles", Fourth European Part. Acc. Conf., London, England (1994). Also see SLAC–PUB–6558.

[Hartman priv] S. Hartman and T. Shintake, private communication.

[Hartman priv] S. Hartman, T. Shintake, N. Akasaka "Nanometer Resolution BPM Using Damped Slot Resonator", SLAC–PUB–95–6908, presented at 1995 Part. Acc. Conf., (1995).

[Hendrickson 1995] L. Hendrickson, private communication and this ZDR, Appendix D (1995).

[Hertzbach 1995] S. Hertzbach, private communication (1995).

[Hirata 1989] K. Hirata, B. Zotter, and K. Oide, *Phys. Lett. B* **224**, 437 (1989).

[Hollebeek 1981] The primary reference on disruption effects and their simulations is: Robert Hollebeek, "Disruption Limits for Linear Colliders", *Nucl. Instr. and Methods* **184**, 333 (1981).

[Irwin 1991] J. Irwin, "Final Focus System Optimization", *Proc. of Linear Colliders. 3rd International Workshop*, LC 91, Protvino, USSR, **Vol. 3**, V. Balakin *et al.*, (ed.) (1991).

[Irwin 1992] J. Irwin, "The Application of Lie Algebra Techniques to Beam Transport Design", SLAC–PUB–5315, published in *Nucl. Instr. and Methods* **A298**, 460 (1990).

[Irwin 1995a] These conditions were derived by J. Irwin (1995).

[Irwin 1995b] J. Irwin, presentation at SLAC (1995).

[Juravlev 1995] V.M. Juravlev, P.A. Lunev, A.A. Sery, A.I. Sleptsov, K. Honkavaara, R. Orava, E. Pietarinen, "Seismic Conditions in Finland and Stability Requirements for the Future Linear Collider", (1995).

[Keller 1993] L. Keller, "Muon Background in a 1-TeV Linear Collider", SLAC–PUB–6385 (1993).

[Lomperski 1993] M. Lomperski, "Compton Scattering off Blackbody Radiation and other Backgrounds of the HERA Polarimeter", *DESY* **93-045** (1993).

[Neal 1968] Neal, R.B., Editor, "The Stanford Two-Mile Accelerator", Chapter 20, (W.A. Benjamin, Inc., New York, 1968).

[Norem 1995] J. Norem *et al.*, "Tests of a High Resolution Beam Profile Monitor", *Proc. of the 1995 Part. Accel. Conf.*, Dallas, Texas, (1995).

[Oide 1988] K. Oide, "Synchrotron-Radiation Limit on the Focusing of Electron Beams", *Phys. Rev. Lett.* **61**, 15 (1988) 1713.

[Oide 1992] K. Oide, "Final Focus System with Odd-Dispersion Scheme", *KEK Preprint* **92–58** (1992).

[Palmer 1990] R.B. Palmer, "Prospects for High Energy e^+e^- Linear Colliders", *Ann. Rev. Nucl. Sci.* **40**, 529 (1990).

[Pitthan 1995] R. Pitthan, "Re-Alignment: It is the Tunnel Floor which Moves, isn't It?", SLAC–PUB–95–7043, invited talk at 4th Int. Workshop on Acc. Alignment, Tsukuba (1995).

[Piwinski 1985] A. Piwinski, "Beam Losses and Lifetime", in CERN Accelerator School, Gif-sur-Yvette, France, *CERN* **85–19** Vol. II (1985).

[Raimondi 1995] P. Raimondi, F.J. Decker, P. Chen, "Disruption Effects on the Beam Size Measurement," SLAC–PUB–95–6882.

[Raimondi 1993] P. Raimondi, P.J. Emma, N. Toge, N.J. Walker, V. Ziemann, "Beam Based Alignment of the SLC Final Focus Superconducting Final Triplets", SLAC–PUB–95–6212. in *Proc. 1993 Part. Accel. Conf.*, Washington, DC (1993).

[Raimondi 1995] P. Raimondi, private communication (1995).

[Rohatgi 1987] R. Rohatgi, H.A. Schwettman and T.I Smith, "A Compact Energy Recovered FEL for Biomedical and Material Science Applications", *Proc. 1987 Part. Accel. Conf.*, Washington, DC, 1, 230 (1987).

[Rossi 1952] B. Rossi, *High-Energy Particles* (Prentice Hall, New York, 1952).

[Ross 1991] M.C. Ross *et al.*, "Wire Scanners for Beam Size and Emittance Measurements at the SLC", SLAC–PUB–5556 (May 1991).

[Ross 1994] M. Ross, *Proc. of Advanced Accelerator Concepts Workshop*, Lake Geneva (1994).

[Roy 1992] G. Roy, "Analysis of the Optics of the Final Focus Test Beam Using Lie Algebra Based Techniques", SLAC–397 (1992).

[Sands 1985] M. Sands, SLAC/AP–47 (1985).

[Servranckx 1990] R. Servranckx, K.L. Brown, L. Schachinger and D. Douglas, "User's Guide to the Program DIMAD", SLAC–0285 (1990).

[Shintake 1994] T. Shintake, "First Beam Test of Nanometer Spot Size Monitor Using Laser Interferometry", KEK-PREPRINT–94–129 (1994).

[Spencer 1995] J.E. Spencer, private communication (1995).

[Spencer 1995a] J.E. Spencer, "The SLC as a Second Generation Linear Collider", *Proc. 16th IEEE Part. Accel. Conf. and Int'l. Conf. on High Energy Accel's.*, Robt. Siemann ed., Dallas, TX, (1995).

[Spencer 1995b] J. Spencer, J. Irwin, D. Walz and M. Woods, "The SLAC NLC Extraction and Diagnostic Line", 1995 Part. Accel. Conf., Dallas, TX (1995).

[Spencer 1995c] J.E. Spencer, and "Beam-Beam Effects and Generalized Luminosity", SLAC–PUB–7051, in *Proc's: Workshop on Electron-Electron Colliders*, C. Heusch ed., Santa Cruz, CA, (1995);

[Telnov 1987] V.I. Telnov, "Scattering of Electrons on Termal Radiation Photons in Electron-Positron Storage Rings", *Nucl. Instr. Methods A* **260**, 304 (1987).

[Tenenbaum 1995] P. Tenenbaum, D. Burke, R. Helm, J. Irwin, K. Oide, and K. Flöttmann, "Beam-Based Magnetic Alignment of the Final Focus Test Beam", SLAC–PUB–95–6769 (1995).

[Walker 1993] N.J. Walker, J. Irwin, M. Woodley, "Global Tuning Knobs for the SLC Final Focus", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC (1993).

[Walker 1987] R.P. Walker, "Calculation of the Touschek Lifetime in Electron Storage Rings", Proceedings of IEEE PAC 1987, Washington, DC, 491 (1987).

[Walz, 1967] D.R. Walz and E.J. Seppi, "Radiolysis and hydrogen evolution in the A-beam dump radioactive water system", Report No. SLAC–TN–67–29, Stanford Linear Accelerator Center, (October, 1967).

[Walz 1967a] D.R. Walz *et al.*, "Beam Dumps, energy slits and collimators at SLAC—their final versions and first performance data", *IEEE Trans. Nucl. Sci.* **3** NS-14, 923–927 (1967).

[Walz 1969] D.R. Walz and L.R. Lucas, "The 'sphere dump'—a new low-cost high-power beam dump concept and a catalytic hydrogen-oxygen recombiner for radioactive water systems", *IEEE Trans. Nucl. Sci.* **3** NS-16, 613–617 (1969).

[Wilson 1995] P. Wilson supplied the scaling laws of Table 11-38, and the analysis of the consequences of thermal changes of the waveguide.

[Woodley 1994] M.D. Woodley, unpublished (1994).

[Woods 1994] M. Woods, "Polarization at SLAC", SLAC–PUB–6694, (October 1994).

[Yokoya 1986] K. Yokoya, ABEL, "A Computer Code for the Beam-Beam Interaction in Linear Colliders", *Nucl. Instr. and Methods* **B251** 1 (1986); Toshiaki Tauchi *et al., Particle Acclerators* **41**, 29 (1993). We thank Mike Ronan for providing the latest version of ABEL (Analysis of Beam-beam Effects in Linear colliders).

[Yokoya 1988] K. Yokoya and P. Chen, "Depolarization Due to Beam-Beam Interaction in Electron-Positron Linear Colliders", SLAC–PUB–4692 (1988). See also: US-CERN School on Part. Accel's. (1990).

[ZEUS/H1] See numerous papers by ZEUS or H1 collaborations on photoproduction in HERA, *e.g.*, C. Kiesling, "Physics from the First Year of H1 at HERA", DESY–94–137 (1994).

[Zimmermann 1995] F. Zimmermann, R. Helm and J. Irwin, "Optimization of the NLC Final Focus System", SLAC–PUB–95–6791, presented at 1995 Part. Acc. Conf., (1995).

[Zimmermann 1995b] F. Zimmermann *et al.*, "Performance of the 1994/95 SLC Final Focus System", SLAC–PUB–95–6790 (1995). presented at 1995 Part. Acc. Conf., (1995).

Contributors

- Gordon Bowden
- Paul Emma
- Karl Brown
- Dave Burke
- Leif Eriksson
- Dick Helm
- Stan Hertzbach
- John Irwin
- Eugene Kraft
- Robert Messner
- Yuri Nosochkov
- Katsunobu Oide
- Tor Raubenheimer
- Jim Spencer
- Peter Tenenbaum
- Fang Tian
- Dieter Walz
- Perry Wilson
- Mike Woods
- Mark Woodley
- Frank Zimmermann

12

The Interaction Region

Contents

12.1 Introduction **816**
12.2 The Luminosity Spectrum **817**
12.2.1 Beam Energy Spread 817
12.2.2 Initial State Radiation 817
12.2.3 Beamstrahlung 818
12.2.4 Measurement of the Luminosity Spectrum 822
12.3 Detector Background Sources **823**
12.3.1 Beamstrahlung-produced e^+e^- Pairs 823
12.3.2 Hadronic Backgrounds from $\gamma\gamma$ Interactions 827
12.3.3 Quadrupole and Bend Synchrotron Radiation 827
12.3.4 Muon Backgrounds 832
12.4 Detector Issues **836**
12.4.1 Effect of Backgrounds on the Detector 837
12.4.2 Vibration Suppression for the Final Focus Quadrupoles 844
12.4.3 Measurement of Sub-nm Displacements by means of a Laser Interferometer 847
12.4.4 An Optical Anchor for the Final Quadrupoles 849
12.4.5 SLD Final Focus Quadrupole Vibration Measurements 850
12.4.6 Measurement of Polarization and Beam Energy 850
12.5 Conclusions **851**

12.1 Introduction

Table 12-1 summarizes some of the machine design parameters important to the physics capabilities of the NLC and to the design of the interaction region (IR) and detector. Three sets of parameters are considered for each of the two machine energies. They define a volume in parameter space within which the luminosity is roughly constant.

Depending on the exact performance of each of the NLC's subcomponents, we will find ourselves somewhere in that space. The rf power system controls the bunch charge and the number of bunches that can be accelerated each machine cycle. The bunch compression system and requirements on momentum spread will determine the minimum bunch length, σ_z. The final doublet β functions, β_x and β_y, must be larger than σ_z. The invariant emittance in y depends on the performance of the damping rings and on how well the linac and final focus can transport a low-emittance beam. The final value of ϵ_y achieved, together with β_y, will determine the value of the y spot size, σ_y. This in turn will set the scale for vibration tolerance and field stability of the final quadrupoles. Assuming the damping rings provide the specified ϵ_x, β_x is a semifree parameter, adjusted to give the desired luminosity while keeping deleterious beam-beam interaction effects to an acceptable level.

The choice of X-band rf for the NLC sets the interbunch separation at 1.4 ns (or 42 cm). To have each bunch interact only with its partner, there must be a crossing angle at the IP. This crossing angle is put into the NLC at the big bend, just after the collimation section located after the linac and before the final focus. The bend angle then also helps to reduce the muon flux that results when the beam interacts with the collimators. The size of the collimator apertures are measured in terms of the number of beam widths. Physical limitations and wakefield effects imply minimum collimator apertures corresponding to $7\sigma_x \times 35\sigma_y$. Beam tails and these apertures determine the level of muons produced. The apertures, the final-focus lattice, and assumptions on what the non-Gaussian profile of the beam may be determine the production of synchrotron radiation (SR). The design of the masking system and the value of detector's solenoidal field control the backgrounds caused by the SR photons.

The high charge density in each bunch causes particles in one bunch to interact with the overall field of the opposing bunch. This beam-beam interaction results in a luminosity enhancement as the beams are attracted to each other. The resulting acceleration, however, results in the copious production of photons, which will smear the luminosity spectrum as a function of $\sqrt{s}$. The photons can themselves interact coherently with the field of the opposing bunch, or interact with the individual $e^{\pm}$ of the opposing bunch to produce e^+e^- pairs. While predominately produced at low p_t, these pairs can cause problems in the detector and must also be controlled by the solenoidal field and the masking. The $\gamma\gamma$ interactions can cause the production of jets of hadrons with high p_t. The problem is exacerbated by the 1.4-ns bunch structure of the NLC. Particle detectors with good timing resolution will be required to separate the background hits arising from other bunches in the train from those hits produced in the interaction that caused the trigger. Trigger schemes may need to be developed to control the rate at which these backgrounds trigger the detector.

The issue of dealing with the 3.6–7.0-nm y-spot sizes strongly affects the discussion of the IR. The source terms in the problem are the naturally occurring ground vibrations and ground motion driven by local laboratory sources, such as pumps and the flow of cooling fluids. The support structures which stabilize the final doublet against vibration at the nm level will reside within the detector. Any device that might be required to sense or control the inertial or relative movement of the the quadrupoles must be accommodated by the detector. Finally, detectors usually use various fluids to cool their magnet coil and electronics; the influence of fluid flow on the final-doublet vibration will need to be understood.

	500 GeV			1 TeV		
	A	B	C	A	B	C
Trains/s		180			120	
Bunches/train		90			90	
Interbunch spacing (ns)		1.4			1.4	
Electrons/bunch (10^{10})	0.65	0.75	0.85	0.95	1.1	1.25
σ_x-spot size (nm)	256	286	286	202	221	256
σ_y-spot size (nm)	4.23	5.25	6.71	3.39	4.11	5.15
σ_z-spot size (μm)	100	125	150	125	150	150
x beam divergence (μrad)	32.0	28.6	28.6	20.2	18.5	16.0
y beam divergence (μrad)	33.8	35.0	33.5	27.1	27.4	25.8
Minimum collimator apertures	$7\sigma_x \times 35\sigma_y$			$7\sigma_x \times 35\sigma_y$		
$\mathcal{L}$ (geometric) (10^{33} cm^{-2} s^{-1})	5.0	4.8	4.9	11.3	11.4	10.2
H_D - Luminosity Enhancement Factor	1.46	1.45	1.52	1.41	1.43	1.53
$\mathcal{L}$ (with pinch) (10^{33} cm^{-2} s^{-1})	7.3	7.0	7.4	16.0	16.4	15.6

Table 12-1. *List of the NLC parameters that define the operating range of the interaction region and detector.*

12.2 The Luminosity Spectrum

Ideally, the beam energy at the NLC would be a delta function at the design energy. It could then trivially be used as a constraint in any physics analysis. The finite energy spread of the NLC, initial-state radiation effects, and the production of photons in the beam-beam interaction will degrade the delta function to a spectrum. The spectrum can be unfolded from any physics analysis if it is not too broad, where "too" is determined by the specific physics channel of interest, and if we can accurately measure the spectrum.

12.2.1 Beam Energy Spread

Figure 12-1 shows the expected beam energy spread at the IP for the 1-TeV machines. The shape comes from the bunch compression that sets the beam spot size in z. It is similar for the 500-GeV machine.

12.2.2 Initial State Radiation

At the NLC, the effects of initial-state radiation (ISR) will be about the same as that at SLC or LEP II. The scale of the problem is set by the variable $L = \ln \frac{s}{m_e^2}$. L varies from 24.2 to 27.6, 29.0, or 29.8 as $\sqrt{s}$ goes from m_z to 500, 1000, or 1500 GeV. The amount of ISR is irreducible and therefore sets the scale for how small the luminosity smearing due to the beam energy spread and the beamstrahlung must be. The electron energy distribution, $D_e(z,s)$, in the presence of ISR has been calculated by Fadin and Kuraev [Kuraev, 1985]. (See their Eqs. (20) and (21)). Here $\sqrt{s}$ is the center-of-mass energy of the interacting e^+ and e^- and z is the fractional e^+ or e^- energy following ISR.

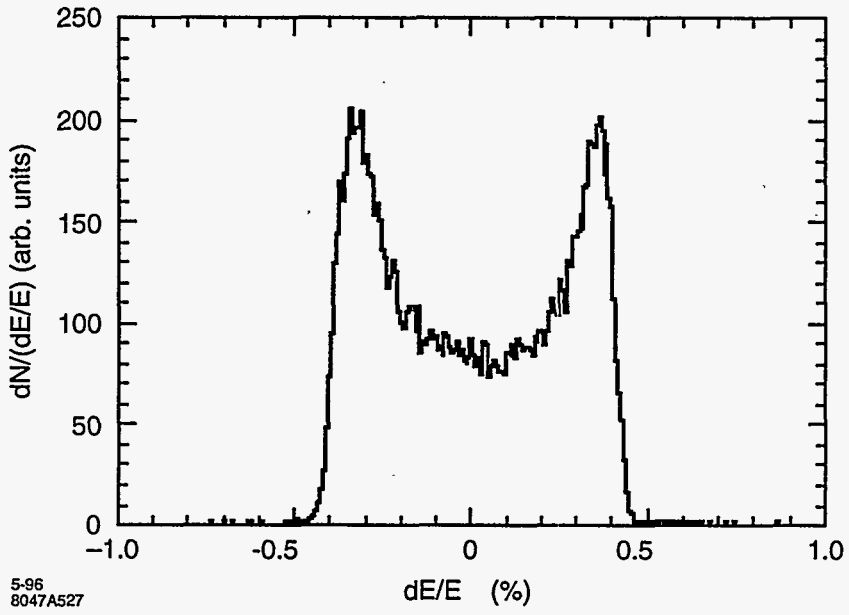


Figure 12-1. *Beam energy spread for the 1-TeV lattice.*

The cross section has the form,

$$\sigma(s) \propto \int\int dx_1 dx_2 D(x_1, s) D(x_2, s) \sigma(s x_1 x_2). \tag{12.1}$$

Defining $s x_1 x_2 = (1 - x)s$ and integrating at fixed x yields,

$$\sigma(s) = \int dx \sigma(s(1 - x)) F(x, s), \tag{12.2}$$

where $F(x, s)$ (Eq. (28) of Fadin & Kuraev) represents the luminosity distribution function.

The dashed curve in Figure 12-2 shows the electron momentum distribution function, $D_e(z, s)$, plotted for $\sqrt{s}$=500 GeV as a function of $z = p_{\text{beam}}/p_{\text{max}}$, the fractional lepton momentum. The solid curve shows the luminosity spectrum $F(x, s)$ for $\sqrt{s}_{\text{max}} = 500\,\text{GeV}$, plotted as a function of $z = \sqrt{s}/\sqrt{s}_{\text{max}}$. Here $1 - z = \sqrt{1 - x}$. Each curve is normalized to unit area. When plotted as in Figure 12-2, one cannot distinguish any difference in either the electron momentum distribution function or in the luminosity distribution over the range of $\sqrt{s}$ considered. Any small differences can be seen by examining Table 12-2, which shows the fraction of the luminosity within 0.1%, 0.5%, 1%, 5%, and 10% of the nominal center-of-mass energy, as well as the mean energy loss and the rms energy spread, for each of the four values of $\sqrt{s}$ considered.

12.2.3 Beamstrahlung

As one bunch of beam particles passes through the electromagnetic field of the opposing bunch, radiation will be emitted. The photons thus produced are called "beamstrahlung" photons. The subject has been extensively discussed in the literature. The results herein are taken from Ref. [Chen, 1990] and Ref. [Chen, 1992].

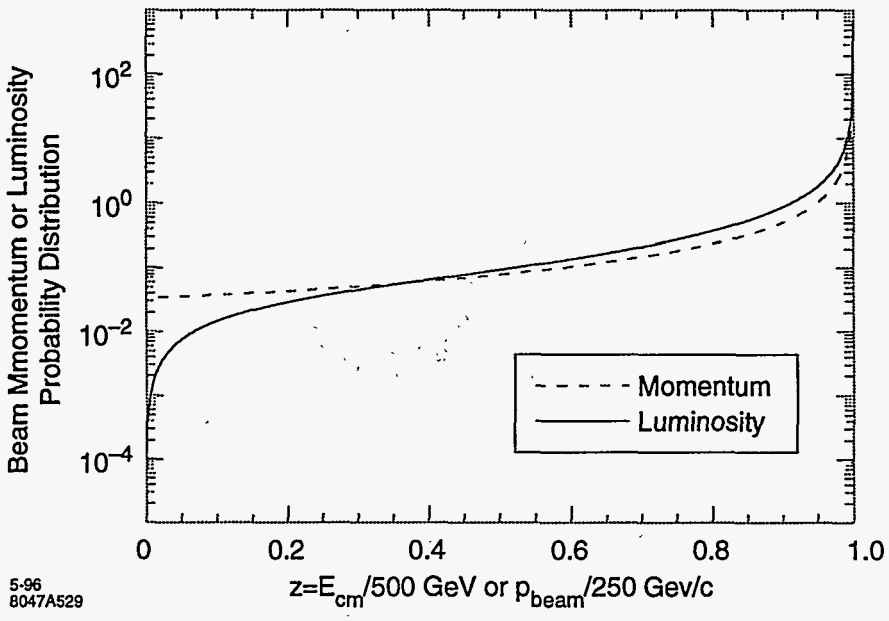


Figure 12-2. *The dashed curve shows the electron momentum distribution as a function of the normalized electron momentum after ISR. The solid curve shows the luminosity distribution after ISR as a function of the normalized center of mass energy. While the plot corresponds to a center of mass energy of 500 GeV, no difference would be visible for any center of mass energy between m_Z and 1.5 TeV.*

	m_Z	500 GeV	1 TeV	1.5 TeV
Mean e^+e^- c.m. energy loss	4.19%	4.77%	5.00%	5.14%
Rms e^+e^- c.m. energy spread	11.2%	11.9%	12.2%	12.4%
% of $\mathcal{L}$ within 0.1% of $\sqrt{s}_{\rm nom}$	54.7%	50.1%	48.3%	47.3%
% of $\mathcal{L}$ within 0.5% of $\sqrt{s}_{\rm nom}$	65.2%	61.2%	59.6%	58.7%
% of $\mathcal{L}$ within 1% of $\sqrt{s}_{\rm nom}$	70.2%	66.6%	65.2%	64.4%
% of $\mathcal{L}$ within 5% of $\sqrt{s}_{\rm nom}$	82.8%	80.5%	79.6%	79.1%
% of $\mathcal{L}$ within 10% of $\sqrt{s}_{\rm nom}$	88.3%	86.7%	86.0%	85.6%

Table 12-2. *Effect of ISR on e^+e^- luminosity spectra.*

The production of beamstrahlung photons is a stochastic process. The probability that a given particle will radiate and the characteristic energy of the radiation are determined by the field density or equivalently by the number of electrons or positrons per bunch and the bunch dimensions. When this density is high relative to the critical electric field, beamstrahlung is more likely to occur. The critical electric field, $E_{\rm critical}$, is defined as that field which, when an electron travels one Compton wavelength in it, does an amount of work equal to $m_e c^2$:

$$eE_{\rm critical}\lambda_c \equiv m_e c^2 \tag{12.3}$$

The magnetic field equivalent is:

$$B_{\rm critical} = \frac{\alpha e}{r_e^2} \sim 4.4 \times 10^9 \text{ Tesla} \tag{12.4}$$

The field strength is measured in terms of the dimensionless quantity

$$\Upsilon \equiv \frac{\gamma B_{\rm bunch}}{B_{\rm critical}} = \frac{E_{\rm bunch}}{E_{\rm critical}} \tag{12.5}$$

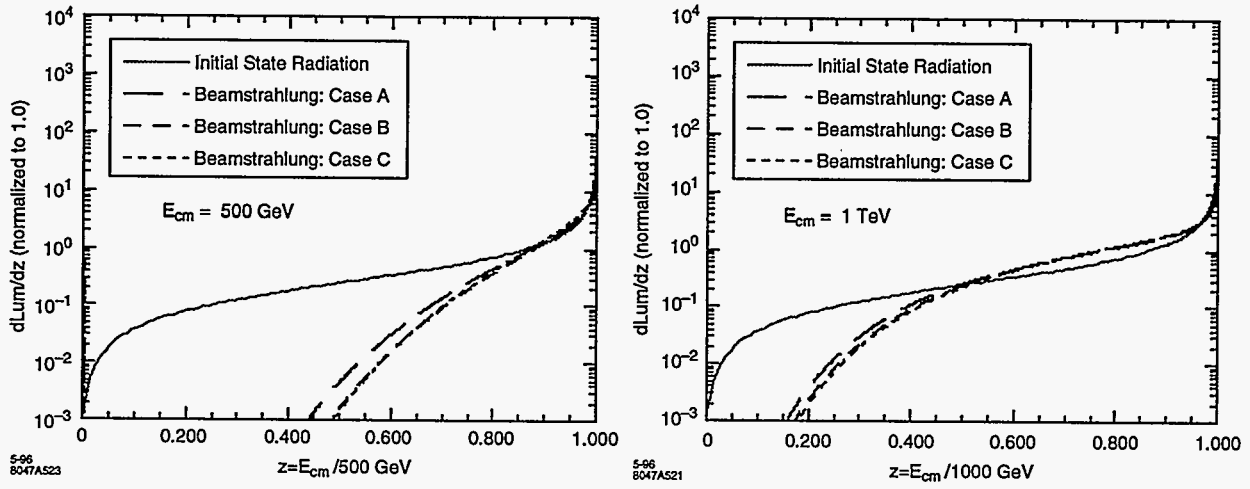


Figure 12-3. *The luminosity spectrum after beamstrahlung at 500 GeV and 1 TeV for the three parameter sets under consideration. The spectrum due to initial state radiation is presented for comparison.*

where

$$B_{bunch} \cong \frac{5eN_e}{6\sigma_z\left(\sigma_x+\sigma_y\right)} \quad . \tag{12.6}$$

In Ref. [Chen, 1992], Chen derives an expression (see his Eq.(24)) for the differential luminosity which is composed of three parts, corresponding to the cases when either no, one, or more than one photon is emitted. Each part is expressed in terms of N_γ, the mean number of photons produced per electron,

$$N_\gamma = \frac{5}{2}\frac{\alpha\sigma_Z}{\gamma\lambda_c}\Upsilon\left(1+\Upsilon^{2/3}\right)^{-1/2}. \tag{12.7}$$

The shape of the curve is generally described by N_γ, by the average energy loss per electron, δ_b

$$\delta_b = \frac{1}{2}N_\gamma\Upsilon\frac{(1+\Upsilon^{2/3})^{1/2}}{(1+(\frac{3}{2}\Upsilon)^{2/3})^2} \quad , \tag{12.8}$$

and by the fraction of the luminosity at the full nominal beam energy,

$$\frac{1}{N_\gamma^2}\left[1-e^{N_\gamma}\right]^2 \quad . \tag{12.9}$$

In Figure 12-3, we plot the luminosity spectra for the three parameter sets at 500 GeV and 1 TeV. The ISR curve is added for comparison. In Figure 12-4, the region of the luminosity spectra within 2% of the nominal beam energy is plotted for the three parameter sets at 500 GeV and 1 TeV. Figure 12-5 presents a comparison of Case A for the 500-GeV and 1-TeV machines.

Table 12-3 summarizes the situation by listing the values of Υ, N_γ, δ_b, and the fraction of the luminosity within 0%, 0.1%, 0.5%, 1%, 5%, and 10% of the nominal center of mass energy. The fraction at 0% is taken from the above expression, while the others are obtained by integrating the distributions in Figure 12-3,

The machine parameter sets have been chosen so that the beamstrahlung induced smearing of the luminosity spectrum is not substantially worse than the luminosity smearing due to ISR. As the ISR part of the smearing

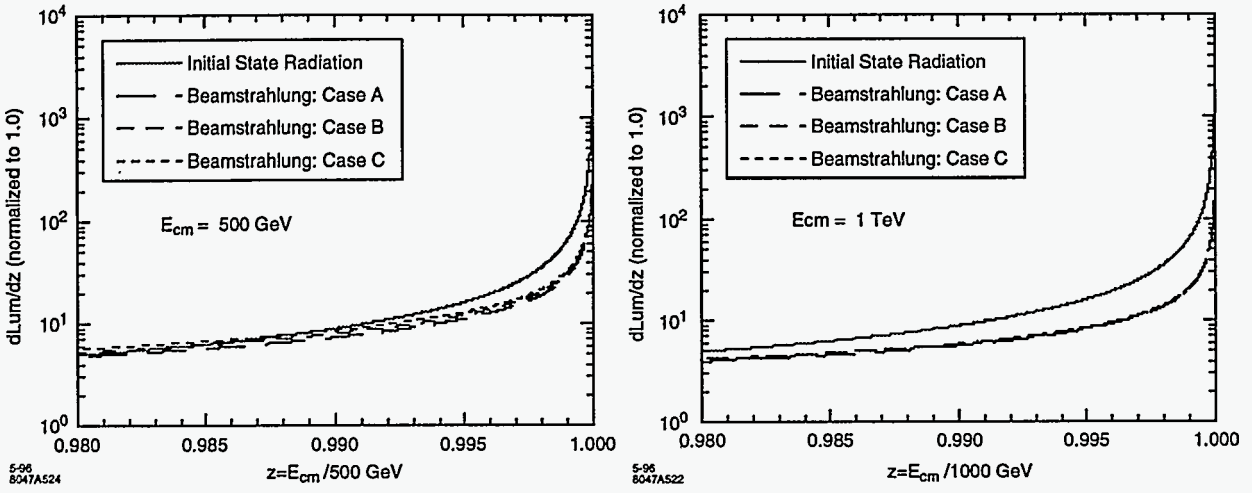


Figure 12-4. *A closeup of the region of the luminosity spectrum due to beamstrahlung near the nominal beam energy at 500 GeV and 1 TeV for the parameter sets listed previously. The spectrum due to ISR is presented for comparison.*

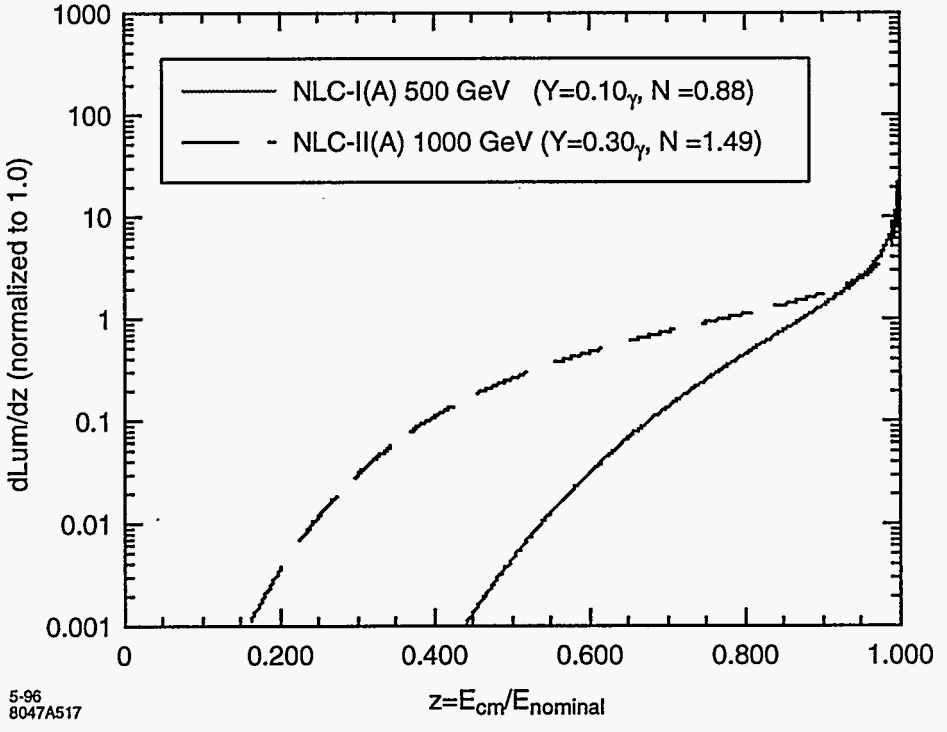


Figure 12-5. *A comparison of the beamstrahlung spectra for case A of the 500-GeV and 1-TeV machines.*

	500 GeV			1 TeV		
	A	B	C	A	B	C
Υ	0.112	0.0924	0.0878	0.334	0.296	0.289
N_γ	0.973	1.017	1.162	1.654	1.781	1.745
δ_b	3.55%	3.21%	3.52%	12.6%	12.6%	12.2%
R.M.S. e^+e^- c.m. energy spread	7.17%	6.44%	6.62%	15.9%	15.5%	15.2%
% of $\mathcal{L}$ within 0% of $\sqrt{s}_{\rm nom}$	40.9%	39.4%	35.0%	23.9%	21.8%	22.4%
% of $\mathcal{L}$ within 0.1% of $\sqrt{s}_{\rm nom}$	48.1%	47.0%	42.7%	28.6%	26.4%	27.2%
% of $\mathcal{L}$ within 0.5% of $\sqrt{s}_{\rm nom}$	54.5%	54.0%	50.0%	33.1%	31.0%	31.8%
% of $\mathcal{L}$ within 1% of $\sqrt{s}_{\rm nom}$	59.0%	58.8%	55.0%	36.3%	34.2%	35.1%
% of $\mathcal{L}$ within 5% of $\sqrt{s}_{\rm nom}$	75.6%	76.8%	74.4%	49.8%	48.1%	49.2%
% of $\mathcal{L}$ within 10% of $\sqrt{s}_{\rm nom}$	85.6%	87.2%	85.9%	60.2%	59.1%	60.1%
% of $\mathcal{L}$ within 100% of $\sqrt{s}_{\rm nom}$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 12-3. *Summary of quantities parameterizing the effect of beamstrahlung on the luminosity spectrum.*

is calculable and small, it should be comparatively easy to unfold from the data. Comparing Tables 12-2 and 12-3 we find that at 500 GeV all three parameter set choices result in smearing below the level expected from ISR. At 1 TeV, there is roughly twice as much smearing as that from ISR.

12.2.4 Measurement of the Luminosity Spectrum

The significance of the luminosity smearing depends on the physics channel under study. For discovery physics, some measure such as the fraction of the luminosity with $\sqrt{s} > 90\%$ of nominal is probably the best figure of merit. However, attention is generally focused on the region within 1% of the nominal energy, as it is critical for an excellent determination of the top quark mass. D. Miller has argued [Frary] that the measurement of the top quark mass will require machine parameters that result in very little smearing and that the detector have excellent forward tracking so as to use the collinearity distribution of Bhabha events to unfold the luminosity spectrum.

In Figure 12-6, we show the cross section for $t\bar{t}$ production as a function of nominal center-of-mass energy for $m_t = 180\,\mathrm{GeV/c^2}$. The theoretical cross section, indicated as curve (a), is based on the results of Peskin and Strassler [Peskin, 1991] with $\alpha_s(M_Z^2) = 0.12$, infinite Higgs mass, and nominal Standard Model couplings. Each energy-smearing mechanism, initial-state radiation (b), beamstrahlung (c), and beam energy spread (d), has been successively applied. Hence, curve (d) includes all effects.

A comment on the beam energy spread is in order. The expected shape of the single-beam energy spread is given in Figure 12-1. The luminosity-weighted center-of-mass energy spread, $\Delta E_{\rm cm}/E_{\rm cm}$, is calculated from the single-beam distribution given its dependence on the bunch spatial distribution, as discussed in Chapter 4. The resulting center-of-mass energy distribution is given in Figure 12-7, where the calculation has been simplified by ignoring variations in energy in transverse space. In this case, the input single-beam energy spread has FWHM of 0.8%, corresponding to the distribution shown in Figure 12-1. The resulting distribution in $\Delta E_{\rm cm}/E_{\rm cm}$ is strongly peaked at zero with an rms of 0.38%, as indicated in the Figure 12-1. It is expected that the single-beam energy spread can be comfortably adjusted within the FWHM interval 0.6% to 1.0%. For the top threshold, it is clear that the smaller width is preferred, and curve (d) of Figure 12-6 was calculated using the 0.6% width. The large top mass of about $180\,\mathrm{GeV/c^2}$ presents a relatively broad,

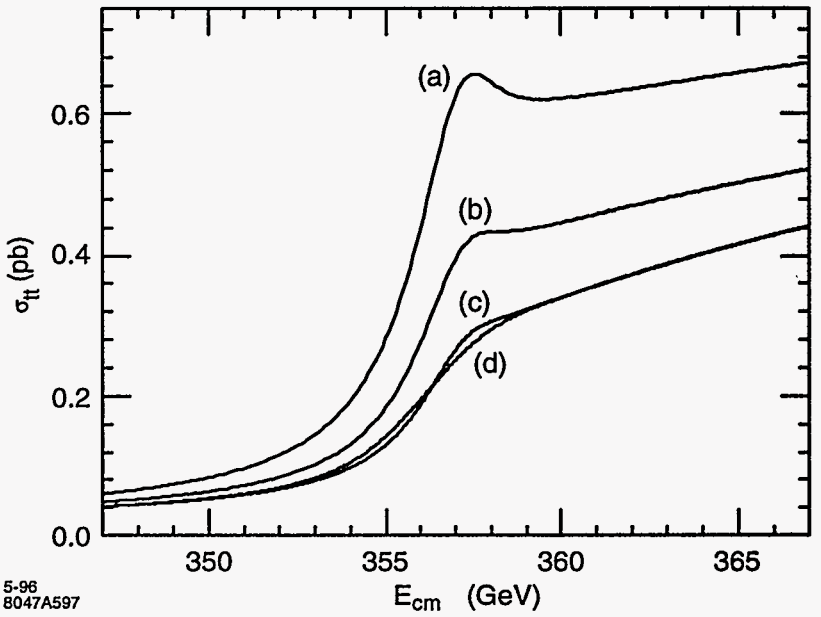


Figure 12-6. *Production cross section for top quark pairs near threshold for $m_t = 180\,\mathrm{GeV}/c^2$. The theoretical cross section is given by curve (a), to which the energy re-distribution effects have been applied. Curve (b): ISR; curve (c): ISR and beamstrahlung; curve (d): ISR, beamstrahlung, and beam energy spread.*

featureless threshold shape which is not nearly as sensitive to the $\Delta E_{\mathrm{cm}}/E_{\mathrm{cm}}$ distribution as would be expected if the top mass were smaller. Figure 12-8 shows the change in shape of the threshold cross section as the single-beam energy spread is increased.

12.3 Detector Background Sources

12.3.1 Beamstrahlung-produced e^+e^- Pairs

As the beamstrahlung photons travel in the high-field region of the opposing bunch, they can produce e^+e^- pairs. If deflected to large enough angles, the pairs can enter the detector and cause unwanted backgrounds. The probability for pair production is set by the Υ parameter. When $\Upsilon \gtrsim 0.3$ the pairs are coherently produced; that is, the virtual electron-positron pairs accompanying the beamstrahlung photon can exchange energy-momentum with the strong electromagnetic field and be kicked on-shell. When $\Upsilon \lesssim 0.3$ the incoherent pair creation processes become dominant. In these, the e^+e^- pairs are created from individual scattering of real beamstrahlung photons through the Breit-Wheeler ($\gamma\gamma \rightarrow e^+e^-$) and Bethe-Heitler ($e^{\pm}\gamma \rightarrow e^{\pm}e^+e^-$) processes, and from the scattering of virtual photons through the Landau-Lifshitz ($e^+e^- \rightarrow e^+e^-e^+e^-$) process.

This phenomenon is dealt with in two ways. Most importantly, as long as $\Upsilon \lesssim 0.3$ the number of coherently produced pairs per bunch crossing is negligible. This has been a design constraint of all colliders to date and seems reasonably easy to achieve at the start up energy of 500 GeV. Note the values of Υ listed in

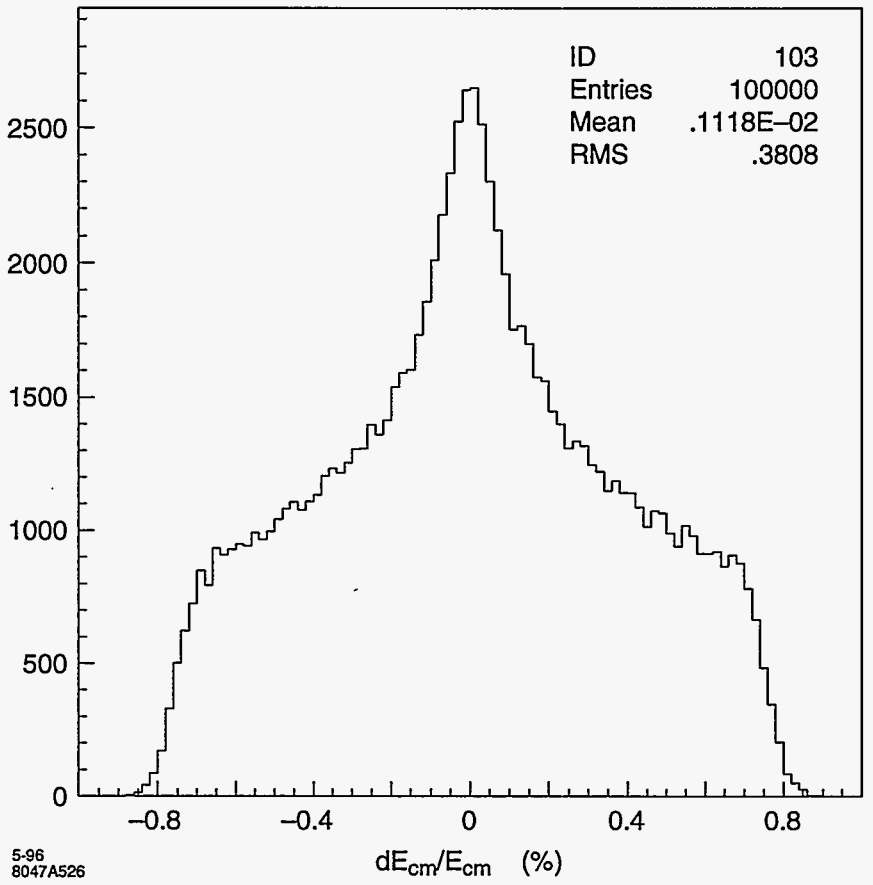


Figure 12-7. *Distribution of the center-of-mass energy, $\Delta E_{cm}/E_{cm}$, due to the convolution of single-beam energy spread distributions for the two beams. The single-beam energy spread in this case has a FWHM of 0.8%, corresponding to the distribution given in Figure 12-1.*

Table 12-3 for the various collider designs. Figure 12-9 plots the contributions from each of the sources of pair production per bunch crossing versus Υ. Table 12-4 lists the number of pairs produced for the different parameter lists at 500 GeV and 1 TeV.

The preceding chapter introduced the ABEL Monte Carlo program used to simulate the beam–beam interaction. The ABEL program divides each bunch longitudinally, produces beamstrahlung photons and e^+e^- pairs as the bunches overlap, and propagates the particles through the high fields of the bunch overlap region. Figure 12-9 and Table 12-4 are calculated using the parameterizations of ABEL described in Ref. [Chen, 1990]. In the last run of experiment E144 at the SLAC FFTB, values of $\Upsilon \sim 0.25$ were reached. Coherent pairs were observed at about the level predicted by the ABEL simulation. In its upcoming run, E144 should probe the region of $\Upsilon \sim 0.5$–0.6. This experimental confirmation of the background level anchors the estimates of detector hit densities arising as a result of pair creation which is discussed in later sections of this chapter.

Secondly, a masking scheme can be designed which, in conjunction with the detector's solenoidal field, protects the detector from the majority of the deflected pairs. The mask takes the form of a truncated cone which begins about 0.5 m from the interaction point and has an opening angle determined by the maximum

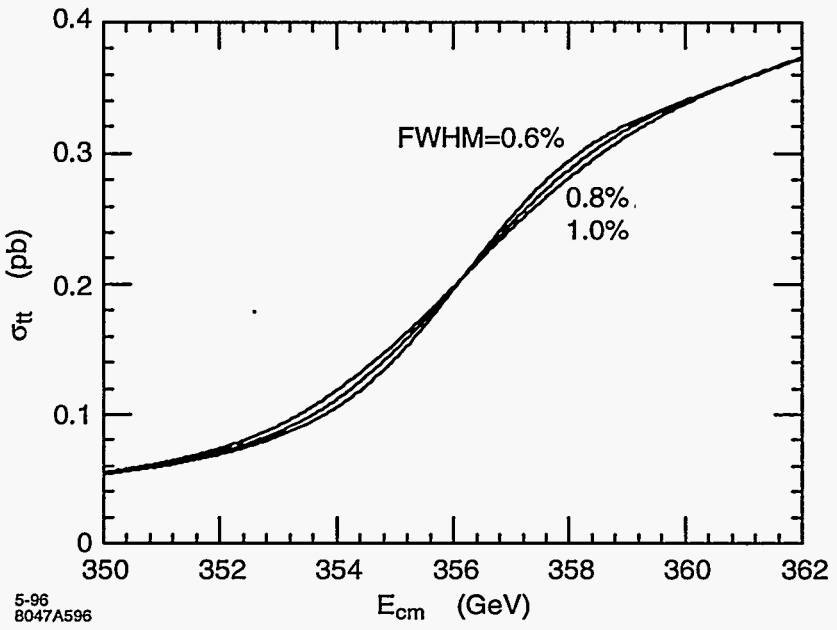


Figure 12-8. *Comparison of $t\bar{t}$ threshold shape, with all effects included, for different single-beam energy spreads. The three curves correspond to single-beam energy spread distributions with FWHM of* 0.6%, 0.8%, *and* 1.0%, *as indicated.*

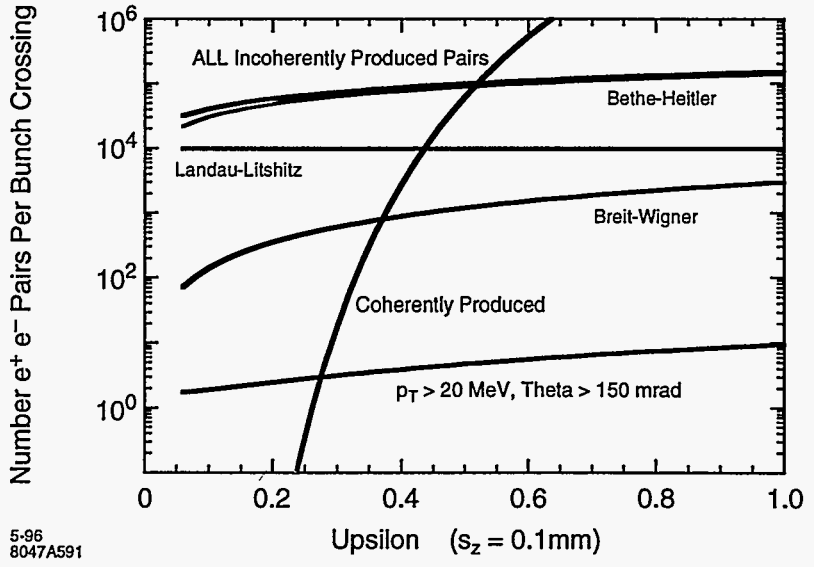


Figure 12-9. *Relative contribution of pairs from all sources as a function of* Υ.

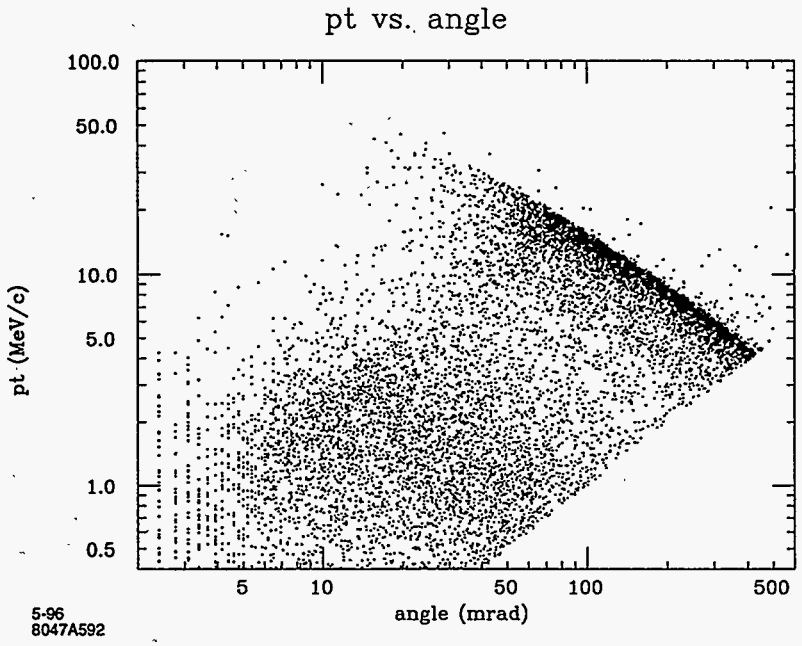


Figure 12-10. *p_t vs. θ distribution for pairs.*

kick the pair can receive from the field of the opposing beam. It defines a "deadcone" within which the detector is blind.

In coherent production, the pairs are not produced with significant intrinsic p_t; rather, roughly speaking, the e^+e^- are deflected by an angle proportional to $\sqrt{D_x/\epsilon} * \sigma_x/\sigma_z$, where D_x is the disruption parameter, ϵ is the ratio of the electron or positron momentum relative to the beam momentum, and σ_x and σ_z are the x and z beam spot sizes. The incoherent pairs have a more uniform energy distribution than the coherent process and can result in more particles with p_t and scattering angle θ outside the dead cone.

The pair partner with the same sign as that of the bunch that produced it will tend to oscillate within the field of the oppositely charged opposing beam. In the flat beam designs under consideration, there will be more oscillations in the vertical plane than the horizontal plane. The vertical oscillations will tend to cancel and the exit angle of the particle will be predominately in the horizontal plane. The oppositely charged member of the pair will see a defocusing force from the opposing beam and be quickly deflected out of the beam. From outside the beam, it sees the field of an elliptic cylinder charge distribution; up to terms of order $\log(D_x/\epsilon)$ the result for the maximum scattering angle is the same as that for the same-sign particle, although in this case the deflection angles in both the x and y planes are comparable.

Figure 12-10 shows the scatter plot of transverse momentum and scattering angle of the pairs generated by ABEL. The simulation used a 10-MeV cut on pair-member energy and a 2-mr cut on angle. The two bands seen in the plot corresponds to the opposite sign partners in the higher p_t region and the same sign partners in the lower p_t region.

	500 GeV			1 TeV		
	A	B	C	A	B	C
Detector dead-cone (mrad)	83	80	79	95	94	100
e^+e^- pairs per bunch X from coherent production	2.6E-13	1.4E-17	9.9E-19	76	13	9.1
e^+e^- pairs per bunch X from incoherent production	16.4E3	16.9E3	17.2E3	57E3	59E3	57E3
$N_{beamstr-e^\pm}$/bunch $\theta > 150$ mr, p_T >20 MeV	2.1	2.1	2.4	6.8	7.4	7.1

Table 12-4. *Summary of quantities parameterizing the backgrounds, before they interact in the detector.*

The particles travel in a helix from the IP. The solenoidal B-field strength and the distance of the conical mask from the IP determine the maximum particle momentum, and thus ϵ, that needs to be considered. All particles with lower momenta will curl up inside the cone. The maximum deflection angle can then be calculated from the formula alluded to above and the required radius of the conical mask determined. The dead cone is on the order of 100 mr and is tabulated for the various machine designs in Table 12-4. The thickness of the mask must be determined by detailed EGS or GEANT simulations to be adequate to stop the debris of the interacting electrons and positrons. Typically, the angle defining the outer dimension of the mask is 150–200 mr. Table 12-4 tabulates, for each of the parameter sets, the number of particles with p_t and θ large enough so that they fall outside the mask and hit the detector. Dealing with these particles depends on the time or bunch structure of the machine and on the timing capabilities of the detector.

12.3.2 Hadronic Backgrounds from $\gamma\gamma$ Interactions

In addition to the electromagnetic processes discussed above, the beamstrahlung photons can interact to produce hadrons and jets of hadrons. By folding the beamstrahlung spectrum into cross section estimates using an equivalent photon approximation, these hadronic rates can be estimated.

Hadronic events produced by beamstrahlung $\gamma\gamma$ annihilation are expected to be benign. Most of the hadronic events are minimum-bias events with small transverse momentum and small center-of-mass energy. Furthermore, the two photons involved in the collision usually have very different energies so that the hadronic system is highly boosted along the beam direction. Monte Carlo studies have demonstrated that the mean energy deposited in a detector from an hadronic background event will be 8 and 11 GeV for colliders with $\sqrt{s}$=500 and 1000 GeV, respectively. Here it is assumed that the detector has typical electromagnetic and hadronic calorimetry over all solid angle with the exception of a hole (dead cone) with $|\cos\theta| > 0.985$. The energy deposition with a larger solid angle dead cone of $|\cos\theta| > 0.900$ will be 3.3 GeV and 4.4 GeV for $\sqrt{s}$=500 and 1000 GeV, respectively.

12.3.3 Quadrupole and Bend Synchrotron Radiation

Linear collider designs obtain a small beam spot at the IP by using strong focusing magnetic quadrupole lenses close to the IP. SR generated by particles passing through these quadrupoles and bend magnets in the final focus is a potential source of background in the detector. SR backgrounds in the SLD detector at the SLAC Linear Collider (SLC) are in reasonable agreement with calculations. The SLD/SLC model implies that the nominal Gaussian beam core generates a small SR background in the final "soft" bend magnet, and negligible SR background in the quadrupoles. However, the non-Gaussian beam tail or beam halo can result

in significant background due to synchrotron radiation generated in the quadrupoles (QSR). This source is controlled by collimation of the beam and by limiting the angular divergence at the IP. The QSR background is quite variable, and measures that control it can reduce luminosity.

Apertures in the beam line near the IP are matched to the design angular divergence and collimation. The SLD has a carefully designed set of internal masks to limit the SR background in the detector, especially the drift chamber. The SR background at SLD is reasonably well modeled by a nominal Gaussian core and a beam tail at the IP that is a very broad Gaussian containing 1% of the nominal charge. This "1% flat tail" is purely arbitrary and difficult to measure directly. The SR background generated by the tail depends directly on the fraction of the beam it contains and how it is collimated.

SR backgrounds in the SLD were modeled with EGS4, starting with SR photon fluxes calculated with a modified version of QSRAD. QSRAD was written to study SR generated by the PEP-I beam in the last two quadrupoles prior to the IP. The program traces weighted rays from a Gaussian beam profile through the specified magnetic optics and produces a geometric fan of synchrotron radiation with uniform power density and constant critical energy for each magnetic element. These fans are then traced, and a tally is made of the fraction of each fan that strikes opaque surfaces with specified apertures. The distribution of photon critical energies is accumulated for each surface and converted to a photon energy distribution. The code provides additional information to characterize the SR photons incident on each surface, and this is the input data for an EGS4 model of the masking and detector.

At SLD the SR due to the final soft bend has a critical energy of 70 keV and results in 7×10^8 photons per pulse incident on a mask 1.3-m down beam from the IP. This is $\approx 1.5 \times 10^{10}$ keV per pulse at 120 Hz. EGS4 calculations predict that this would result in 280 photons entering the SLD drift chamber per pulse. This corresponds to about 0.5% wire occupancy. The major component of SLD background from quadrupoles in the model is due to photons striking the same mask. There are 1.5×10^8 photons per pulse with a critical energy of 1 MeV. This is about 5×10^{10} keV per pulse and, according to EGS4, results in 73 photons in the drift chamber. However, this background, due to the tail, is very sensitive to the model, and can be made considerably larger by varying the assumptions about the IP divergence angle and the beam collimation, thereby introducing new sources. We believe that variations in these parameters contribute to the fluctions in backgrounds actually observed. The energy spectra for the NLC calculation contain far more SR energy than calculated for the SLD case. The spectra in Figure 12-11 correspond to a critical energy of 200 keV and contain between 1.2×10^{11} keV and 1.5×10^{11} keV per train of 90 bunches.

Preliminary calculations of NLC synchrotron radiation have been performed for the 1-TeV machine with 500-GeV beams using the currently available final focus optics [Helm, TLCFF28]. The QSRAD code has been enhanced to include bend magnets in addition to quadrupoles, and to allow SR from individual magnets to be turned on or off. Optical elements within 300 m of the IP have been included in the calculation. SR from more distant magnets cannot reach the IP region directly because of the bends centered at 183 and 261 m from the IP. The beam at the IP was taken to be 226 nm × 3.57 nm with angular divergence 22.6 μr × 28.6 μr, corresponding to the 1-TeV parameter set A of Table 12-1. Calculations were done for a single bunch containing 9.5×10^9 electrons, and, unless otherwise indicated, all results quoted here are for a single bunch. Collimation was modeled as perfect rectangular collimators at $7\sigma_x$ and $35\sigma_y$. SR from beam tails was calculated for an arbitrary 1% flat tail, 9.5×10^7 electrons per bunch uniformly distributed over the collimator aperture, and the reader can scale this up or down as desired. In reality, detector elements will integrate backgrounds over a bunch train, and the single-bunch numbers should be multiplied by 90 bunches per train for background considerations.

The first studies of SR in the NLC final focus lead to the relocation of the final bends and other modifications. In the current design the final bend ends 151 m from the IP, apertures in the final focus have been increased,

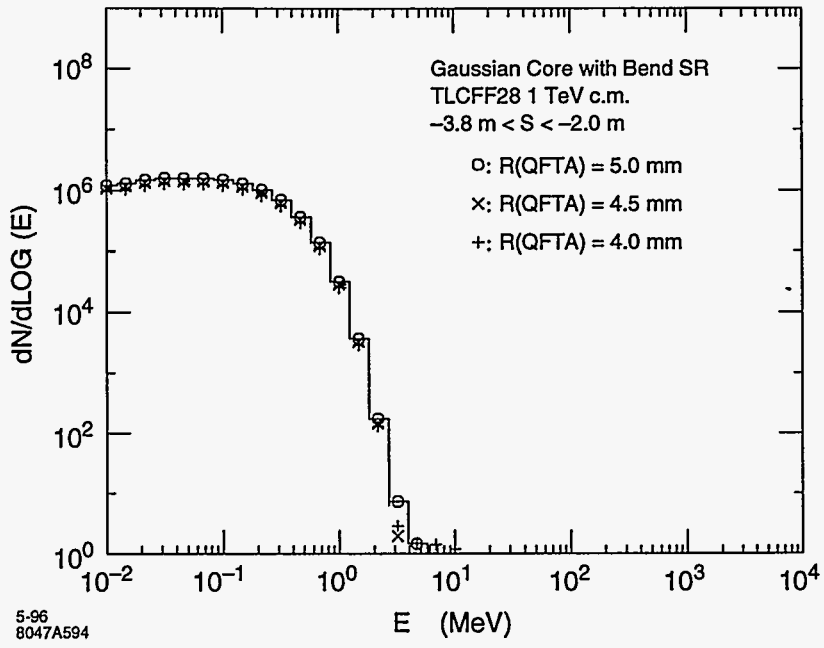


Figure 12-11. *Energy distribution of synchrotron radiation photons, from particles in the Gaussian core of the beam, striking the inner aperture of the QFTA quadrupole. There are only minor variations in the range of radii considered.*

and the quadrupole closest to the IP (QFTA), with the smallest aperture, is shorter than in the previous lattice.

Our experience with EGS4 simulations of SLD backgrounds and initial calculations of SR at the NLC indicate that the most serious source of SR background in this model of the NLC is likely to be the SR photons incident on the inner surface of the beam pipe through QFTA, the magnet closest to the IP. The calculations reported here are for this source only. We consider here the inside of a cylindrical surface of length 1.8 m, between 2.0 and 3.8 m from the IP. This includes the interior of QFTA and a mask that is likely to precede the magnet. The calculation is done for several quad apertures, as discussed below.

The Gaussian beam results in 1.2×10^7 photons above 10 keV energy incident on the inside of a cylindrical QFTA of 4.5-mm radius. These photons are all from the final "soft" bend (SBSR), and have a critical energy of 200 keV. The number of incident photons changes by only 10% if the radius of the magnet aperture is changed by 0.5 mm. The energy spectra are shown in Figure 12-11. These numbers are independent of assumptions about the beam, other than parameters that are well defined. Furthermore, the interior of QFTA can be shielded from the SBSR by a mask in the horizontal plane 40 m from the IP. The half-aperture of this mask would be 3 mm, about $10\sigma_x$, whereas the beam is collimated at $7\sigma_x$. Although the mask is required on only one side of the beam, it would be symmetric to minimize wakefield effects, which must still be evaluated before this mask is accepted as part of the design.

The effect of the tail is calculated without the SR from the bends, because that has been included with the Gaussian core. Results for the flat tail are very dependent on assumptions about collimation and the population of the beam tail. With the assumptions above, the 1% tail calculation results in 3.7×10^5 photons per bunch with an average energy of 5.3 MeV incident inside QFTA if the aperture is 4.5 mm in radius. This SR is all from QFT5 and QFT6, 134 m and 150 m from the IP.

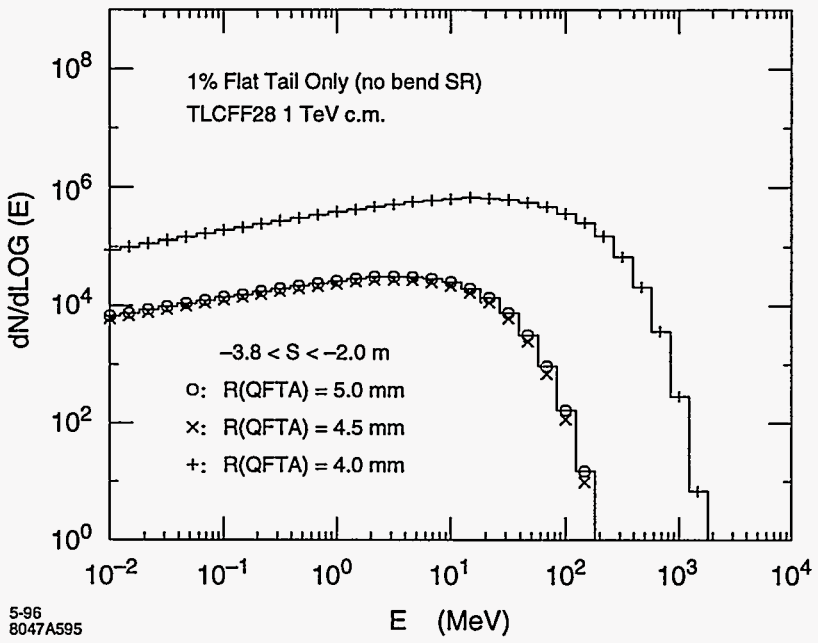


Figure 12-12. *Energy distribution of synchrotron radiation photons, from particles in the assumed 1% flat tail of the beam, striking the inner aperture of QFTA, the innermost quad. Photons produced by bends and photons produced by the Gaussian beam core are not included in this plot. The spectra for 5.0-mm and 4.5-mm radius do not differ significantly. When the aperture is small enough to intercept QSR from the superconducting quadrupoles the incident QSR is much larger, as shown for a 4.0-mm radius.*

There is also a large flux of 5.1×10^7 SR photons per bunch, incident where the aperture decreases between QFT1 and QFTA. Calculation shows that this can be reduced by an order of magnitude, shifting the flux upstream to a protection collimator 10 m from the IP. This SR is primarily in the vertical plane, and 90% can be intercepted by a 5-mm mask, which is large compared to the 2-mm $35\sigma_y$ beam envelope at this location.

If the radius of QFTA is increased to 5.0 mm, there is a small increase in the number and average energy of the incident photons. The energy spectra are shown in Figure 12-12. However, Figure 12-12 also shows that reduction of the radius to 4.0 mm results in a large increase in both the number and energy of the incident photons. The total SR energy incident on the inner bore of QFTA increases by more than two orders of magnitude. This is because the smaller aperture intercepts a large flux of energetic QSR generated by the beam tail in the superconducting quadrupoles QFT1 and QFT2. This must be considered as the design evolves, because the only way to compensate for a smaller QFTA aperture is with tighter vertical collimation. In this lattice the collumator aperture must be reduced by one σ_y to compensate for each 125-μm reduction in QFTA radius.

The exit aperture to the dump line across the IP will have to be larger than QFTA in order to avoid a large flux of QSR incident on the face of the first dump line quad. The non-zero beam crossing angle makes this possible. For a 4.5-mm radius QFTA the first dump line quad aperture should have a 7-mm radius. Figure 12-13 shows the $7\sigma_x$ and $35\sigma_y$ beam envelopes. In the figure the optical elements are subdivided and the SR fan produced by each element traced to the IP. It is clear that the SR fan in the vertical plane from the final doublet is what sets the exit aperture. The radiation that does hit QFTA is seen to come from Q6, approximately 150-m upstream of the IP.

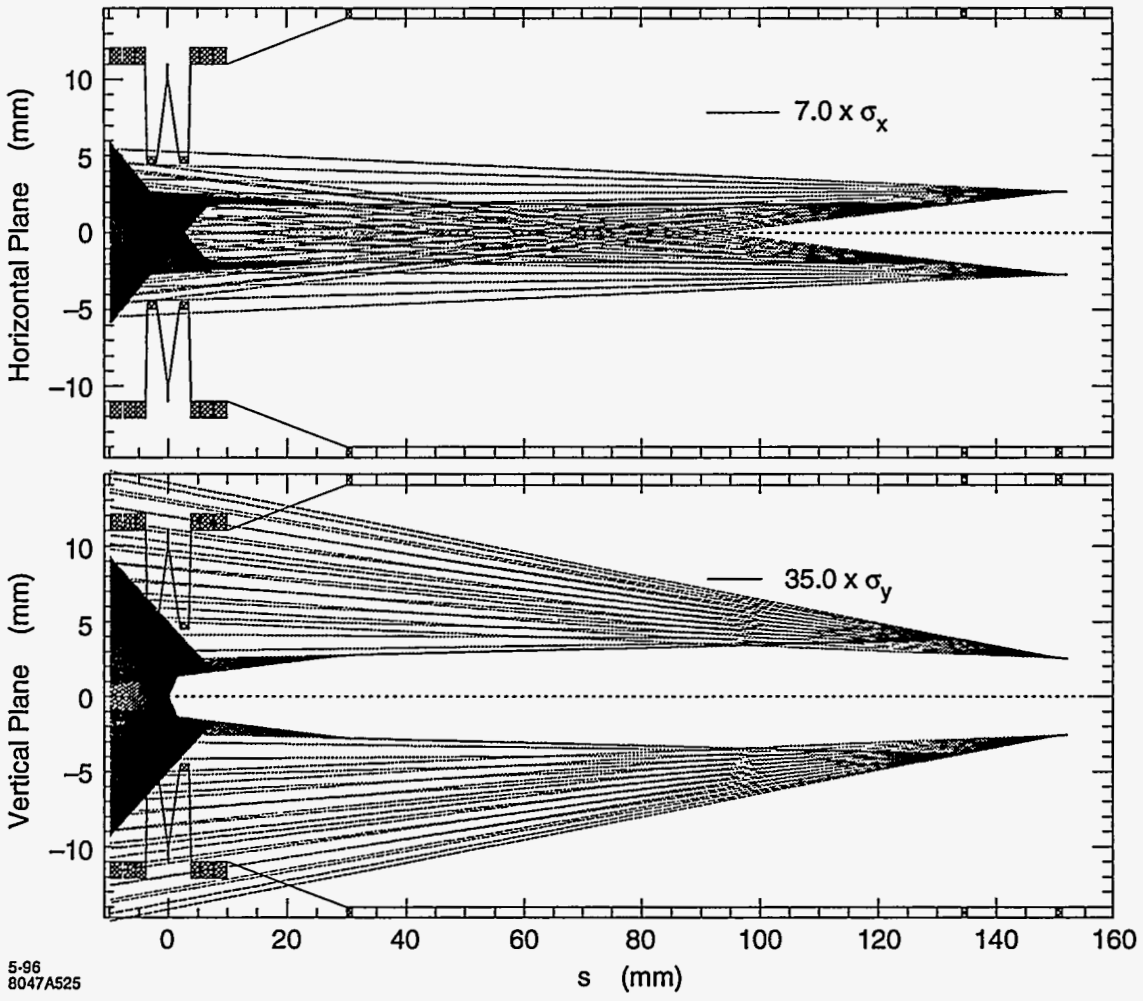


Figure 12-13. *The $7\sigma_x$ and $35\sigma_y$ beam envelopes and the SR fans emitted for the lattice used for these calculations.*

These studies should continue in conjunction with design of the final focus and the dump line. More detail must be included in the models of the apertures to identify those sources which are most likely to result in detector backgrounds. It is not obvious that the 1% flat tail assumption, which seems to work at the SLC, is reasonable for the the NLC. Because of its importance in estimating detector backgrounds, an effort should be made to estimate the likely beam halo at NLC.

The results of these SR calculations are used as input to EGS4 and GEANT models of a detector and its masking, and the first results are described later in this chapter. This study is far from complete, and detailed results from the EGS calculation will be used to propose further modification of the lattice and magnet apertures. The severity of the backgrounds modeled here may influence the choice of detector technology. The machine design has been able to incorporate suggestions motivated by the background calculations, and considerable progress has been made in reducing the potential for backgrounds due to synchrotron radiation.

12.3.4 Muon Backgrounds

The same collimators which limit beam phase space and protect the detector from SR backgrounds produce considerable numbers of muons as the incoming beam interacts with them, primarily through the Bethe-Heitler process $e^{\pm}N \rightarrow e^{\pm}\mu^{+}\mu^{-}N$. The problem was first encountered in 1988 with the Mark-II detector at the SLC. To quantitatively study the production and transport of the muons the program MUCARLO [Feldman] was written. It successfully reproduced the experimental results for the number and spatial distributions of muons hitting Mark-II. The backgrounds were reduced to acceptable levels by moving the primary collimation to a point at the end of the linac, approximately 1500-m from the IP and separated from the detector by the SLC arcs, and installing magnetized iron spoilers in the final focus tunnels. At the NLC, this philosophy is duplicated by locating the collimation section at the end of the linac, $\sim$ 5.2-km from the interaction point, and placing the final focus in a tunnel section dug at 10 mr with respect to the linac, the "Big Bend."

In 1990, the MUCARLO program was modified [Keller, 1991] for use with beams of up to 250-GeV energy. Using the TRANSPORT deck for a 500-GeV-c.m. collider available in June 1990 (FFN09, R. Helm), the placement of five toroidal spoiler magnets was optimized to maximize the number of electrons that could hit a collimator before producing one muon that would make it to the detector.

In 1993, MUCARLO was again modified. Beams of energy up to 500 GeV were allowed, and muon production by direct e^{+} annihilation, $e^{+}e^{-} \rightarrow \mu^{+}\mu^{-}$, and photopion production, $\gamma A \rightarrow X\pi(\pi \rightarrow \mu\nu)$, were included. Using the TRANSPORT deck for a 1-TeV-c.m. collider available in June 1992 (TLCFFN5, R. Helm), the muon background study was repeated [Keller]. Most of the figures presented in this section are from that study. Figure 12-14 shows a schematic of the collimation and final focus beam transport sections being considered at that time. The collimation section has a series of six Hi-Z collimators and a total bend of 6.14 mr followed by a big bend with a total bend of 10 mr followed by the final focus with a reverse bend chromatic correction section of $\pm$2.21-mr bends. The model includes a 3.05-m-square cross section concrete tunnel through sandstone, concrete support girders under the beam elements, and dipoles and quadrupoles which include return flux in the iron and pole tips. The detector is assumed to have a 4.5-m-radius cross section centered on the IP. There is a series of magnetized iron spoilers of alternating polarity distributed through the final focus.

Figure 12-15 shows how the spoilers are arranged in the tunnel at a given location. The field in the iron was modeled using the two-dimensional program POISSON.[7] Each spoiler is 9.1-m long with a winding slot width and height of 3.2 cm and 126 cm, respectively. Each set of "tunnel-filler" spoilers weighs 750 tons and would cost \$2–3 million installed. For magnetized iron with a field of 16 kg, the ratio of bend angle to scattering angle is $\theta_{\text{bend}}/\theta_{\text{MCS}} \approx 3\sqrt{L}$, where L is the length of the spoiler in meters.

An alternative to large tunnel-filler magnetized iron spoilers has been proposed, consisting of nested iron cylinders with opposite-polarity azimuthal magnetic fields as shown in Figure 12-16. The idea is that the nested cylinders are located downstream from each muon source and are long enough to either range out muons or cause enough energy loss so that the muon is unlikely to reach the detector. A version of the nested cylinder idea was tried in MUCARLO, and the results are presented in Figure 12-17.

After the muon exits the source, the Monte Carlo program swims it in 30-cm steps through the tunnel. When material is encountered, the muon scatters, loses energy, and bends (if a magnetic field is present). The trajectory of each muon is followed until the muon either stops or reaches the IP. For the purposes of this study, the details of the detector, *e.g.*, trackers, calorimetors, and muon walls, are not included. A muon which reaches the IP within a radius of 4.5 m from the beam line is counted as a detector hit.

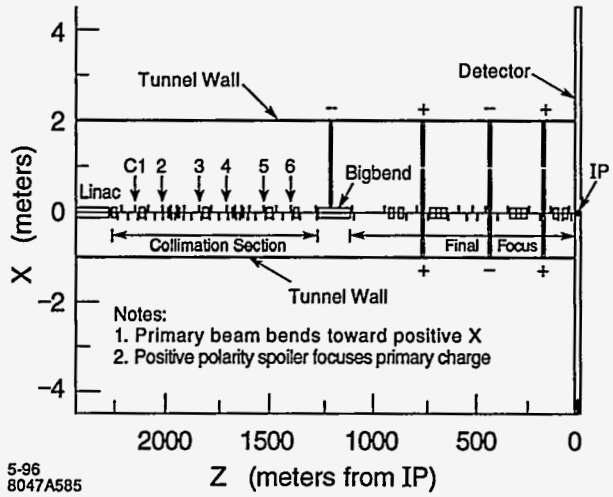


Figure 12-14. *Plan view of the 1992 NLC beam line in the tunnel with magnetized iron spoilers.*

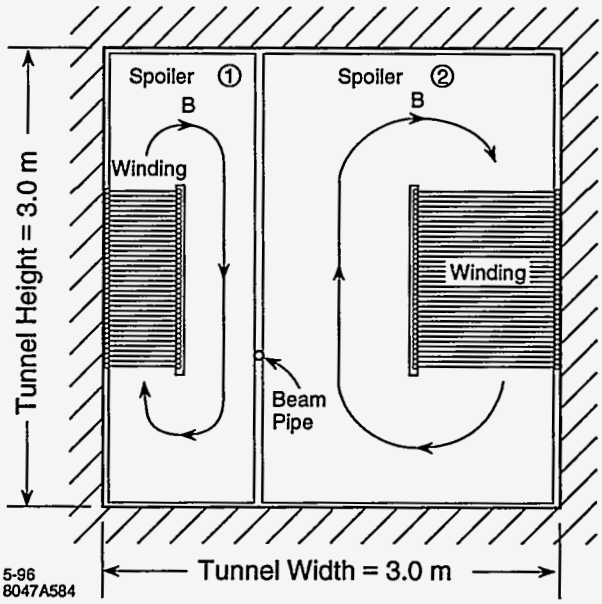


Figure 12-15. *Two side-by-side magnetized iron spoilers filling a 3 × 3-m beam tunnel.*

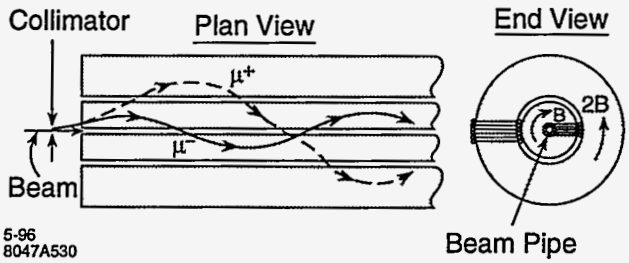


Figure 12-16. *Nested, magnetized iron cylinders of opposite polarity to channel both μ^+ and μ^-.*

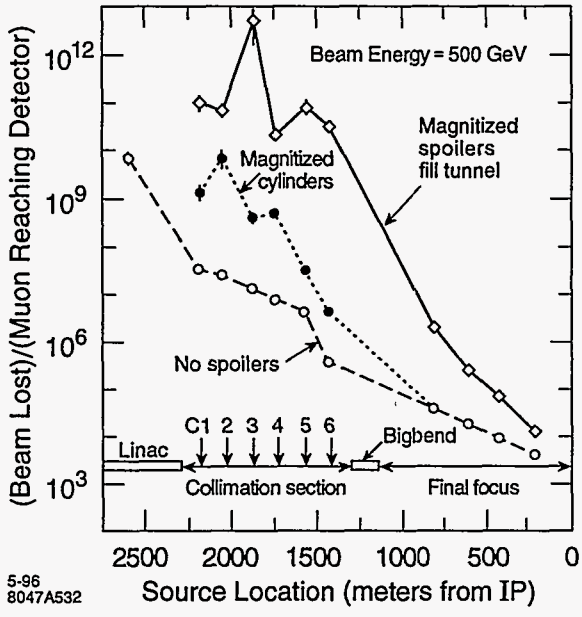


Figure 12-17. *Results of muon Monte Carlo calculations, based on the 2-km final-focus design of 1992, for three conditions: no spoilers, and two types of magnetized iron spoilers.*

Figure 12-17 shows the number of beam particles which must hit a collimator to produce one muon in the detector as a function of source location in the beam line. As a worst case, an e^+ beam was chosen for this study so that direct annihilation production could be included. The source points include the six Hi-Z collimators in the collimation section and collimators at four high-beta points in the final focus, which are potential scrapers of beam-gas coulomb scattering or beam-gas bremsstrahlung.

From the curve labeled "No spoilers" it is seen that a beam loss of 10^{10} per bunch train anywhere in the collimation section will result in a muon background which is many orders of magnitude away from the design goal. The curve labeled "Magnetized spoilers fill tunnel" in Figure 12-17 is the result of adding magnetized iron spoilers which fill the tunnel at three locations in the final focus and one magnetized iron piece in the tunnel aisle next to the 10 mr big bend. It is seen that for all six collimators in the collimation section, more than 10^{10} beam particles must be lost to produce one muon in the detector. This satisfies the design goal of allowing a 1% continuous beam loss in the collimation section.

The curve labeled "Magnetized cylinders" in Figure 12-17 shows the results of filling all drift spaces in the collimation section with magnetized iron cylinders described above. The result is considerably worse than for magnetized iron spoilers which fill the tunnel and does not meet the design goal. This is because the magnetized cylinders must be interrupted by beam elements, especially dipoles in the chromatic correction sections, which disperse muons away from the cylinders and therefore disrupt the channeling orbits.

For the case of the magnetized iron tunnel-fillers, Figure 12-18 shows histograms of muon-production momenta and final momenta for muons which hit the detector from the six sources in the collimation section. It is seen that the initial muon momentum must be greater than 300 GeV/c to hit the detector from sources outboard of the big bend. Figure 12-19 shows the spatial distribution of muons from the collimation section which reach the IP. It is seen that increasing the detector dimensions by 2–3 m in the horizontal direction would intercept significantly more muons.

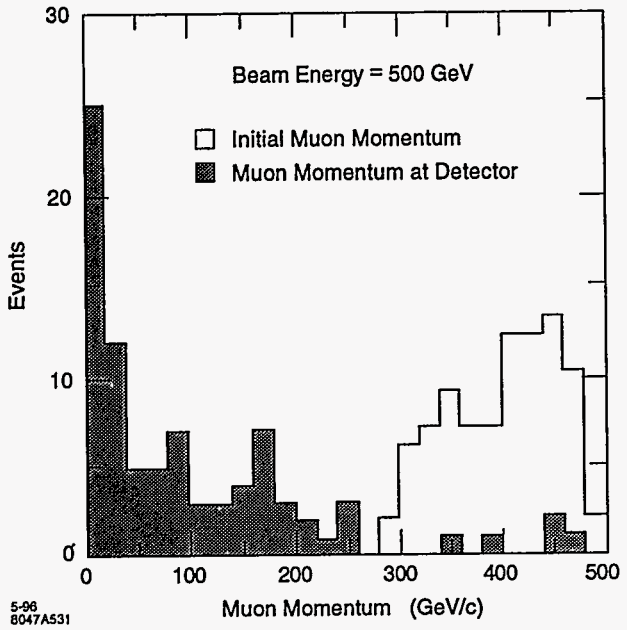


Figure 12-18. *Momentum distribution of muons which hit the detector from six sources in the collimation section. The normalization is arbitrary.*

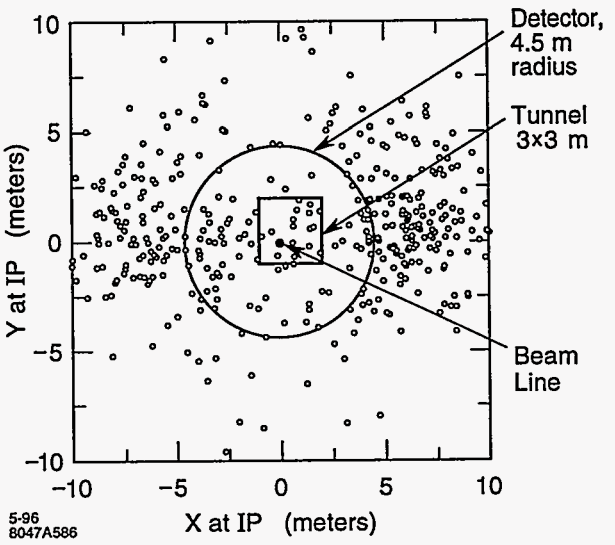


Figure 12-19. *Spatial distribution of muons which reach the IP from six sources in the collimation section. The normalization is arbitrary.*

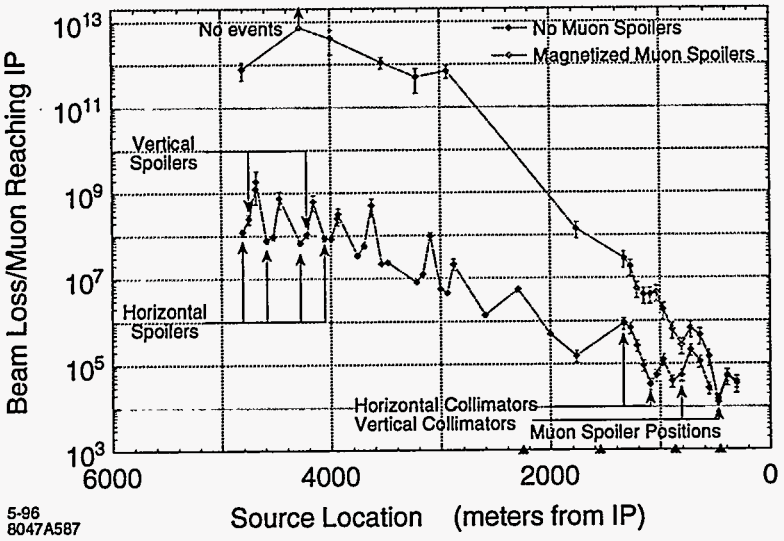


Figure 12-20. *Results of muon Monte Carlo calculations updated for the collimation and final focus design available in early 1996. Data are presented for calculations with and without a muon spoiler system. The data from the previous calculation (Figure 12-17) are also plotted for reference.*

For sources in the collimation section of a 1-TeV center-of-mass linear collider, a system of magnetized iron spoilers which fills the tunnel and meets the design goal of allowing a continuous 1% beam loss, or 10^{10} beam particles per bunch train, was found. This is more than a three orders of magnitude improvement over the case with no muon spoilers. For the case of magnetized iron cylinders which fill all drift spaces in the collimation section, the design goal was not met.

These studies were recently repeated using the current design for the collimation section, big bend, and final focus (TLCBD01B, R Helm). While the tunnel size, magnet design, and magnet support are important inputs to the program, we still assume the original 3×3-m-square tunnel cross section and SLC-like magnet and support. Figure 12-20 shows the results for two configurations: no spoilers and with tunnel filling magnetized spoilers. It can be seen that the newer design, with its longer beamline, allows for a larger number of beam particles to be lost per muon arriving in the experimental hall. Nonetheless, muon spoilers are still required to attain the design goal. The spoiler locations used here were determined by scaling those of the older, shorter final focus design. Additional optimization of the spoiler system is in progress.

12.4 Detector Issues

There has been much less thought devoted to the NLC detector than to the accelerator itself. In physics studies conducted in Japan and Europe, the detector is envisioned to follow the solenoidal architecture typified by OPAL, ALEPH, or SLD, scaled appropriately for the increase in center-of-mass energy and having, depending on the study, improved tracking, vertexing, or calorimetry. Table 12-5 gives a list of possible physics-motivated specifications as summarized in a recent review.[TRC 1995]

Two issues appear to drive the conceptual design of the detector: the required momentum resolution of the charged particle tracking system and the requirements for "special" support systems for controlling vibrations of the final focus. A pixel vertex detector (presumably CCDs), a moderate resolution electromagnetic and hadronic calorimeter, and a muon identifier of instrumented steel plates are non-controversial.

Recently, a "straw-man detector" design has been presented [SLAC 1996] which is based on a compact five-layer tracking system using double-sided 5-μm-resolution silicon strips in 4T magnetic field. Such a system would have a momentum resolution around 0.03%×P(GeV), and would be resistant to the expected backgrounds and instabilities of the NLC. By keeping the coil inside the electromagnetic calorimeter it remains relatively thin (2–3 radiation lengths) and short (2–3-m in length). Figure 12-21 shows a quadrant this detector.

The nominal nanometer vertical size of the NLC focus drives the question of the quadrupole support system. The design possibilities include a rolling support of the quadrupole from a retractable end door such as in the SLD; a static steel end door supporting the quadrupole with retractable muon system and calorimeter for interior access; and actual intrusion of a static steel and concrete buttress into the door region for quadrupole support. Measurements of the differential motion of the quadrupole end of SLD have been made (see Section 12.4.5). It appears that, with appropriate care, seismic concerns will not overly compromise the detector design.

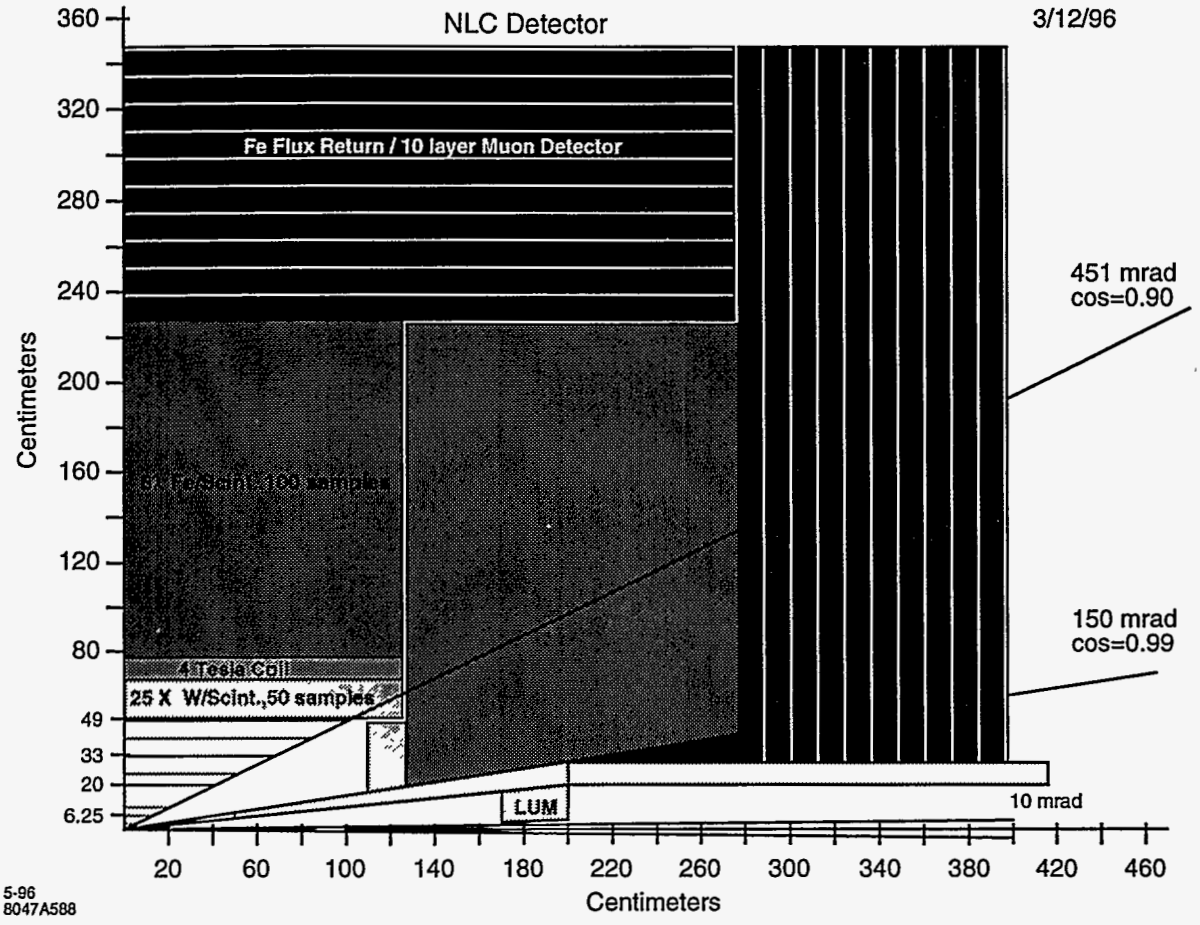


Figure 12-21. *A compact NLC Detector, robust against backgrounds. The scale is in cm.*

12.4.1 Effect of Backgrounds on the Detector

The zeroth-order masking design for the IP is shown in Figure 12-22. This geometry has been programmed into EGS4, and full simulations of the effects of the interactions of beamstrahlung-produced pairs, QSR-produced photons, and lost beam particles are in progress. Tungsten is used for Mask M1, M2 and the luminosity monitor. The beam pipe made of 750-μm-thick beryllium is located at $r = 1$ cm near the IP and of 500-μm-thick stainless steel for the rest. The beam pipe is pulled back to a larger radius as soon as it is past the maximum z required by the vertex detector. The rf shield made of 200-μm-thick copper is necessary to avoid the wakefield effect. The solenoid field in the detector and the quadrupole field in the final quads are included, and charged particles are transported properly in the magnetic field. However, non-axial field components of the solenoid or fringing field of the quads are not considered. The vertex detector is modeled with four layers of sillicon+beryllium at the radii of 1.2, 2.7, 3.8, and 4.8 cm with an angular acceptance of $|\cos\theta| < 0.9$. The layer thickness and the locations of the outer three layers are chosen after the new SLD vertex detector. In addition to the vertex detector layers, there are massless scoring planes at $r = 10$, 15, 20, 25, and 30 cm with $|\cos\theta| < 0.9$. The tracking chamber is located at $r = 30$ cm and $|z| < 100$ cm.

	LEP/SLC-Style	ee500 1991	ee500 Typical	JLC Detector	1 TeV Detector	Units
Tracking $\frac{\delta p_t}{p_t^2} = C$						
$C =$	8	5-100	10	1	2	$\times 10^{-4}$gev/c^{-1}
E-M Calorimeter						
$\frac{\delta E}{\sqrt{E}} =$	0.2	0.02-0.15	0.1	0.15	0.10	$\sqrt{\text{GeV}}$
Hadronic Calorimeter						
$\frac{\delta E}{\sqrt{E}} =$	0.9	0.3-1.0	0.8	0.40	0.65	$\sqrt{\text{GeV}}$
Energy Flow						
$\frac{\delta E}{\sqrt{E}} =$	0.65	0.3-0.8	0.5	0.3	0.4	$\sqrt{\text{GeV}}$
Vertexing						
$\delta(IP) = A \bigoplus \frac{B}{p}$						
$A =$	25	5-20	10	11	10	μm
$B =$	100	50-100	50	28	50	μm GeV/c
Hermetic coverage						
$\lvert\cos\theta\rvert <$	0.96	0.70-0.99	0.95	0.98	0.98	

Table 12-5. *Examples of detector performances used in physics studies.*

The complete history of secondary electrons, positrons, and photons produced from the incident e^+e^- pairs and synchrotron photons was traced using EGS, and the background level in the detector was calculated.

The figure of merit we have traditionally used for the maximum allowable background level is one charged particle hit/mm^2/train-crossing for a pixel-based vertex detector, and 10,000 photons/train-crossing for a drift chamber.

This work has resulted in changes to the final-focus lattice and to the design of the final quadrupole doublet. The process has been iterative and is still in progress; the reader is warned not to assume that we have presented a unified and consistent set of results throughout the relevant sections of the design report.

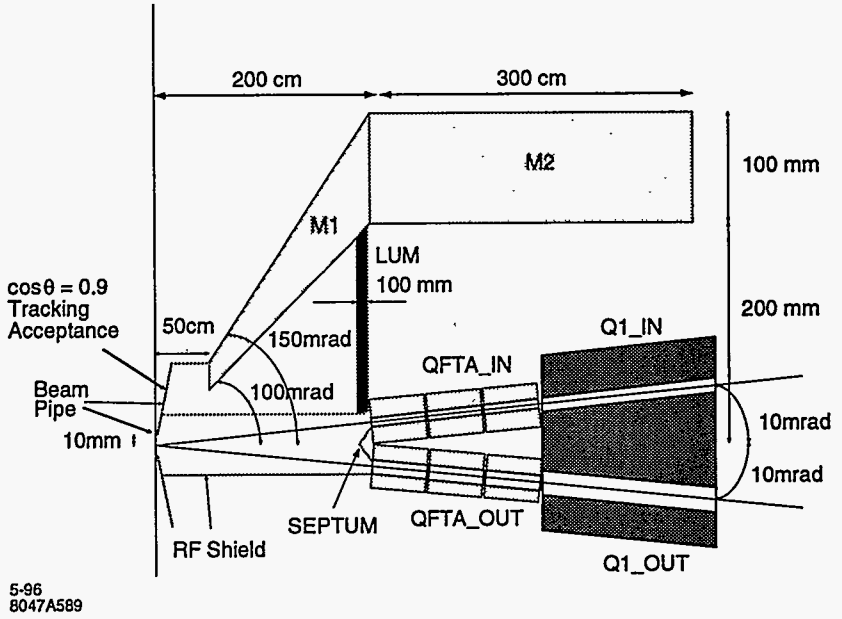


Figure 12-22. *First pass at NLC masking:*

M1 Mask	Tapered Tungsten Cone beginning at z=0.5 m and ending at z=2.0 m, with inner and outer angles of 100 and 150 mr, respectively.
M2 Mask	Tungsten annulus with 10-cm wall and inner radius 20 cm, $2.0 < z < 5.0$ m.
QFTA_IN/OUT	Incoming/Outgoing Sm2Co17 FF quads rotated 10 mr in x–z plane. QFTA_IN - Inner/Outer radii = 4.5/20.0 mm QFTA_OUT - Inner/Outer radii = 7.5/20.0 mm Longitudinally divided into three 5.0-cm-long segments with 2.0-cm gaps. L^* = line at 10 mr to IP is 2.0-m long.
Q1_IN/OUT	Superconducting magnet with Q1_IN aperture = 5 mm; Q1_OUT aperture = 8 mm. Extends from 3.5 to 5.0 m.
Beam Pipe	750-μm Be beam pipe with 100-μm Titanium liner at 1.0 cm, at $z \pm 2.1$ cm, which is joined onto a 500-μm Stainless sectioned flared at 451 mr until r = 7.56 cm, after which it proceeds to the M1 mask and follows its inner contour.
RF Shield	200-μm Cu extension at r = 2.75 cm, z = 2.1 cm, extending at constant radius to z = 165 cm, then proceeding via two "legs" to join electroplated inner radius of in-/out-going quads.
SEPTUM	200-μm Cu cone beginning at z = 189 cm and proceeding to z = 2.0 m at an angle θ, where $\theta = \arctan((20\,\text{mm} - 7.5\,\text{mm})/11\,\text{cm})$.
LUM	10-cm Tungsten with back end at 195 cm, outer radius tapered at 150 mr at M1, inner radius tapered at 10 mr at radius of outgoing quad.

Beamstrahlung-produced pairs

The ABEL simulations of the beam-beam interaction have been described in the final-focus chapter (Chapter 11) of this report. We have concentrated on the data set which describes the 1-TeV NLC parameter set A machine design listed in Table 12-1. The simulations used a 10-MeV cut on pair-member energy. This cut is responsible for the hard edge at the lower right corner of Figure 12-10. As the bremsstrahlung probability for high-Z materials is not negligible at 10 MeV, the cutoff energy should be lowered to 1 MeV in ABEL. The effect of lowering this cut is under study.

Particles with a large p_t will hit the vertex detector directly, while particles with a small p_t will traverse the detector following the axial magnetic field and hit the final quad face producing secondary charged particles which come back to the IP and hit the vertex detector. Figure 12-23 shows the average hit densities of e^+e^- per train at $r = 1$ and 2 cm as a function of z for two solenoid field strengths. The solid histogram represents the particles directly hitting the layer, while the dashed histogram is for the backscattered particles. Since the interaction of e^+e^- at a few MeV is very dependent on the detailed geometry and materials, the calculation presented in Figure 12-23 used massless scoring planes at r=1 and 2 cm. Since no interactions were simulated at these layers, one particle could contribute multiple hits. Therefore, the numbers should be considerd as an upper-limit. With this caveat, the hit density of directly hitting particles is about 2–3/mm^2/train at r = 1 cm and less than 1/mm^2/train at $r = 2$ cm; an acceptable level for a pixel-based vertex detector. The hit density of backscattered particles is, however, about ten times higher. Low Z coatings on the quadrupole and luminosity monitor faces will substantially reduce the flux of soft backscattered particles. Improvements to the IR layout and masking to reduce the backscattering probability are underway.

This study has tried to find the minimum feasible radius for a vertex detector. The reader must remember that at $r = 3$ cm, the direct pair background is essentially zero in a 4-T field. Furthermore, at 500 GeV-center-of-mass, backgrounds are down by another factor of 4.

Figure 12-24 shows the number of photons per train striking a scoring plane at $r = 30$ cm, the nominal position of the inner wall of a central tracking chamber as a function of the solenoidal field of the detector. The number of electrons crossing the scoring plane is negligible. Since the majority of photons are produced by the interactions of pairs at the Mask M1 face and at the rf shield, the background can be reduced substantially by using higher solenidal field strength so that more pairs are curled up within the deadcone. However, even at 4 T, the total number of photons is 80K/train, which may not be tolerable for a drift chamber.

Figure 12-25 shows the number of photons per mm^2 per train striking the vertex detector layers and scoring planes from $r = 1.2$ cm to $r = 30$ cm for three solenoid fields of 2, 3, and 4 T. These hit densities can be used to optimize the position of a vertex detector or intermediate tracker. As Figure 12-25 shows, 10/ mm^2 as the maximum *photon* hit density per train, *i.e.*, before conversion, the background would seem to be completely manageable for a pixel-based vertex detector.

At this point, the lesson we can draw from these studies is that having a beam-pipe radius less than 2 cm will require a field strength greater than 2 T. Most of the photons produced come from the interaction of the pairs with the beam pipe and from backscattering from the surfaces which are the closest in z to the IP, namely the luminosity monitor at $z = 190$ cm, the front face of the quadrupole doublet at 2 m, and the septum of the rf shield, currently at 189 cm. The IR design should pull back the the beam pipe to a larger radius as soon as it is past the maximum z required by the vertex detector, limited only by wakefield considerations to have a smooth transition from the FF quadrupole inner aperture to the beam pipe. If, when all residual questions as to the normalization of the simulation are understood, the effect of the photon hits in the central tracker are supportable, we may be able to decrease the distance between the IP and the quadrupole face, which would make the FF optical design much easier. At present the conical tungsten

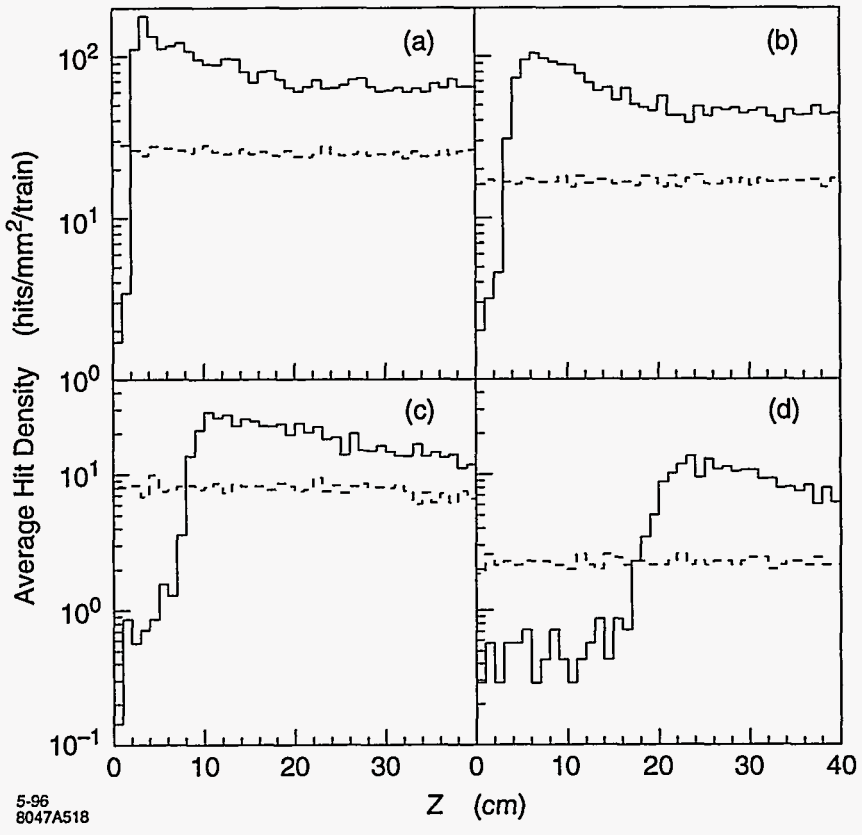


Figure 12-23. *The electron pair hit density per mm^2 per train of 90 bunches as a function of z. The solid line counts hits in scoring planes as the pairs leave the IP; the dashed line counts hits in the scoring planes resulting from the low p_t particles that have gone forward and hit the quad faces, septum, and luminosity monitor. The four parts correspond to: a) $r = 1$ cm and $B = 2T$, b) $r = 1$ cm and $B = 4T$, c) $r = 2$ cm and $B = 2T$, and d) $r = 2$ cm and $B = 4T$. The result is an EGS4 calculation using as input the ABEL beam-beam simulation at a center-of-mass energy of 1 TeV. In each case the scoring plane has an angular acceptance corresponding to $\cos\theta = 0.90$.*

mask M1 between 100 and 150 mr seems completely adequate to shield against pair induced backgrounds. The long tungsten skirt around the quadrupole pair, mask M2, appears necessary in the z area around the quadrupole face. It does not appear to be necessary that M2 neither be 3-m long, nor 10-cm thick.

Synchrotron Radiation Backgrounds

Section 12.3.3 discusses in detail the photon flux incident on the inner aperture of QFTA for the various lattice and final quadrupole options under discussion. Since, as described there, the synchrotron radiation from the Gaussian beam can be reduced to a negligible level by using an upbeam collimator, the detector backgrounds were calculated only for the synchrotron photons from the 1% flat-tail. The photon flux and energy spectrum of the tail described in Section 12.3.3 were used as an input for the EGS simulation.

Most of the photons incident on the inner surface of QFTA are absorbed in the quad, and particles penetrating the magnet can be stopped by Mask M1 and M2. However, the synchrotron photons hitting within about

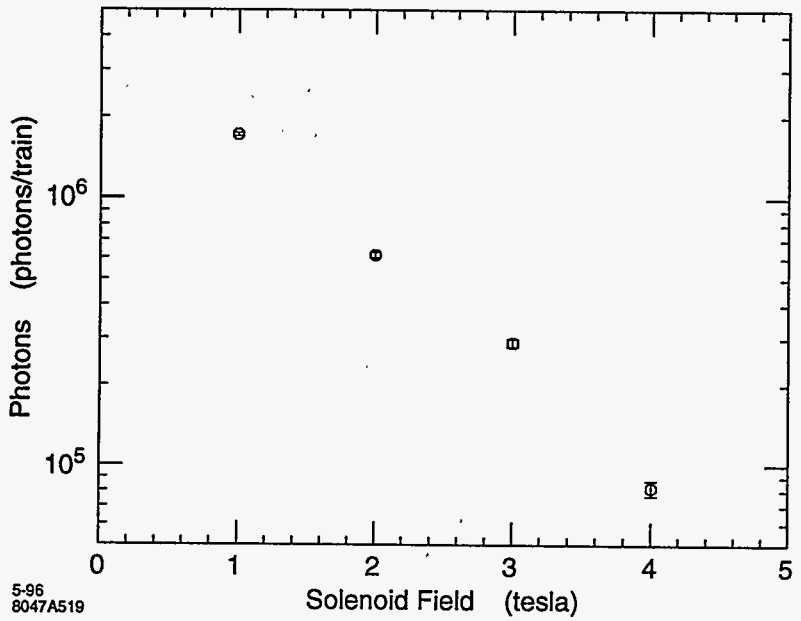


Figure 12-24. *The number of photons that strike a scoring plane at $r = 30$ cm arising from the interaction of e^+e^- pairs produced by the beam-beam interaction. The number of hits is shown as a function of the detector's solenoid field for center-of-mass energy of 1 TeV.*

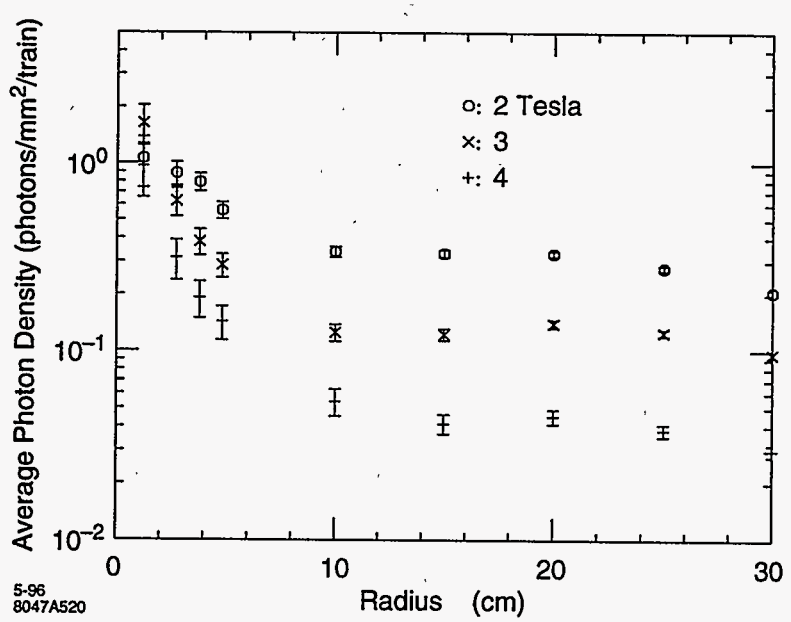


Figure 12-25. *The number of photons per train crossing that strike scoring planes at the indicated radii arising from the interaction of e^+e^- pairs produced by the beam-beam interaction. The number of hits is shown as a function of radius for three solenoid field strengths and 1 TeVcenter-of-mass energy.*

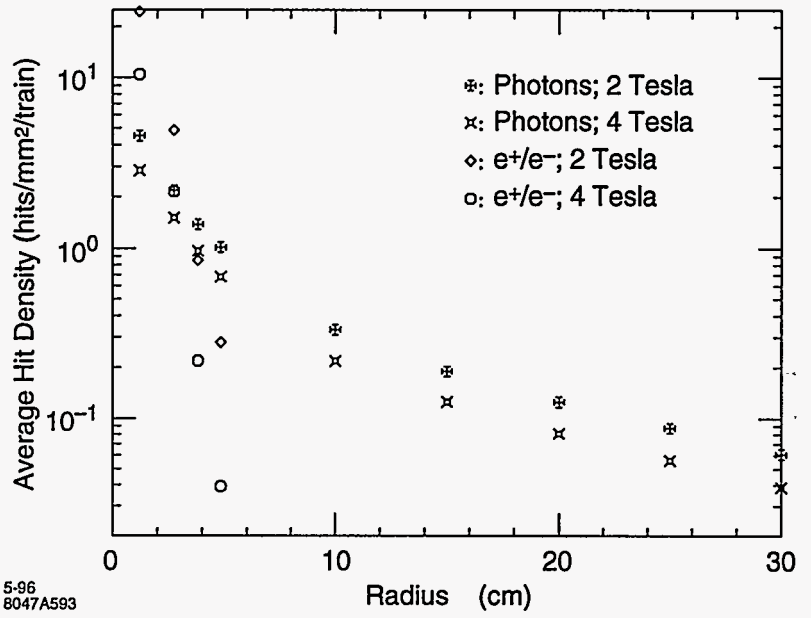


Figure 12-26. *The number of photons and e^+e^- hits per train crossing that strike scoring planes at the indicated radii arising from the interaction of photons in the inner bore of the QFTA quadrupole. These hits are produced by the quadrupole SR in the final focus lattice at 1-TeV-c.m. energy.*

10 cm of the quad exit produce secondary photons and electrons/positrons which come out of the quad and become a potential source of the detector backgrounds. The number of e^+e^- coming out of QFTA is 400K/train and the number of photons is 2×10^6/train. Those electrons/positrons that come out of the quad traverse the detector following the solenoid field and hit the vertex detector, while those secondary photons with angles between 10 and 30 mr with respect to the beam line interact with the vertex detector and rf shield within $z = \pm50$ cm or hit the downstream M1 face, contributing major backgrounds in the detector. Figure 12-26 shows the number of photon and electron/positron hits per mm^2 per train-crossing striking the vertex detector layers and massless scoring planes from $r = 1.2$ cm to $r = 30$ cm. The simulations were made for 2 and 4-T solenoid fields. At $r = 1.2$ cm, the e^+e^- hit density reaches as high as 10–20 hits/ mm^2 per train-crossing, and it may be intolerable even for a pixel-based vertex detector. While the photon density in the vertex detector is less than 5 photons/ mm^2 per train-crossing and is tolerable, the total number of photons striking the tracking chamber at $r = 30$ cm and $z=\pm100$ cm is 167K for 2 T and 110K for 4 T per train-crossing. The large photon flux may not be tolerable if we want to use a drift chamber.

The preliminary calculation described above has indicated that the SR backgrounds are potentially serious at $r = 1$ cm even for a pixel-base vertex detector and if a drift chamber is used. However, at 20–25-mm radius, such as that used by SLD at SLC, the situation is very much easier. While the backgrounds from the beam-beam interaction can be reduced by using a higher-strength solenoidal field, as shown in Figure 12-24, the SR backgrounds are not affected in the same way. Rather, the machine collimation and final-focus scheme must be designed so that the photon flux striking the inner surface of QFTA is substantially reduced. Furthermore, at 500 GeV-center-of-mass-energy, the likely startup point for an NLC, the number of SR photons goes down a factor of 2 and the average energy of the photons goes down a factor of 6, reducing the severity of this background source substantially.

Track Reconstruction Limits to the IR design

We will assume for the present that the innermost tracking detector will be a pixel-based device. The SLD has found that its vertex detector, VXD2, effectively composed of two layers of CCDs with $22\,\mu\text{m} \times 22\,\mu\text{m}$ pixels, is very robust against the backgrounds seen at the SLC, which result in average occupancies at the level of 10^{-4}, or 0.4 hits per mm^2, summed over its readout of 19 beam crossings. SLD is currently installing a new vertex detector, VXD3, consisting of three layers of larger, less massive CCDs with the same pixel size. The extra layer will allow VXD3 to be a self-tracking device. VXD3 serves as an excellent model for an NLC detector inner tracking device.

To begin to estimate quantitatively the effects of backgrounds on track reconstruction, we have employed the detailed VXD3 Monte Carlo simulation and the standard SLD track reconstruction code to study the efficiency and purity for reconstructing the correct tracks resulting from hadronic Z^0 decays at resonance as a function of the density of randomly generated background hits. The mean number of tracks in the VXD3 acceptance is approximately 15. The standard SLD track reconstruction code begins with a track found in the SLD central drift chamber. For the purposes of this study, the Monte Carlo tracks were smeared appropriately and track banks formed to simulate a central tracking device. These tracks are then extrapolated to the VXD3 and linked to pixel clusters layer-by-layer using a Kalman Filter technique. Pixel clusters (hits) in the VXD3 resulting from background can affect the reconstruction procedure in two ways: background hits can merge with real hits acting to worsen the resolution, and background hits can cause track mislinks which can pull the track away from the real hits. The purity of the VXD3 linked hits, defined as the ratio of the number of real Monte Carlo hits to the total number of hits linked per track, measures the former effect. The efficiency of linked VXD3 hits, defined as the ratio of the number of hits linked to the number of hits expected to link per track, measures the latter effect. Figures 12-27 and 12-28 show how the efficiency and purity vary as a function of background hit density ranging from 0 hits/mm^2 to 1 hit/mm^2. Both the purity and efficiency drop only a few percent across this range display the robustness of this procedure against random backgrounds in the VXD3. This work is currently being extended to study the ability to tag B mesons resulting from top and Higgs production at 500-GeV-c.m. energies. This study will continue to use VXD3 as a model vertex detector, but will use a field strength of 2 T as well as using typical resolution and efficiency parameters for the NLC detector central tracking device. It is hoped to employ the B mass tagging algorithm which has become successful at the SLD.

12.4.2 Vibration Suppression for the Final Focus Quadrupoles

A schematic of the interaction region is presented in Figure 12-29, and indicates the four quadrupoles nearest the IP for the incoming beam. Also indicated are approximate locations of the detector tracking, calorimetry and muon identification systems, and the detector solenoidal coil. The quadrupole nearest the IP, QA, is a permanent magnet; Q1 is a superconducting magnet; and the two Q2 magnets are normal magnets. QA and Q1 are vertically focusing, while the Q2s are horizontally focusing. For a 5-nm vertical spot size at the IP, the tolerances for uncorrelated vertical vibrations of QA and Q1 are at the nm level. The vertical vibration tolerances for the Q2 magnets are about a factor of 3 less severe.

QA is chosen to be a permanent magnet to achieve a transversely compact magnet with the high-focusing gradient field needed. (Note that the separation of the incoming and outgoing beams is only 40 mm at z = 2 m from the IP.) Q1 is superconducting to allow tuning of the final focus at different beam energies. The drift distance from Q1 to the IP, l^*, is chosen to be 2 m due to geometric constraints and to allow space for vertex and luminosity detectors. QA must not be located in the fringe field region of the detector solenoid,

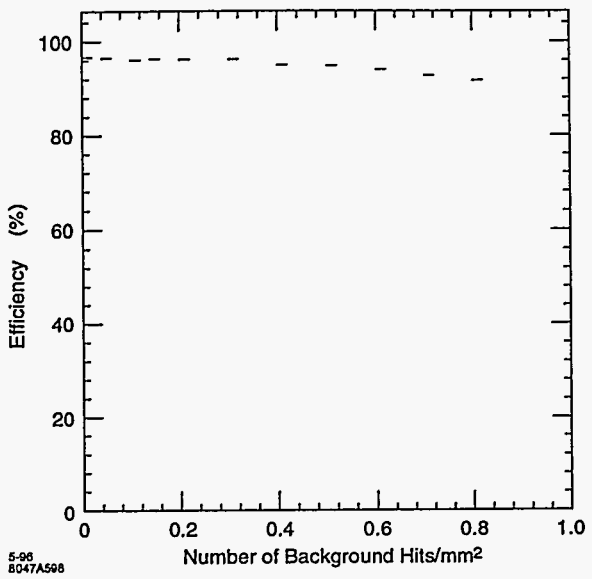


Figure 12-27. *Track efficiency vs. background level.*

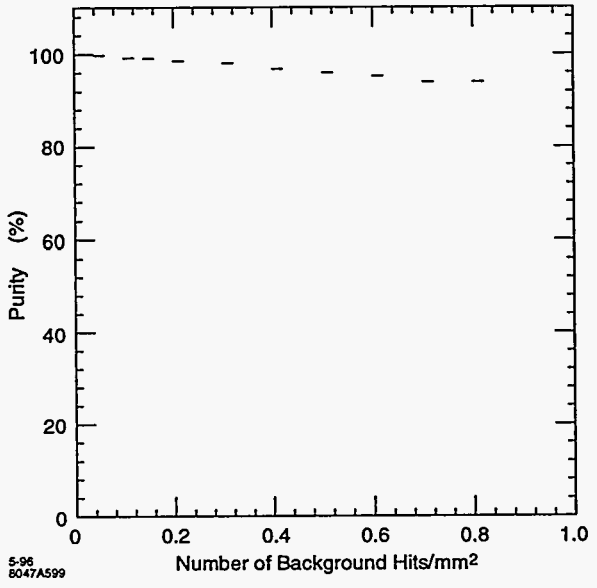


Figure 12-28. *Track purity vs. background level.*

since strong transverse fringe fields can demagnetize it. QA and Q1 are located well within the detector volume and will need to be supported from it.

The Q2 magnets are located outside the detector volume. Though their vibration tolerances are rather severe, we will assume that these magnets can be anchored mechanically to bedrock in the tunnels leading to the IP. As discussed in Appendix C and below, seismic motion in bedrock has a negligible impact on the NLC luminosity. We will not consider further the vibration stability issues for Q2, but will now focus our attention on the more difficult situation for QA and Q1. Figure 12-29 proposes that QA and Q1 be linked by optical transport arms of a laser interferometer to bedrock. This will be discussed below.

Appendix C gives a detailed description of the theory and measurement of ground motion. As is discussed there, the quadrupole alignment tolerances are very insensitive to disturbances with wavelengths which are long compared to the local lattice betatron wavelength. The dominant seismic effect is due to the microseismic peak, which has an amplitude of about 100–200 nm and a frequency of about 0.15 Hz. The wavelength of this disturbance, however, is many kilometers and so does not cause a relative misalignment of the final focus quadrupoles. Disturbances with frequencies below 1 Hz show highly correlated motion over separations of up to hundreds of meters, and the residual uncorrelated motion for the opposing final focus quadrupoles should be correctable by means of slow feedback. This feedback can monitor the deflection angles of the colliding beams to determine corrections to be made to the incoming beam trajectories. For frequencies above 1 Hz, seismic motion at quiet sites is less than 1 nm and therefore less than the tolerance for the final focus quadrupoles. Thus, if one could fix the FF quadrupoles to bedrock tens of meters below the earth's surface, seismic motion would have a negligible effect on colliding beam luminosity.

In practice, the final quadrupoles will be mounted inside a large particle detector. The structural characteristics of such a detector weighing thousands of tons and housing many layers of particle-detection apparatus in a strong 2–4-T magnetic field are similar to those of a large cube of gelatin at the nanometer scale. It will not be possible to construct independent supports for the final quadrupoles since the detector encloses nearly 4π of solid angle around the collision point. The final quadrupoles will have to be supported from the

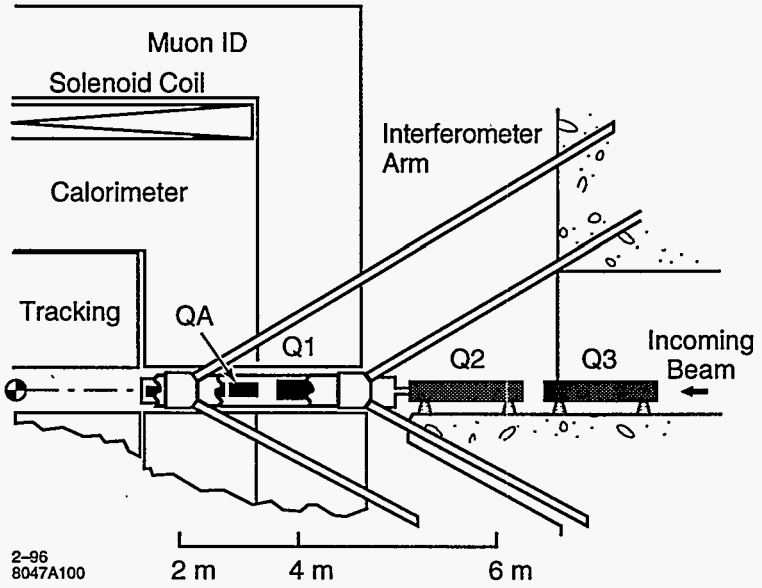


Figure 12-29. *IR final-focus quadrupoles.*

detector. Changing thermal gradients, cooling-system noise, cultural noise, and amplification of noise sources by the mechanical structure of the detector and quadrupole supports, will all contribute to mechanical noise.

Measurements made at SLAC (Section 12.4.5) and elsewhere indicate that even with good mechanical design, one may expect residual vibrations for the final quadrupoles at the level of 50-nm rms for frequencies above 1 Hz. This level of vibration will have to be measured and corrected to better than 1 nm.

Both optical and inertial measurement techniques are feasible for detecting sub-nm motion. The inertial devices are described in Appendix C. and these have been used extensively at SLAC and elsewhere to characterize ground motion. However, it may not be feasible to use these devices for the final quads that are buried inside the magnetic field of a large particle detector. An alternative method is to use a laser interferometer. A description of how to measure sub-nm displacements with an interferometer is given in Section 12.4.3, and a preliminary description for how to use such a device to stabilize the final quads is described in Section 12.4.4.

If motion of the final quadrupoles buried inside the detector can be measured with respect to bedrock, there are several approaches to stabilization. One indirect method would apply these signals to an external correction magnet to stabilize beam trajectory either by dithering its current or by piezoelectric positioning. A more direct approach would be active piezoelectric positioning of the final quadrupoles themselves. This approach has all the advantages of linearity and stability intrinsic to direct closed-loop feedback regulation around a null.

A number of activities are currently planned to address vibration issues for the IR. These include more measurements of vibrations of quadrupoles in the FFTB. The FFTB is also planning to commission an rf BPM capable of measuring beam position and beam jitter to better than 10 nm. If this is successful, they can try to correlate the beam jitter with the quadrupole jitter, and if a correlation is observed to try to correct for it. More vibration measurements are needed to try to quantify effects of cultural noise and vibration amplification by mechanical structures. Simulation work is needed to understand the vibration measurements being made and to develop engineering guidelines for a mechanical design of the IR and

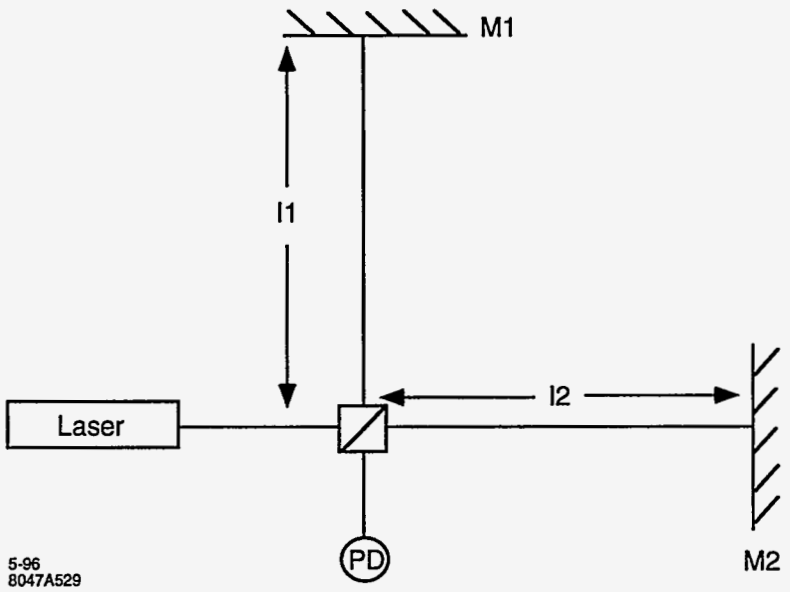


Figure 12-30. *A simple laser interferometer.*

Detector. In Sections 12.4.3 and 12.4.4, we describe an interferometer scheme to optically anchor the FF quadrupoles to bedrock. We plan to propose developing such a scheme with a mockup of an FF quadrupole to demonstrate that 50-nm vibration jiter can be measured and corrected for. This proposal will require developing independently a laser interferometer to measure the jitter and a piezo-mover system to correct it. The complete mockup would integrate a simulated FF quadrupole with its interferometer and piezo-mover systems.

12.4.3 Measurement of Sub-nm Displacements by means of a Laser Interferometer

The relative displacement of two objects at the subnanometer level can be measured by a laser interferometer. The LIGO experiment [Abramovici, 1992] proposes to measure the relative displacement of two masses, separated by 10 km, to a precision of 10^{-18} m. At the NLC, we only need to measure the relative displacement of the final quadrupole doublets, separated by 10 m, to a precision of 10^{-9} m. This may seem easy by comparison, but the geometry and accessibility of the final quadrupoles, buried inside the NLC detector, is much more complex. LIGO sensitivity will be at frequencies greater than 100 Hz, while the NLC is sensitive to frequencies below 50 Hz, where motion amplitudes are much greater. The measurement technique proposed by LIGO uses a laser interferometer somewhat more complex than the one indicated in Figure 12-30, but this figure serves to illustrate the measurement technique.

In Figure 12-30, we wish to detect relative changes in the lengths of the optical arms, l_1 and l_2, by monitoring intensity changes on the photodiode. The photodiode signal is given by

$$\begin{aligned} I &= N_\gamma(1+\cos\Phi) \\ \Phi &= \Phi^0 + \delta\Phi \end{aligned} \tag{12.10}$$

where N_γ is the number of incident photons, Φ^0 is the equilibrium phase difference between the two interferometer arms, and $\delta\Phi$ is due to changes in $(l_1 - l_2)$. The intensity change at the photodiode due to a change in $(l_1 - l_2)$, is given by

$$\begin{aligned} \delta I &= N_\gamma \sin\Phi^0 \delta\Phi \\ \delta\Phi &= 2\pi \cdot \frac{\delta(l_1 - l_2)}{\lambda} \end{aligned} \tag{12.11}$$

where λ is the wavelength of the laser light. One can choose to set $\sin\Phi^0 = 1$, in which case one gets

$$\frac{\delta I}{I} = \delta\Phi \quad . \tag{12.12}$$

For $\lambda = 633\,\text{nm}$, and $\delta(l_1 - l_2) = 1\,\text{nm}$, one will have $\delta\Phi = 0.02$ and this will cause a 2% intensity change on the photodiode.

There are many tolerances to be satisfied by the laser and optical system in this interferometer. These include tolerances on the laser's intensity stability, frequency stability and power. There are also tolerances on pressure and temperature fluctuations in the optical transport arms, which require that the transport arms be evacuated.

The laser intensity stability is required to be less than 2% for sub-nanometer displacement measurements. (Actually, the photodetector and its signal processor can be configured to detect and correct for laser intensity fluctuations. However, lasers with 1% intensity stability are commercially available.)

To determine the frequency stability required, we assume that the two optical path lengths can be equalized to better than 1 mm. Then

$$\begin{aligned} \frac{\Delta f}{f} &< \frac{10^{-2}}{2\pi} \times \frac{633 \times 10^{-9}}{2 \times 10^{-3}} \\ \frac{\Delta f}{f} &< 5 \times 10^{-7} \end{aligned} \tag{12.13}$$

The laser must have sufficient power that photon statistics will not cause significant noise in the measurement. For frequencies up to 100 Hz, we require

$$\begin{aligned} P &> \frac{(10^{-2})^2\gamma}{10^{-2}\text{sec}} \times \frac{2\text{eV}}{\gamma} \times (1.6 \times 10^{-19}\frac{\text{J}}{\text{eV}}) \\ P &> 3.2 \times 10^{-13}\ \text{W} \end{aligned} \tag{12.14}$$

Pressure fluctuations in the laser transport arms will cause a change in the index of refraction, n, and hence a change in the optical path length. In air, the dependence of the index of refraction on pressure and temperature is given by $(n-1) = 1.2 \times 10^{-4}[P(\text{Torr})/T(^\circ\text{K})]$. An optical path length of 15 m and a temperature of 300°K, will require $\Delta P < 10^{-4}$ Torr.

Temperature fluctuations will cause a change in the optical path length due to the strain coefficients for the mirrors and beam splitter. If we assume a strain of 10^{-6} per °C, then a 1-cm-thick optic can change the path length by 10nm/°C. Thus, the temperature will need to be stabilized to 0.1°C.

A laser which satisfies these requirements is a commercially available frequency-stabilized HeNe, with the following characteristics:

$$f = 473.61254\,\text{THz} \tag{12.15}$$

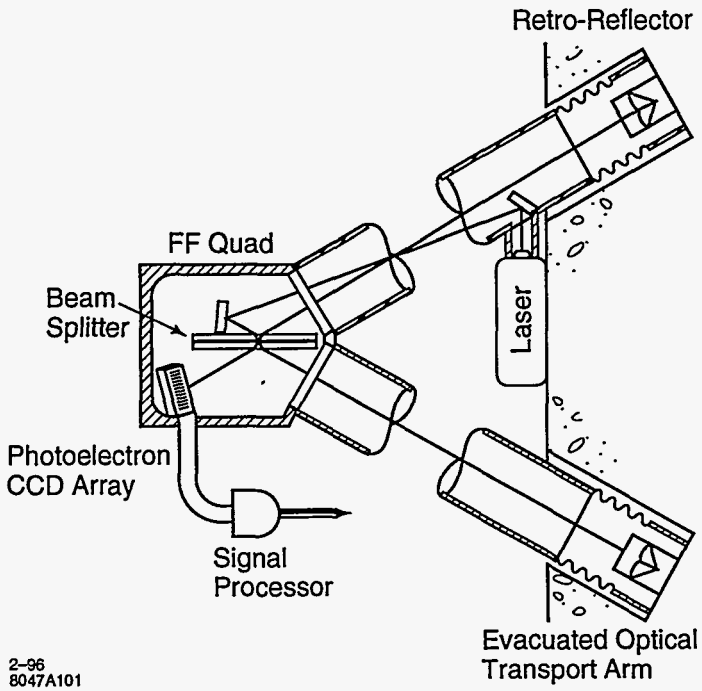


Figure 12-31. *Interferometer for Optical Anchoring of an FF quadrupole.*

$$\frac{\Delta f}{f} < 10^{-8} \tag{12.16}$$

$$\frac{\Delta I}{I} < 5 \times 10^{-4} \tag{12.17}$$

$$P = 1\,\text{mW} \tag{12.18}$$

$$\text{Cost} < \$4000 \tag{12.19}$$

12.4.4 An Optical Anchor for the Final Quadrupoles

To gain the necessary stability for collision of nanometer beams, the final quadrupoles must be rigidly connected to bedrock outside the detector. This has been done in the past by optical interferometry. [Wyatt, 2982] The final quadrupoles could be anchored to points buried deep into the walls of the detector hall by building arms of an interferometer out from the final quadrupoles to corner cube reflectors mounted to stable rock. Optical paths are angled at 60° as shown below in Figure 12-31. (Note that the choice of a 60° angle is not critical, and that smaller angles can be chosen to work equally well.)

For a 60° geometry, the differential change in the optical path between the two arms d equals the transverse motion of the beam splitter S. Each quadrupole would require individual interferometers to measure y transverse motion (we assume that measurement and correction of x vibrations is not necessary). Optical transport arms from the quadrupole to bedrock retroreflectors would be 15–20-m long and would require vacuum pipes 60–80-mm in diameter out through the detector. For each QA or Q1 quadrupole, we assume that the quadrupole acts as a rigid body (not a trivial assumption for nanometer vibrations of a 1.5-m-long

quadrupole; we will need to investigate over what lengths this assumption is valid), and that independent interferometer systems will be needed to stabilize each end of a quadrupole.

As described in Section 12.4.3, a commercial frequency-stabilized HeNe laser can be used for this system. The photodetector and its signal processing system are required to detect phase changes corresponding to less than 1/500th of a fringe (*i.e.*, less than 1 nm). We are currently investigating commercial systems to do this, but it is possible this will need to be custom-built. There are at least two companies who appear to make appropriate piezo movers with sub-nanometer resolution and accuracy, and we are investigating the specifications for these.

12.4.5 SLD Final Focus Quadrupole Vibration Measurements

Preliminary results are available from vibration measurements of the SLD final focus quadrupoles. The SLD final focus quadrupoles are a superconducting triplet supported from SLD's endcap door. The measurements were made with the STS-2 seismometers described in Appendix C.

In the final-focus tunnel leading to the collider hall and SLD, the seismometers measured about 10 nm of rms vertical motion above f = 1 Hz. On the triplets themselves, the measurements were in the range of 30–50 nm of vertical rms motion above $f = 1$ Hz. These measurements were made with the solenoid cooling water on and the cryogenic He flow on. There was no observed effect from the He flow, though there appeared to be an effect due to the cooling water. The solenoid was not powered during these tests.

It appears that triplet vibrations do not cause loss of luminosity at the SLC where the beam sizes are greater than 500 nm. And it appears that one might expect vibrations for the final focus magnets inside an NLD to be on the order of 50 nm or less for $f > 1$ Hz. (One might expect less at a quieter geographic site, and when a support design is engineered giving consideration to vibrations at the 10-nm level.) If such is the case, this vibration would need to be measured and corrected for to about 1-nm accuracy. It also appears that superconducting magnets are acceptable in the final focus, so that one is not limited to permanent magnets.

12.4.6 Measurement of Polarization and Beam Energy

One key advantage of doing physics at the NLC is that the electron beam will be highly polarized. The experimental challenge is to measure this polarization accurately enough for the physics channel under study. At the SLC, this is accomplished to an accuracy of $\simeq 0.5\%$ by colliding the electron beam with a longitudinally polarized photon beam of known polarization and using the spin asymmetry in the cross section of the resulting $e\gamma$ Compton-scattering interaction to determine the beam polarization.

Two major mechanisms induce depolarization during the beam-beam interaction. They are discussed in Ref. [Yokoya, 1988]. These are the classical spin precession under the collective field of the oncoming beam, and the spin-flip effect from beamstrahlung (Sokolov-Ternov effect). The latter is typically the most serious problem. For the range of parameters of the collider designs under discussion, one can place the limit $\Delta P < 0.04 n_\gamma$. We need to determine the precision to which the polarization needs to be measured (this will later be moved to the Physics section being developed). The details of the polarization measurement are described in the Final Focus Chapter 11 Section 11.8 on the Extraction Line.

The SLC measures the beam momenta to 20 MeV, or about 0.05%, using a spectrometer-like system which measures the separation and width of two synchrotron radiation stripes produced before and after the beam

is bent by the field of a very well measured magnet. Again we need to specify the precision to which this needs to be measured. The details of the measurement are described in the Final Focus Chapter 11 Section 11.8 on the Extraction Line.

12.5 Conclusions

More work is needed at all levels. For example, the same tools used for the other background sources can be used to place limits on the allowable number of particles in the beam tails which may strike the quadrupole aperture. We have not discussed the possibility of using fast detector timing as a means of resolving background processes, nor the general question of how backgrounds will affect trigger rates. Backgrounds have not been discussed in the context of radiation damage to detectors.

References

[Abramovici, 1992] A. Abramovici *et al., Science* **256**, 325 (1992).

[Barklow, 1992] T. Barlow, P. Chen, and W. Kozanecki, "Beamstrahlung Spectra in Next Generation Linear Colliders", SLAC–PUB–5718 Rev., (April 1992).

[Chen, 1990] P. Chen and K. Yokoya, "Beam-Beam Phenomena in Linear Colliders", US/CERN School on Particle Accelerators (1990).

[Chen, 1992] Pisin Chen, Differential Luminosity Under Multiphoton Beamstrahlung, *Phys. Rev. D* **46**, 1186 (1992).

[Feldman] G. Feldman, MUCARLO, Version 1.0 (1988).

[Frary] M.N. Frary and D.J. Miller, in Ref. [Munich, 1991], p. 379.

[Helm, TLCFF28] R. Helm, Transport Deck TLCFF28.

[SLAC 1996] R. Jacobsen, *Proc. 1996 SLAC LC Physics Workshop*, (1996).

[TRC 1995] International Linear Collider Technical Review Committee Report 1995, G.A. Loew, Chairman, SLAC–R–95–471, (December 1995).

[Keller, 1991] L.P. Keller, "Calculation of Muon Background in a 0.5-TeV Linear Collider", SLAC–PUB–5533 (April 1991).

[Keller] L.P. Keller, Muon Background in a 1.0-TeV Collider, SLAC–PUB–6385 (October 1993).

[Kuraev, 1985] E.A. Kuraev and V.S. Fadin, *Sov. J. Nucl. Phys.* **41**, 466 (1985).

[Mokhov] N. Mokhov, MARS13 muon production and transport code.

[Peskin, 1991] M.E. Peskin and M. Strassler, *Phys. Rev. D* **43**, 1500 (1991).

[Wyatt, 2982] F. Wyatt, K. Beckstrom, and J. Berger, *Bull. Seismological Soc. of Am.* **72**, 1707 (1982).

[Yokoya, 1988] K. Yokoya and P. Chen, "Depolarization due to Beam-Beam Interaction in Electron-Positron Linear Colliders", SLAC–PUB–4692 (1988).

[Munich, 1991] P.M. Zerwas, ed., e^+e^-*-Collisions at 500 GeV: The Physics Potential, Parts A and B* (Workshops at Munich, Annecy, Hamburg, 1991), DESY 92-123A&B (1992).

Contributors

- G. Bowden
- M. Breidenbach
- R. Frey
- S. Hertzbach
- L. Keller
- T.W. Markiewicz
- T. Maruyama
- R. Messner
- T. Usher
- M. Woods

13

Multiple Bunch Issues

Contents

13.1 Introduction 856
13.2 Major Impacts of Multibunching 857
13.2.1 Electron and positron sources 857
13.2.2 Damping rings 859
13.2.3 Bunch compressors 861
13.2.4 Control of multibunch beam break-up in low-frequency linacs 862
13.2.5 Main linacs 862
13.2.6 Final focus, interaction region, and beam dumps 864
13.3 Machine Protection and Operations 864
13.4 Instrumentation Specifications 865
13.5 Experimental tests related to multibunch issues 865
13.6 Summary and Conclusions 865

13.1 Introduction

Obtaining the full design luminosity in the NLC requires that a train of about 90 bunches be accelerated on each machine pulse while preserving the emittance and stability of the beam. Operating a linear collider in multibunch mode has an impact on the entire machine. Many of the multibunch issues have already been discussed at some length in previous chapters. In this chapter we give a general overview of multibunch issues, and we will point out the most crucial multibunch problems, *i.e.*, those problems that have significantly affected the overall machine design and required the most effort to solve. One important example is control of multibunch emittance growth in the main linacs; this strongly impacts the design of the accelerator structures. We summarize the proposed solutions to the problems posed by multibunch operation.

Multibunch issues exist in the SLC, since there are three bunches (the e^+ and e^- colliding bunches, and the e^- bunch used to produce the positron bunch for the next machine pulse) accelerated down the SLC linac on each machine pulse, and there are two bunches circulating in the damping rings during normal operation. However, this is a small number of bunches compared to the 90 bunches per pulse in the NLC design, and the bunches are closer together in the NLC (1.4-ns apart, as opposed to about 60 ns in the SLC linac). Multibunch issues also exist in long-pulse operation of the linac at SLAC. Multibunch beam break-up was encountered when the SLAC linac was turned on, and detuning of the dipole modes was used to help control it. Also, it was necessary to control the energy spread of the long-pulse beam.

Furthermore, the NLC main linacs are at higher frequency (11.424 GHz) compared to 2.856 GHz in the SLC linac. We have chosen to go to this higher (X-band) frequency in the main linacs of the NLC, because of the savings in power and the higher accelerating gradient obtainable. Even if we had not chosen X-band, control of multibunch beam break-up would still be an issue, but it is nevertheless much more severe at X-band. The X-band accelerator structure is smaller and closer to the beam, resulting in much stronger wakefields in the main linacs of the NLC, unless additional measures are taken to reduce these wakefields. This has been the major force driving the design of new types of accelerator structures for the NLC, namely the Gaussian-detuned structure and the damped detuned structure (DDS) discussed in Chapters 7 and 8. Once the design and fabrication techniques are developed for the main linac accelerator structures, it is also convenient to apply them to the design of the accelerator structures for the other lower-frequency linacs (S-band and L-band) that are part of the NLC design.

Regulation of the bunch charges is a second very important issue, because of its impact on multibunch energy control, particularly in the main linacs. As was discussed in Sections 7.4.5 and 8.2.8, the method chosen for multibunch energy control is to fill the accelerator structure with a field profile that simulates that of the beam-loaded steady-state in the structure. The ideal profile depends on the charge in the bunches. If the charge of each of the bunches in a train jitters by as little as a percent from pulse to pulse, this compensation of the beam loading in the main linac is upset. This places tight tolerances on the sources and may also necessitate collimation and feedforward systems to control the charge profile over the length of each train.

A third major multibunch issue is polarization of the electron beam. We wish to obtain a train of 90 bunches of electrons with at least 80% polarization. A gun with a strained GaAs cathode capable of achieving this is under development. Although such a gun is somewhat beyond what has been achieved at present, it is believed to be well within reach on the timescale needed for NLC.

These three problems—control of long-range transverse wakefields, regulation of average current from pulse to pulse, and obtaining high polarization from the e^- source—are the multibunch-related beam dynamics issues which we have identified as most critical. In addition, the development of new instrumentation, in particular new diagnostics such as structure beam position monitors, will be critical to the success of multibunch

operation in the NLC. In the next section, we will turn to a survey of these and other multibunch issues starting from the beginning and proceeding to the end of the machine.

13.2 Major Impacts of Multibunching

Multibunch issues are in general closely tied to other issues in the NLC design. For example, interbunch and intrabunch dynamics cannot be considered completely independently of each other. As noted above, the transverse dipole wake left in a linac by the bunches at the front of a train exerts transverse forces on subsequent bunches, and thus directly affects the growth of the projected multibunch emittance at the end of the linac. However, other effects come into play in determining the final emittance. The longitudinal wake left by a given bunch affects the energy and energy spread of subsequent bunches in the train. The transverse single bunch emittance can be blown up by the combination of intrabunch energy spread and transverse kicks, since particles of different energies will filament onto different trajectories unless the dispersion is zero. On the other hand, single-bunch filamentation can damp the motion of the bunch centroid and thus reduce its effectiveness as a driver of the transverse wake. For these and other reasons, single-bunch and multibunch trajectory correction and emittance control are strongly interrelated and must often be considered together.

Because of the many inter-relationships between multibunch issues and other issues, and because it was logical to organize the bulk of this design report according to geographical regions, many of the studies dealing with multibunch issues have been discussed at length in other chapters. However, in this section, for the convenience of the reader who wants an overview of multibunch effects in the NLC, we summarize the major findings of these studies and refer the reader to sections of the design report containing further details. Since some features of the design having to do with multibunching rely heavily at this time on simulations, we summarize briefly some of the simulation methods used in obtaining the results presented here and in other chapters. We will describe the solutions found to the multibunch problems and indicate where we expect further work will be concentrated as the detailed implementation of these solutions evolves.

The final goal of the NLC is to obtain two opposing trains of bunches with suitable properties for doing experiments at the interaction point (IP). The nth bunch in each train must meet its counterpart in the other train sufficiently close to the nominal IP. Thus the transverse offsets in x and y from the nominal incoming orbits must be small compared to the respective transverse bunch sizes. The two bunches meeting at the IP at a given time should pass through each other with maximum overlap. The centroid energy deviation and the energy spread of each bunch should each be no more than a few tenths of a percent. Perhaps the most difficult tolerance is that the projected multibunch emittances in each train are to be kept close to the desired transverse single-bunch emittances of $\gamma\epsilon_x = 4 \times 10^{-6}$ m-radand $\gamma\epsilon_y \approx 10^{-7}$ m-rad. Obtaining these properties at the IP requires careful control of various parameters in the other regions of the machine. In addition to tolerances on machine components, there will be a need for feedback, feedforward, special instrumentation, and beam-based correction techniques to operate in multibunch mode.

13.2.1 Electron and positron sources

There are several important multibunch issues that must be considered in designing the e^- and e^+ sources (for details on the electron and positron sources, see Chapters 2 and 3). One is bunch-to-bunch charge uniformity within a given train. Another is train-to-train total charge uniformity. The electron gun must be capable of producing a train of 90 bunches that are only 1.4-ns apart, and the population of each bunch

needs to be up to 2.8×10^{10} (NLC-IIc with 20% overhead). Furthermore, the polarization of the electron beam should be at least 80%.

A laser modulator using rf-driven resonant Pockels cells is used to turn the approximately 100-ns DC laser pulse into a train of nearly square pulses with period 1.4 ns and width 1 ns. It will also be possible to obtain trains with bunch-to-bunch spacing of 2.8 or 5.6 ns, rather than the nominal 1.4 ns.

As was noted in the previous section, the train-to-train charge jitter tolerance is very tight because of its effect on multibunch beam-loading and thus on keeping the overall energy spread of each train within tolerance. A conventional DC gun (with a strained GaAs cathode to produce polarized electrons) has been chosen for the baseline NLC design, but an rf gun, which could inject flat, smaller-emittance beams, is under consideration as a possible upgrade. One major reason (though not the only one—survival of the cathode is probably the most important reason) for choosing the DC gun over the rf gun is the difficulty of achieving the laser intensity stability tolerance in the higher-bandwidth laser that would be required in the rf gun, in order to obtain bunch trains with the required charge intensity stability. The tolerance of $< 0.5\%$ rms laser intensity stability is not easy to obtain even for the baseline NLC polarized electron source. However, it is expected that it can be obtained by improvements to a feedforward system of the type used in the oscillator for the existing SLAC polarized gun.

A intensity-limiting aperture that scrapes away about 17% of the beam before it enters the injector bunching section will be used to reduce the intensity jitter below the very small required tolerance of about 0.5%. This tolerance is what is required to achieve the desired beam loading compensation in the X-band main linacs.

Beam-loading compensation in the various accelerator structures that are part of the e^- and e^+ sources is another significant multibunch issue in this region. Two basic methods of beam-loading compensation were considered. One possible energy compensation scheme (Δt scheme) is to turn the beam on before the structure has completely filled. The additional filling of the structure while the beam is passing through compensates the linear part of the beam loading. The slope of the SLEDded rf pulse can also be adjusted to compensate the quadratic "droop" in energy over the train.

Another possible scheme (Δf scheme) utilizes additional cavities driven at $\pm\Delta f$ away from the nominal central frequency. The resulting variation in phase from bunch to bunch can be used to cancel some of the variation in beam loading.

The Δf scheme has the advantage of being relatively easy to tune (by changing the amplitude of the fields in the Δf accelerator sections). However, the Δt scheme was selected for most of the source linacs because the Δf scheme gives a single bunch energy spread that is too large. A combination of the Δt and Δf schemes will be used in the e^+ capture linac, the e^- capture section, and the bunch compressor S-band prelinac.

Long-range transverse wakefields must be kept small enough to prevent multibunch beam blow-up in the source linacs. Satisfactory control of the multibunch emittance was achieved by using Gaussian-detuned structures for the positron booster linac (see Section 3.6.1) and the electron injector linac (see Section 2.4.5).

Multibunching also presents additional demands on the positron target (see Section 3.4), which must be capable of withstanding the peak and average power in the e^- beam impinging upon it. The design will be based on that of the positron target of SLC, with improvements to allow higher beam power and better intensity stability.

13.2.2 Damping rings

The damping rings of the NLC are larger and more complicated to design than those of the SLC. Each SLC damping ring contains at most two bunches at a time, while the NLC damping rings each contain four trains of about 90 bunches each; the beam loading in the NLC damping rings is much heavier than in SLC. Also, since the bunches are not distributed uniformly about the circumference (there is a gap between trains to allow time for the kickers to inject and extract a train from the ring on each machine pulse), there is a variation in the synchronous phase along each bunch train.

The required emittances of the bunches extracted from the damping rings are $\gamma\epsilon_x = 3 \times 10^{-6}$ m-radand $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad. The number of electrons per bunch in the damping ring of NLC-I is about 1×10^{10}, and goes up to about 1.3×10^{10} for NLC-II and NLC-III. The maximum charge per bunch in the most extreme design variations under consideration is about 1.5×10^{10}, which for four 90-bunch trains in a ring of circumference ~220 m, leads to a maximum average current of about 1.2 A.

The two main damping rings (one for electrons and one for positrons) each damp four 90-bunch trains at a time, with one train being extracted and one train immediately injected in its place on each machine pulse. The reason for simultaneously injecting and extracting a train from each ring on each machine pulse is to minimize transients in the rf cavities that would be produced by changes in the average ring current. In addition to the main damping ring, there is a pre-damping ring for the positrons which damps three 90-bunch trains at a time.

Although the trains are separated by many buckets, they can still affect each other through long-range wakefields, unless these wakefields are quite heavily damped. Injection and extraction of bunch trains must be done with minimal disturbance of other trains in the ring. There are gaps of about 60 ns between trains, so the kicker rise and fall time must be comfortably less than this, and ringing of the kicker pulse must be minimized. The separation of 60 ns between trains is about equal to the rise and fall times of the kickers existing at present in the SLC. A flattop of about 130 ns is needed to accommodate the 90-bunch train. None of the requirements on the kickers are especially difficult, although the positron pre-damping ring kicker will need to kick more strongly due to the relatively large aperture.

Beam loading and synchronous phases

As noted above, the beam loading seen by the damping ring rf system will vary in time, due to gaps between trains, and the synchronous phases of the bunches in a train are different due to the different amount of beam loading seen by each bunch. Unless the bunch-to-bunch variation in beam loading is compensated within the ring (*e.g.*, by a special higher-harmonic cavity in addition to the regular rf cavities), it must be compensated further downstream, presumably in the bunch compressors. The variation in phases along the bunch train is very nearly linear if no phase compensation is performed in the damping ring. The present bunch compressor design is able to perform compensation for this phase variation (see Section 5.4.7). However, two methods of compensating the phases while still in the damping ring are also being considered, in an effort to simplify the requirements on the bunch compressors. One possibility is to vary the generator voltage as a function of time; this requires that the klystron have sufficient power and bandwidth and is currently under study. Another possibility is to use passive, lower-frequency harmonic cavities to partially compensate the phase variation (see Section 4.4.4); the main disadvantage of this scheme is that the pattern of synchronous phases versus bunch number becomes very nonlinear, and it would be difficult to remove the residual phase variation downstream, if this were necessary.

The variation in synchronous phases due to changes in charge of a bunch train has also been studied in simulations (see Section 4.4.4). Even for a change in average charge of 5%, the resulting phase variations could be easily compensated by a damping ring phase feedback system.

Coupled-bunch Instabilities

In addition to the effect on the synchronous phases, the longitudinal wakefields (both the fundamental and higher-order modes in the cavities) produce longitudinal coupled bunch instabilities. Preliminary rf cavity design and coupled-bunch simulations (see Section 4.4.5) indicate that it should be possible to damp the longitudinal higher-order modes (HOMs) to keep the threshold for longitudinal coupled-bunch instabilities comfortably below the radiation damping rate.

The transverse wakefields, due to both the rf cavities and the resistive wall of the vacuum chamber, can produce transverse coupled-bunch instabilities in the damping rings. Assuming that the cavity HOMs are damped to have Qs less than 300~500, then the resistive-wall impedance dominates. A bunch-by-bunch feedback system will be needed to damp any modes that are not suppressed by radiation damping, coherent head-tail damping, or Landau damping. Even if there is sufficient damping present that all the normal modes of oscillation are stable, interference between modes can produce transient blow-up of the beam. This transient behavior can be important in damping rings since the storage times are short. In addition, for sufficiently strong wakes and long trains of bunches, the transient could be large enough to cause beam loss at injection. A bunch-by-bunch feedback system along the lines of that designed for the PEP-II B-factory at SLAC will be used to suppress these effects.

Several coupled-bunch simulation programs have been used in calculating longitudinal and transverse coupled-bunch instabilities in the damping rings. Some are based on a semi-analytic, normal-modes approach, in which the bunches need not be symmetrically placed on the circumference. Interference between the modes can produce transient blow-up of the beam even if all these modes are long-term stable. Given the coherent frequencies and normal modes, the Laplace transform can be used to obtain the motion of the bunches, taking the initial conditions into account [Thompson 1991a]. Alternatively one may use a computer tracking method to obtain the offset of each bunch as a function of time. This is straightforward and computationally efficient provided that the number of bunches is not too large and the wakefields do not persist for too many turns. Several tracking codes are in use for NLC damping ring calculations [Thompson 1991b, Thompson 1991c, Byrd 1993].

A new code to investigate coupled-bunch mode-coupling was also developed [Berg 1995]; however, this turns out not to be a significant effect in the damping rings (see Chapter 4).

Ions and other effects

A possible multibunch issue in the damping rings (and also in the main linacs) stems from the fact that the bunches in a train are coupled not only by long-range wakefields but also by the fields due to ions in the beam line [Raubenheimer 1995]. If there is significant collisional ionization and if ions remain trapped between the passage of successive bunches, then bunch-to-bunch coupling can be mediated by the ions (similar to the way that surrounding structures mediate transverse wakefields). Ions can also produce a focusing variation between bunches, which may lead to filamentation of the trajectories of different bunches. Control of these effects may put stringent requirements on the vacuum in the damping rings and linacs, according to simulations of the beam dynamics with ions present. It should be noted that our concerns about ions are

based mainly on simulation results, and experimental studies of the effects of ions on multibunch operation are needed.

Calculations predict that ions produced by the beam scattering with residual gas in the electron damping ring vacuum chamber can produce a fast transverse instability within a bunch train (see Section 4.4.6). Simulations and analytical estimates suggest that a vacuum pressure of 10^{-9} Torr or better may be required to control this instability. If this is not adequate, additional gaps in the bunch train may be used to clear the ions, but obviously this is a somewhat inelegant solution. Other solutions, such as "detuning" the ion frequencies or lowering the equilibrium emittance may be possible. More work, including experimental tests, is still needed.

A different multibunch instability may occur in the positron damping ring (see Section 4.4.7). This occurs when an electron cloud is produced in the vacuum chamber from photoelectrons and their collisions with the walls to produce secondary electrons. This electron cloud can couple to the transverse motion of the bunches and lead to an instability. There is some evidence for such an instability in the KEK Photon Factory and in CESR. Theoretical predictions of the coupled-bunch instability growth rates [Ohmi 1995] agree approximately with what has been observed. Present estimates for the NLC positron main damping ring give a characteristic growth time scale of about 200 ns. If this estimate is correct, it should be taken into account in the feedback systems being designed to combat coupled-bunch instabilities. Estimates also need to be made for this possible instability in the pre-damping ring.

Experience obtained in high-current, multibunch storage rings currently in operation or under construction is of course important to the design of the NLC damping rings. The electron-positron instability will be studied in the APS as it begins operation with positrons and in the PEP-II Low Energy Ring (LER). The ion-electron instability will be studied at the ALS, the PEP-II LER, and the KEK Accelerator Test Facility.

13.2.3 Bunch compressors

The main multibunch issue in the bunch compressor is compensation of the multibunch beam loading. Multibunch beam break-up must also be controlled in the various linacs that are part of the compressor design; these are discussed in the next section.

Compensation of beam loading

As noted above, the bunch compressors may need to perform the compensation of phase offsets produced by differential bunch-to-bunch beam loading in the damping rings. In addition, there is beam loading in the bunch compressors themselves, and the resulting bunch-to-bunch energy differences must be kept sufficiently small. It is possible to compensate the multibunch beam loading in the bunch compressor by using two rf systems having slightly different frequencies; this was assumed in the initial design studies and satisfactory results were obtained (see Section 5.4.7). However this "Δf" scheme of beam loading compensation has the disadvantage of being somewhat nonlocal, since the beam energy spread grows between the off-frequency compensation sections. As in the injectors, a combination of this "Δf" method and the "Δt" (early injection) method will be used to obtain even better results.

13.2.4 Control of multibunch beam break-up in low-frequency linacs

The transverse emittance of the multibunch trains must be controlled in the low-frequency linacs, throughout the front end of the NLC (in the sources and compressor regions). The multibunch beam break-up can be controlled by using Gaussian-detuned or damped detuned structures, as in the main linac. The pre-linacs are at lower frequency than the main linacs, so the wakefields are not as strong. However the beam is at lower energy, which makes it more susceptible to wakefield kicks. Simulations show that methods similar to those studied for the main linacs, namely the use of detuned or damped detuned accelerator structures, will control the break-up.

In some of the S-band (2.856 GHz) linacs, detuning alone may not be quite sufficient to control multibunch beam break-up. Thus an S-band damped detuned structure (DDS) is being designed for use in all the S-band linacs. It will have a total detuning frequency spread of about 6%. The modes will be damped to Qs of about 1000.

In the L-band (1.428-GHz) positron linac, Gaussian detuning with a 10% total spread is sufficient to control multibunch beam break-up (see Section 3.6.1).

13.2.5 Main linacs

As was noted at the beginning of this chapter, one of the most important issues in the design of the main-linac accelerator structures is control of the transverse wakefield. The achievement of an X-band accelerator structure that will accomplish this has been one of the major efforts in the design of the NLC. This damped detuned structure (DDS) has been discussed at greater length in Chapters 7 and 8. The structures are detuned by varying the individual cell dimensions in such a way that there is an approximately gaussian (truncated) distribution of frequencies of the fundamental dipole mode. The dipole modes in the structures are damped via ports leading into manifolds running parallel to the structures. Construction of the first DDS is nearing completion, and DDSs will be part of the complement of accelerator structures for the NLCTA.

Calculation and measurement of long-range wakefields

Calculation of the long range wakefields in the new accelerator structure designs being proposed for the NLC has been the focus of much effort. Over the past several years, increasingly sophisticated models of the wakefields in increasingly complex accelerator structures have been developed.

A fairly good representation of the long range wake was obtained in initial simple models of detuned structures, by regarding the structure as a collection of uncoupled oscillators corresponding to the synchronous modes of the periodic structures that one could construct from each of the cells in the structure. However, a more complete and accurate treatment includes the effects of the small couplings between the oscillators. A discussion of two such models [Bane 1993], a single-passband model and a model which takes into account the mixing of the two lowest dipole passbands, was given in Section 7.4.2. These are equivalent circuit models that give the best representation we have so far obtained for the wakefields in the detuned accelerator structure without damping (except for copper losses, which are taken into account in the models via perturbation theory).

These coupled, equivalent circuit models can be further extended to include the interaction between the accelerator structure and the damping manifolds in a damped detuned structure (DDS). The first such

models treated the manifolds as coaxial lines, ignoring the periodicity introduced by the openings from the cells into the manifold [Kroll 1994]. Only a single passband was included in most of this work. A more complete model has now been developed that takes into account the mixing of the two lowest dipole passbands, as well as the periodicity of the manifold [Jones 1996].

A crucial element in the design of the NLC main linacs is an accelerator structure in which the transverse wakefields are greatly reduced below those that would occur in a conventional disk-loaded structure. Experimental verification of the performance of such structures is therefore essential. The Accelerator Structure Setup (ASSET) facility in the SLC has been used to measure the wakefields in a Gaussian detuned X-band structure (see Section 8.2.10). Damped detuned structures will also be tested in ASSET as they become available during the coming months.

Beam dynamics simulations in main linacs

The results of beam dynamics simulations in the main linacs, including multibunch effects, have been discussed extensively in Chapter 8, and we will not repeat that discussion here. In this section we will briefly describe the simulation tools used to obtain those results. These simulations incorporate the calculated long range wakefields, to calculate the multibunch beam blow-up to be expected, the bunch train injection tolerances, structure misalignment tolerances, and the effects of various correction schemes. As was noted in Chapters 7 and 8, it has been found that the structure internal misalignment tolerances are very tight, due to the effect on multibunch emittance growth. Also, control of the multibunch energy spread imposes tight tolerances on the variation of train current from pulse to pulse.

Several codes have been used to study multibunch beam break-up in the main linac and in the other linacs in the NLC. These include: LINACBBU [Thompson 1990], MBLINAC [Thompson 1991d], and LTRACK [Bane 1987]. A program that can handle combined single- and multibunch emittance control and trajectory correction was developed [Kubo 1995] and used to make the initial studies on these issues.

The main simulation tool used so far to study multibunch energy compensation in the main linacs is the program MBENERGY [Thompson, 1993]. In this simulation, one may take account of input rf pulse shaping and timing, the dispersion of the rf pulse as it transits the structure, the longitudinal distribution of charge within the bunches, the long range wake (LRW) including both the fundamental (accelerating) mode and higher order modes (HOMs), the short range wake (SRW), and phasing of the bunches with respect to the crests of the rf.

Ion effects may be important not only in the damping rings, but also in the main linacs [Raubenheimer 1992]; the basis for this expectation is almost entirely calculations and simulations. Results of these simulations are described in Section 7.4.6 and suggest that the vacuum tolerances in the main linac will be very tight, although achievable.

A new linac code, the Linear Accelerator Research (LIAR) code has recently been developed [Assmann 1996] to do the many detailed tolerance studies that will be needed over the coming months. This code is designed to flexibly accommodate new features, and is now beginning to be used for multibunch tolerance studies. It will be possible with this code to do more complete simulations that incorporate multibunch effects along with other effects to get a more accurate assessment of the many tolerances required to preserve the emittance of the multibunch beam.

13.2.6 Final focus, interaction region, and beam dumps

There are not any multibunch problems in the final focus region that are as difficult as those that must be faced in other parts of the machine. However, multibunching does have an impact on the design of the final focus, interaction region and beam dumps.

"Parasitic encounters", *i.e.,* kicks on bunches in the incoming train due to bunches in the outgoing train, must be kept sufficiently small. This is the main reason for the introduction of a crossing angle at the interaction point. The crossing angle would entail a significant loss of luminosity since the bunches are long and thin, if the bunches were allowed to be non-parallel when they collide.

To avoid this loss in luminosity, special rf cavities are used to rotate the bunches away from their direction of travel just before they collide, so that the longitudinal axes of opposing bunches will be parallel when they pass through each other. It has been checked that these "crab cavities" do not produce unacceptable kicks due to bunch-to-bunch wakefields (see Section 11.7.2).

Another component affected by multibunch operation is the beam dump, which must be able to absorb the large amount of power in the multibunch beam. The requirements on the beam dumps are very stringent, due to the large amount of total charge, small emittance, and high energy of the beam. A dump using water as the main absorbing material has been designed (see Section 11.A) to handle the electromagnetic shower from the multibunch beam of up to 750 GeV.

13.3 Machine Protection and Operations

The long bunch trains at design emittance are capable of seriously damaging certain components (including the main linac accelerator structure) on a single pulse. Thus, start-up and recovery procedures will be strongly influenced by multi-bunch issues (see Chapter 16 and Section 7.8). The control system must be capable of handling a variety of bunch patterns and modes of operation, and the stability (*e.g.,* temperature of certain components) of the machine must be preserved as one cycles through the various modes of operation.

Even a single bunch at the design emittance may be capable of damaging the machine. It is anticipated that sacrificial titanium spoilers and collimators would be placed so as to prevent damage to crucial machine components. Such spoilers would be able to withstand a single bunch of nominal emittance and intensity, but not a full nominal-emittance bunch train. One must begin with single bunches having relatively large emittance and work up from there as the machine is tuned.

Note also that once operation with the full bunch train has been established at nominal emittance, the repetition rate must be kept high enough that the beam trajectory cannot change too quickly between pulses. Changes must be monitored and the beam must be shut off or have its emittance blown up before the trajectory wanders far enough to damage accelerator structures or other critical components.

Controlling the emittance, energy, energy spread, and trajectories of all the bunches in a train of 90 bunches is not trivial, even in simulations. There are a number of aspects of multibunch running that will require detailed online simulation and control. One such example is fine-tuning the input rf pulse in linac accelerator sections, as part of a feedback system to improve the multibunch energy compensation. Obviously this is only one of many examples—the control system software will be required to do detailed online simulations related to emittance control, trajectory correction, feedback, etc throughout the machine.

13.4 Instrumentation Specifications

Meeting the required specifications at the IP also imposes certain requirements on instrumentation of various parts of the machine. These include: BPMs that can resolve bunches (or a few bunches) within a train, multibunch emittance measurement stations, multibunch energy measurement stations, and fast kickers (bandwidth sufficient to correct alignment of trains). Design of such instrumentation is underway, and discussions are given in Sections 7.3.4 and 7.10. and in Chapter 15.

The tolerances on the alignment of the structures with respect to the beam are very tight due to the need to control the transverse multibunch wake; the tolerance is only a few microns on some scales. Thus, one of the most important issues is instrumentation of the accelerator structure to measure its offsets with respect to the beam, via detection and analysis of the beam-induced dipole wakefield. This method of measuring the alignment of the structures by looking at signals derived from the dipole wake is discussed in Section 7.10.3. and experimental tests of the method are being carried out in the SLC.

13.5 Experimental tests related to multibunch issues

A crucial element in the design of the NLC main linacs is an accelerator structure in which the transverse wakefields are greatly reduced below those that would occur in a conventional disk-loaded structure. Experimental verification of the performance of such structures is therefore essential. The Accelerator Structure Setup (ASSET) facility in the SLC has been used to measure the wakefields in a Gaussian detuned X-band structure (see Section 8.2.10). Damped detuned structures will also be tested in ASSET when they become available.

We of course wish to verify the multibunch beam dynamics simulation results as soon as possible, and so a number of multibunch beam dynamics experiments will be done in NLCTA (see Section 8.2.9), including detailed tests of the multibunch energy compensation scheme and measurements of multibunch beam breakup. Studies of the latter will be greatly facilitated by an upgraded injector for NLCTA that produces a bunch train with a charge per bunch and bunch spacing similar to that in the NLC designs.

Experience obtained in high-current, multibunch storage rings currently in operation or under construction is of course important to the design of the NLC damping rings. Bunch-by-bunch feedback systems similar to that needed to suppress transverse coupled bunch instabilities in the NLC damping rings are being built for the PEP-II B-factory and other high-current storage rings.

13.6 Summary and Conclusions

The proposed NLC design relies heavily on multibunching to achieve the design luminosity. Some of the resulting tolerances are very tight—two important examples are the alignment of the X-band accelerator structures in the main linacs, and variations in the bunch populations from pulse to pulse. However, we believe that they are achievable by the methods we have proposed.

Experience at the SLC provides considerable guidance in pursuing solutions to the problems posed by multibunch operation. While in some ways SLC experience is limited by the fact that the number of bunches

per pulse is small compared to that in NLC, there is much that can be done in SLC that is directly relevant, particularly in the characterization of the long-range wakefields and in the development of instrumentation. Experience with high-current storage rings, such as PEP-II, which have average currents even greater than that proposed for the NLC damping rings will also be valuable, especially in refining the design of feedback systems for the NLC damping rings. As has already been emphasized, one of the most crucial elements of the NLC design is an accelerator structure for the X-band main linacs that adequately controls the transverse wakefields. Here we must rely on simulations to guide the design and ongoing experimental tests to verify that the structures work as planned. Tests in the SLC will also be important to the development of instrumentation, such as structure beam position monitors that use measurements of the induced dipole wakefields to infer the alignment of the structures with respect to the beam.

Our goal so far has been to find satisfactory conceptual solutions to the problems we will encounter in multibunch operation at the NLC. Considerable detailed design and engineering remains to be done, for example, on the various feedback systems that will be required to control the multibunch emittance and energy spread throughout the machine, as well as the associated instrumentation. Another major focus will be to continue the detailed engineering of the most practical ways to manufacture the large number of accelerator structures for the X-band linacs.

References

[Assmann 1996] R. Assmann, C. Adolphsen, K. Bane, T. Raubenheimer, R. Siemann, K. Thompson, SLAC–AP–Note 103 (1996).

[Bane 1987] K.L.F. Bane, 1987a, "Wakefield Effects in a Linear Collider", in *Physics of Particle Accelerators*, edited by M. Month and M. Dienes, *AIP Conf. Proc. No. 153* (AIP, New York, 1987).

[Bane 1993] K.L.F. Bane and R.L. Gluckstern, 1993a, "The Transverse Wakefield of a Detuned X-Band Accelerating Structure", SLAC–PUB–5783, in *Particle Acclerators* **42**, 123 (1993a).

[Berg 1995] S. Berg, 1995, "Transverse Coupled Bunch Mode Coupling and Growth Rates for the NLC Main Damping Ring", NLC–Note 11, (February 1995).

[Byrd 1993] J. Byrd, 1993, "Study of Coupled Bunch Collective Effects in the PEP-II B-Factory", Presented at 1993 Particle Accelerator Conference, Washington, DC (1993).

[Jones 1996] R.M. Jones, K. Ko, N.M. Kroll, R.H. Miller, and K.A. Thompson, submitted to 1996 European Particle Accelerator Conference.

[Kroll 1994] N. Kroll, K. Thompson, K. Bane, R. Gluckstern, K. Ko, R. Miller, R. Ruth, "Manifold damping of the NLC detuned accelerating structure", 6th Workshop on Advanced Accelerator Concepts, Lake Geneva, WI, (1994a).

[Kubo 1995] K. Kubo, K.L.F. Bane, T.O. Raubenheimer, K.A. Thompson, NLC–Note–14 (1995).

[Ohmi 1995] K.Ohmi, "Beam and Photo-electron Interactions in Positron Storage Rings", KEK preprint 94–198 (1995).

[Raubenheimer 1992] T. Raubenheimer and P. Chen, "Ions in the Linacs of Future Linear Colliders", SLAC–PUB–5893, presented at LINAC 92, Ottawa (1992).

[Raubenheimer 1995] T.O. Raubenheimer and F.Zimmermann, "A Fast Beam-Ion Instability in Linear Accelerators and Storage Rings", SLAC–PUB–6740, submitted to *Phys. Rev. E* (1995a).

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

[Thompson 1991a] K.A. Thompson and R.D.Ruth, "Transverse coupled-bunch instabilities in damping rings of high-energy linear colliders", SLAC–PUB–4801, in *Phys. Rev. D* **43**, 3049 (1991).

[Thompson 1991b] K.A. Thompson, "Simulation of Longitudinal Coupled Bunch Instabilities", ABC–Note–24, (7 February 1991).

[Thompson 1991c] K.A. Thompson, "Extension of Coupled Bunch Ring Simulation to Transverse Case", ABC–Note–57, (31 December 1991).

[Thompson 1991d] K.A. Thompson, unpublished (1991).

[Thompson, 1993] K.A. Thompson and R.D. Ruth, "Simulation and Compensation of Multibunch Energy Variation in NLC", SLAC–PUB–6154, in *Proceedings of 1993 Part. Acc. Conf.*, Washington, DC (1993);

Contributors

- Kathy Thompson

14

Control System

Contents

14.1 Introduction 870
14.2 NLC Requirements 870
14.2.1 Feedback 871
14.2.2 Synchronized measurement 872
14.2.3 Data acquisition and processing 872
14.2.4 Special architecture for damping ring applications 873
14.2.5 Tuning 873
14.2.6 Modeling and simulation 873
14.2.7 Accelerator/Detector Coupling 874
14.2.8 Reliability and Availability 874
14.3 Architectural implications 875
14.4 The Control System Model 876
14.4.1 Operator Consoles 878
14.4.2 Application and Server Computing Resources 879
14.4.3 High Speed Networks (FDDI & Ethernet) 879
14.4.4 Front-end Computers 880
14.4.5 Data Acquisition Crates 880
14.4.6 Instrumentation Modules (VXI & GPIB) 880
14.4.7 Radio Frequency Control, Phasing & Feedback 880
14.4.8 Dedicated Control Networks 880
14.4.9 Timing and Beam Rate Control 881
14.4.10 Machine Protection Systems 881
14.4.11 Equipment & Tunnel Access Control 882
14.4.12 Application Software 882
14.4.13 Software Application Bus 883
14.4.14 Software Development Environment 883
14.4.15 Industry Standards 883

14.1 Introduction

Control system implementation is notoriously dependent on available technology. "Notoriously dependent," because its technology base—computers, communications, and electronics—is developing at the most rapid rate of all components used in accelerators. Thus design proposals for systems to be built in five to ten years from the date of a proposal are often little more than examples of how dated ideas can become in that period of time. In spite of that fact, some rough cut on control system design must be made now. The control system will interact with most other components of the accelerator. It is in everyone's interest to define the most important features—in terms of functionality, not implementation—of hardware and software interfaces early. This will promote uniformity where appropriate and will give other systems an opportunity for review.

In this chapter, we will acknowledge that proposals for a control system design are transient and focus on two points:

1. Although the implementation design of a control system may be transient, the functional requirements that such a system must meet are much less so. Operations at the SLC and the Final Focus Test Beam have given us valuable lessons for the control system needs of an NLC [Humphrey 1992]. Thus, the first section of this chapter will describe some of the key functional requirements for NLC operations that the control system must supply and the second section will discuss some of the implications of these requirements for the architecture.
2. A serious proposal to build NLC must include a cost estimate for all significant components. The cost estimate for the control system can only be based on a specific implementation; therefore, an implementation model, based on presently available technology, is proposed in the third section of this chapter.

14.2 NLC Requirements

A control system should be tailored to its users and the challenge they face. Much of its bulk is devoted to device controller interfaces, networks, consoles, etc.; those parts will not be discussed at length in this chapter, though some rough idea of the numbers of such devices will be needed for the cost estimate and for discussions of reliability (Chapter 17). But, there are many aspects of a linear collider that impose requirements on the control system that are distinct from normal accelerator control systems. Most of our understanding of these requirements comes from experience with the SLC and its control system, which has evolved considerably from its initial implementation. They include:

- Beam-based feedback
- Measurement synchronization with beam pulses and high-bandwidth data acquisition
- Special handling of damping ring applications
- Automated procedures for tuning
- Machine modeling and simulation
- Accelerator/detector coupling.

14.2.1 Feedback

Beam-based feedback, of the type described in Appendix D. should be considered a possible engineering solution for all tight tolerances which are static enough to be within the correction bandwidth of the feedback system. Furthermore, beam-based feedback should be considered for all systems whose initial fabrication errors place them outside of predicted tolerances. However, for systems with disturbances outside of the feedback bandwidth, or "within the bunch train," other solutions such as feedforward or dedicated, very high-bandwidth feedback systems should be considered.

There are several aspects to this overall approach:

- Use of the control system, where effective, to offset the cost of meeting tight tolerances by modifying or rebuilding the system affected.
- Development of feedback to control instabilities of all time scales, especially mid to long term. In the SLC "mid term" is the highest frequency presently addressed by the fast feedback system—120 Hz.
- Strong effort to counter fast instabilities with high-powered tools and feedforward.
- Consideration of high-bandwidth feedback for systems with very tight tolerances.
- Establishment of *treaty regions* with full separation of phase space and centroid stability.

It is critical to identify which technical issues can be addressed by these means and which cannot. Mechanical, thermal, slow magnetic fields, and low-bandwidth microphonic disturbances have time scales which can be addressed by such feedback systems. Pulsed devices such as kickers and modulators must be stabilized using other feedback techniques (not beam-based).

The present system of feedback at SLC was built after much of the lower-level design was mature. It evolved through several generations of development, each incorporating the lessons learned. The pervasiveness of the use of feedback throughout the accelerator controls was not fully anticipated and therefore the system is not optimally integrated. This requires attention in the detailed design of an NLC control system.

Fast Feedback Rate Considerations

The SLC feedback system is designed in a generalized, database-driven fashion, which contributes greatly to its flexibility. The system uses standard control-system hardware. As a result of this design, unplanned control loops can often be added with only database work, without requiring hardware or software changes. In addition to the beam position monitors and correctors used for steering control, the feedback system is capable of measuring and controlling a wide variety of devices. For example, it is as easy to add a steering loop in the linac as it is to provide control of laser gun timing in the injector. Special-purpose extensions to the linear feedback system have been added to accommodate non-linear cases, such as optimization feedbacks in which the measurement responds parabolically to actuator movement. The system also provides built-in diagnostic and analysis capabilities, and the many sample-only monitoring loops provide a wealth of diagnostic information. These design features have been key to the success of the SLC system, and the NLC design should be equally flexible and extensible in order to support unplanned controls needs.

It must be noted that for this generalized feedback system to perform properly sufficient resources must be available. The specific system used in the SLC employs dedicated point-to-point links in order to minimize communication overhead. A typical launch feedback loop at the end of the SLC Linac stabilizing angle and

position in two planes for both electrons and positrons absorbs approximately 70% of the available bandwidth of an INTEL 486 single board computer running at a 66-MHz clock rate.

The feedbacks for the NLC are planned to run at the full beam rate of 120 or 180 Hz. The steering loops control the average trajectory of the bunch train rather than individual bunches, so the Q BPMs are used, which measure the average of the train. Dipole corrector magnets are used, for which the feedback controls the magnetic field by setting a DAC (digital to analog converter) which alters the current from a power supply. Correctors and other actuators need to respond (to make 90% of a requested change) in a single 180-Hz period.

For the NLC, linac feedbacks will use typically use 10 BPMs and four correctors (two horizontal and two vertical). Communications and CPU must support processing of this local data at the 180-Hz rate. In addition, a cascade correction system for the NLC will require 180-Hz communications over long distances between linac loops (in diagnostic sections).

14.2.2 Synchronized measurement

The collider will be a pulsed machine with a relatively low repetition rate. A linear collider does not have the tendency to stabilize inherent in a storage ring machine: every pulse is a new beam. Therefore it is crucial that the facility exist to take measurements synchronized to particular pulses throughout the accelerator. These will include beam measurments, hardware diagnostics, and state information from the detector. This will allow pulse-to-pulse correlations of all different aspects of the collider and is the only way to efficiently trace sources of beam jitter and subtle hardware failures. A full sample – data for every pulse – over an extended time may be needed to see both high and low frequency instabilities.

Pulse-oriented sampling is also needed to provide an estimate of electromagnetic fields and other analog monitors at the beam pulse frequency. This means that the acquisition of analog signal data should be synchronized with the beam or line frequency.

14.2.3 Data acquisition and processing

High data-acquisition bandwidth is required to characterize pulses for optimization or feedback. Ideally, the BPM system should produce phase space centroid information of each bunch in the train at the full 180-Hz rate. This will require a control system data acquisition that has a throughput equivalent to the maximum beam pulse rate. The system should be able to acquire and process data more or less indefinitely. With a large number of samples acquired at the pulse rate, fine details of machine performance can be examined such as, for example, the frequency structure of narrow excitation lines. (Such lines have been observed in the SLC beam motion frequency spectrum, and are sometimes caused by the linac water-cooling pumps).

A more serious challenge is the measurement of phase space volume or emittance. In order to measure the bunch train's transverse and longitudinal volume with the sampling mode currently in use at SLC, a minimum of several hundred pulses is required. At low repetition rate (as when trying to diagnose and correct a problem) this is too slow to allow for effective tune up. A profile monitor system has been proposed that can be used to "fast" scan the entire train in a single pulse.

14.2.4 Special architecture for damping ring applications

At SLC, most of the data acquisition system and scheduling system is optimized for a pulsed linac machine. A special architecture is required that can be used to track the performance of the damping ring hardware throughout a single storage cycle and can track the behavior of the beam at the same time. The B-factory control system, for example, will have to have an integrated timing system that handles both the injected and stored beam. NLC damping rings will need controls that address the problems of injection, extraction and damping.

14.2.5 Tuning

Complex tuning will be required in the NLC. This is true in many places but is most prevalent in the final focus. To limit the luminosity loss due to the tuning procedures, the tuning scans and measurements must be highly automated and fast. Thus, the entire scan, which usually involves changing the strength of many magnets and then measuring the IP spot sizes with a beam-beam deflection scan, would probably be performed locally before the data can be shipped back and processed.

14.2.6 Modeling and simulation

An early decision in the design of the SLC control system was to base many aspects of machine setup, operation, and diagnostics on online accelerator models rather than a look and adjust method of interaction. The adoption of this approach has resulted in the development of a rich suite of applications that forms the foundation for near-automated operation of SLC. This modeling framework has been effectively used throughout the machine life cycle including commissioning, routine operation, diagnostics, and performance upgrade and optimization phases.

Given the complexity, strict tolerances, and the expansiveness of NLC, it is vitally important to have an online modeling environment with appropriate degrees of sophistication to facilitate machine commissioning and operation. At SLC offline simulation has been used to better understand machine behavior and to investigate alternative strategies; this trend should continue at NLC.

General areas of application for modeling and simualtion would include:

- Machine commissioning where the objective is to reconcile the model with the accelerator.
- Routine machine setup based on the design models and specifications.
- Model driven feedback system as described above.
- Near-automated diagnostic capabilities such as emittance measurement, and lattice diagnostics to quickly identify the sources of machine performance degradation.
- Model-based optimization tools such as orbit and lattice properties correction applications to allow rapid fixes for performance degradation.
- Creation of "multiknobs" which allow one to vary a parameter which may depend on many hardware values in a linear or nonlinear manner.

- Ability to use the same model base for online as well as offline machine physics studies.

14.2.7 Accelerator/Detector Coupling

The interaction region is precious not only to the physicists taking data with a detector but also to those running the accelerator. Certain diagnostic information can only be obtained there, but instrumentation is likely to impinge on the detector volume. The situation can be ameliorated somewhat by close coordination between shift personnel in the Accelerator Control Room and the active Detector Control Room as well as direct communication between the control system and the detector data acquisition.

At NLC provision should be made for at least the following:

- Shared timing. The Detector Acquisition must be synchronized with the bunch train. It should also have access to the configuration of the train (number of bunches, spacing). In order to make correlations offline each bunch train should have a unique identifier available to both the Control System and the Detector Acquisition.
- Interlocks. Detector components known to be vulnerable must be in a protected state during potentially damaging accelerator tuning.
- Tuning information. The Detector Acquisition should make available to the Control System a suitable collection of background signals at bunch-train rate during periods of good luminosity or fine tuning. This information is needed both for online tuning and for offline analysis.
- Veto information. The Control System should promote knowledge of anomalous pulses to the Detector Acquisition front-end.
- Polarization state. Both the Control System and the Detector Acquisition need access to this.
- Slow updates. Various quantities – state information, statistics – kept on one side may be of interest to the other.

14.2.8 Reliability and Availability

The availability of Control and Instrumentation Systems in existing accelerators (FNAL, SLC, CESR) has achieved the availability numbers required for the NLC (98-99%). The NLC is 10 times larger, and its construction comes 10 years (or more) after the existing machines. From the reliability point of view, the fact that the NLC is larger is the engineering challenge; the fact that its construction occurs a decade later is the engineering opportunity.

The NLC Control System will use components from the computer, communications and electronics industries. These industries have a long-standing record of improving their product reliability and availability in response to market needs. Thus, one can expect that industry will supply improved component reliability which will, in turn, get us part way to the achievement of our availability goals.

It would be wonderful if we could leave it up to the marketplace to solve this problem. Unfortunately, we cannot. We have already seen in existing accelerators that availability is heavily impacted by local design and operating decisions. Spares availability, and the scheduling of time for checkout and maintenance debugging

are examples of local operating decisions. The use of redundancy in designs, and the design of equipment in ways to decrease the MTTR (Mean Time To Repair), such as hot swapping of spares, are examples of local design decisions.

We have not yet expended engineering effort in putting together a design and operational strategy to achieve the availability goals of the control system. However, the goals appear to be realistic in terms of the experience of presently existing accelerators. The technologies involved for the control system are well known and tested. There are many practitioners in the fields of reliability, availability, and maintenance of large complicated systems. In the fields of electronics, communications and computers, this is a well-established engineering discipline. We know that effort in this area has to start early in the life of a project, since reliability goals must be set and given to designers early in the design phase.

14.3 Architectural implications

To zeroeth approximation, NLC is just a large accelerator requiring a large control system. To this extent NLC can use something like the control system "standard model" prevalent in large modern control systems. But it is also clear that the functional requirements discussed in the previous section can only be satisfied by making substantial perturbations to this model. These fall into the following categories:

Network bandwidth NLC's heavy dependence on feedback will make significant demands on realtime per-pulse bandwidth between device-controlling computers (called "micros" at SLC). There is also a need for real-time communication in order to synchronize measurements. Finally, to employ online modeling and provide information for simulation, large amounts of data must flow from the micros to an arena accessible to online servers and perhaps to a logger. That is, the control system will have a substantial data acquisition component.

Computational resources As pointed out in Section 14.2.1 a single fast feedback loop at SLC consumes the better part a micro. By contrast, less than 100 micros (mostly running at a lower clock rate) are used for standard device monitoring and control of the thousands of devices comprising the control system. The lesson for NLC, with its anticipated heavy use of feedback, is clear. Online modeling will be another cpu-intensive activity.

Interface with detector data acquisition In order to support the communication outlined above in Section 14.2.7, well-defined interfaces (hardware and software) and adequate network bandwidth between the NLC control system and the detector data acquisition will be needed. This kind of communication has proven essential for SLC/SLD (for example, parts of the detector readout are of great value as accelerator diagnostics). but was absent from the original design. The ad hoc methods currently employed lack flexibility and the amount of information which can be transmitted at 120 hertz is pitifully small.

Hierarchical software organization The cost of the control system for the SLC was much larger in proportion to total project cost than that of any other machine yet built. Already 200 to 300 person-years of software resources have been invested and further improvements are planned. This software development effort has been primarily an intellectual challenge as opposed to a bookkeeping or task management exercise.

In many ways, the pace of evolution of the control system has been limited by the time required for the machine physicists to identify and understand a problem and specify a solution.

To appreciate the need for hierarchical organization within the NLC control system it is instructive to look at SLC as it has evolved. SLC's database is in many respects the heart of its control system. During the first years of its existence (1980–1986), approximately 80 kinds of objects (with typically many instances of each type) were defined in it. All but a handful describe hardware. Since that time about another 90 kinds of objects have been added to the database, but of these well over half describe something bearing little resemblance to a physical device, for example feedback loops or model parameters. The database, like the system and application software which access it, is both substantially larger and different in character from initial expectations.

This evolution at SLC was only possible because of the generic nature of the treatment of the lowest level components. Strict rules are followed in all low-level device interfaces, and this structure facilitates continued growth through the application of higher and higher layers. This implies that the design of the lower levels is critical for a control system which must support a complex variety of high- and very high-level applications.

With its tight tolerances which must be actively maintained, and opportunities for subtle interactions among its elements, NLC will continue this trend. In order to get their job done efficiently all who use the control system (operators, machine physicists, hardware maintenance personnel, programmers) must be able to access it at an appropriate level of abstraction so they can control, monitor and analyze the entities of interest to them.

14.4 The Control System Model

As noted in the Introduction, we will describe a control system implementation model that is, we believe, a realistic model for an NLC Control System, if it were being built today. It is realistic because it follows the main thread of control system design used in accelerators recently built (CEBAF, ALS) or under construction (PEP-II). However, we will state the standard disclaimer that this is a design based on current technology, and that the real NLC control system design will be based on the technology available at the time of its construction.

This model is based on what has become a fairly standard approach in modern computer control systems (Figure 14-1); it comprises a set of consoles and servers linked to each other and to a hierarchically lower set of front-end computers via high speed networks. These front-end computers are, in turn linked to devices via dedicated control networks. The devices themselves are expected to often include embedded computers; thus the front-end computers are really communicating with still another architecturally-lower layer of device computers. In the case of low to medium multiplicity devices (1 to 1000 devices), there may be a device interface crate (VME, VXI) which contains a module which controls the device. We expect that the truly high multiplicity devices (Beam Position Monitors, Magnet Movers; greater than 5000 devices) will probably have a cost-optimized design which includes dedicated embedded computers, and communicates to the front-end computers via a digital network link of some kind.

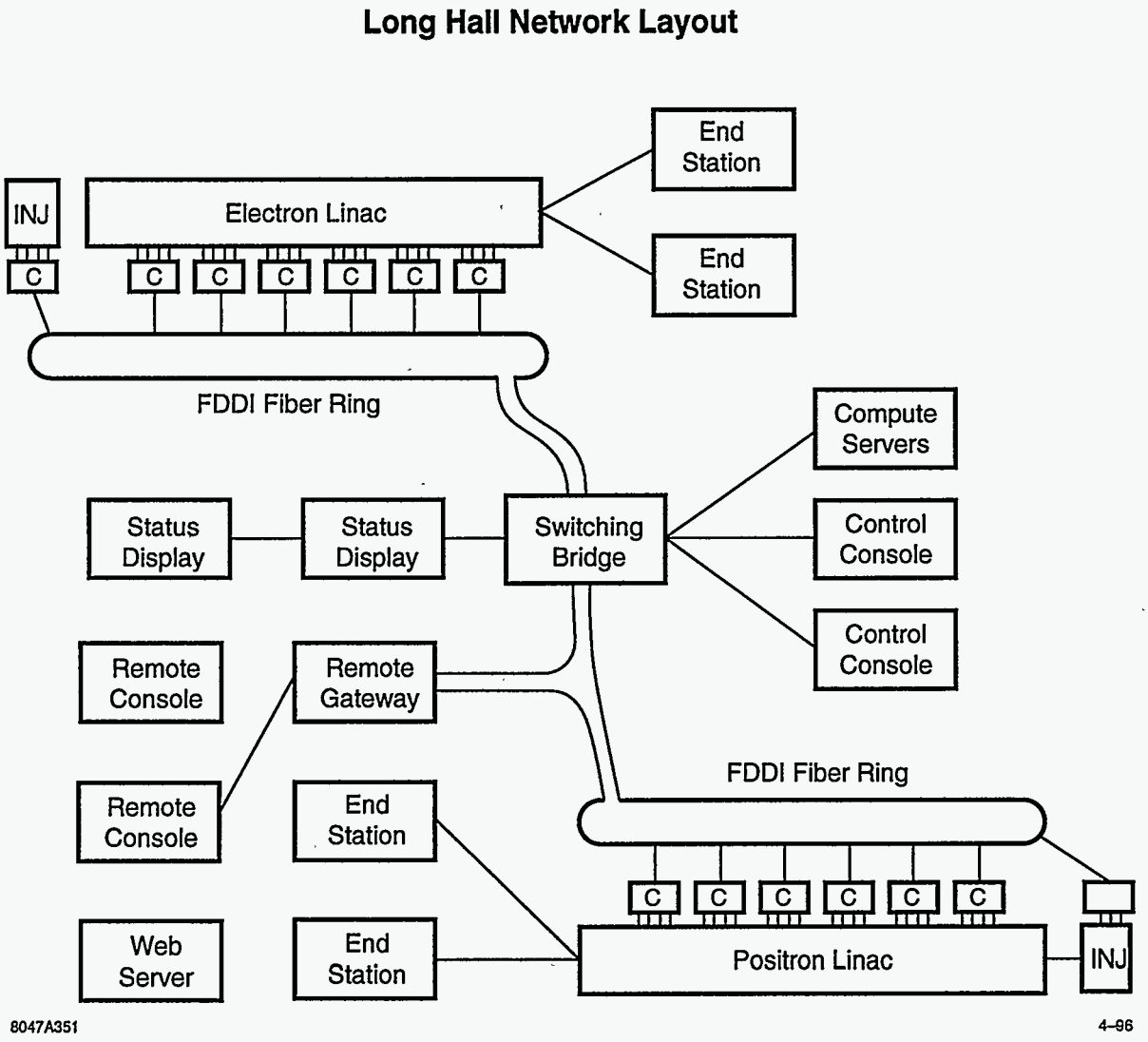


Figure 14-1. *Schematic of the control system layout.*

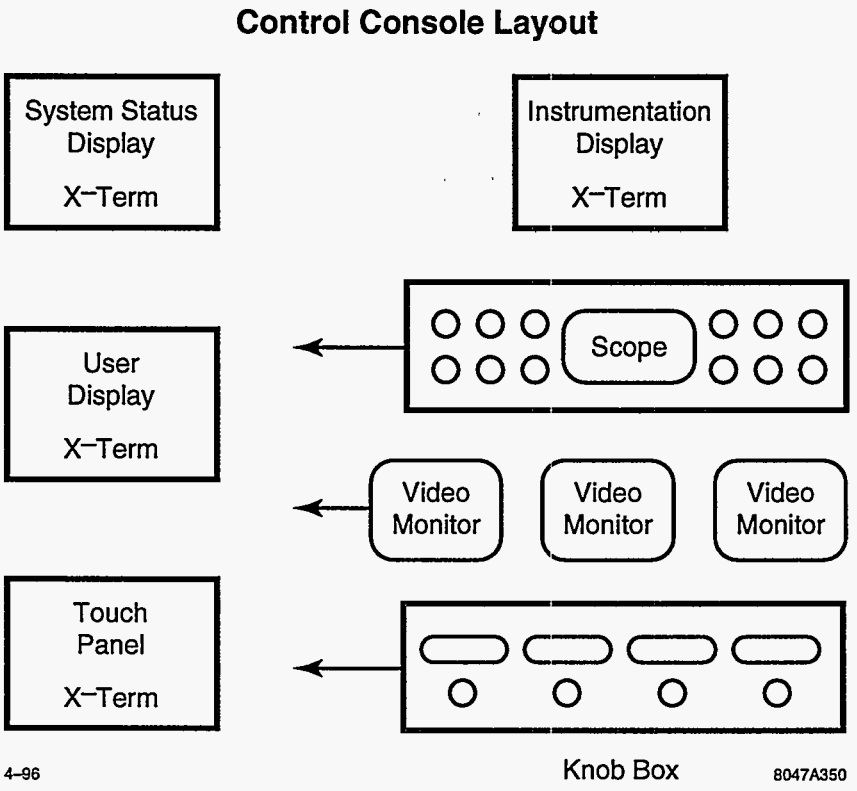


Figure 14-2. *Schematic of operator's console.*

14.4.1 Operator Consoles

Console Computers will be located in the Control Room and will be the Primary Operator Interface (OPI) into the Control System. There will be a Graphical User Interface (GUI) driven by the EPICS Control System based on the X-protocol. The physical hardware will be comprised of a processor device driving (perhaps) two video heads with some pointing capability plus a standard keyboard (Figure 14-2). These workstation processors will have large internal memory and internal hard disks.

These machines are modeled as DEC Alpha Processors running the NT Operating System. These machines are capable of high throughput, excellent number crunching, and a fast network response.

Consoles will use the TCP/IP over Ethernets which are in turn connected to the FDDI backbones via switching bridges. Separate local networks will be used to subdivide the overhead status display screens and the pairs of screens associated with each console such that network or server failures will not bring down the control room facilities.

14.4.2 Application and Server Computing Resources

The control system model is a distributed system, and one component of that distributed system will be a set of around 20 computers which will function as servers of various functions. The functions we have in mind are database, network interface, computational, and file servers.

Database servers will include the Control System Database and the Documentation Database (drawings of devices and cable-plant, wire lists, system documentation, operations guidelines, code management, etc.). Network interface servers will handle connections and security issues for computer communication access outside the control system. Computational servers will supply sufficient computing power for accelerator modeling and simulations. File servers will provide storage for a number of functions, including software and firmware development, the archiving of data from the control system of device histories, operation histories, configurations, alarm message and error reporting, computer system and network management monitoring data, etc. Many of these functions will be accessible to the detector data acquisition; conversely, information from the detector will be accessible to control system applications.

Associated with these system will be large disc farms for storage of the large amounts of data associated with such an accelerator complex. Some form of redundancy such as clustering will allow the server complex to gracefully degrade in the event of individual server or other component failure. We expect that this server cluster will be in relatively close proximity to the control room, so that local very high speed links connecting to the consoles in the control room may be utilized.

Network Security Systems

There are requirements to safeguard the Accelerator from remote access and control by unauthorized individuals. Security systems will be implemented at the onset of the program which allow data to be freely available over external networks, while preventing unauthorized remote users from operating equipment. These security systems will provide encrypted remote sessions such that passwords and private data cannot be swept for unauthorized use.

14.4.3 High Speed Networks (FDDI & Ethernet)

This Distributed Control System depends heavily on networks with substantial bandwidth to function properly. Thus it is critical to use network facilities which can handle anticipated loads at low utilization figures. These networks need to be industry standard facilities as well.

Accelerator Control Networks will be subdivided into subnets to isolate the operation of the machine from another networks activities at the laboratory. There will be a control network Firewall router between general laboratory networks and the control networks to allow isolation if required.

In terms of current technology , FDDI and Ethernet make good sense because they are industry standards, they are versatile, and their performance is well understood. At the time of the final design review, it will be necessary to review these specific selections and to use the then current stable dependable technologies.

There are some special considerations for the networks, in that the length of the network runs will generate challenges for any network technology. Fiber implementations will be important, and microwave facilities may be an option to consider.

14.4.4 Front-end Computers

Front-end computers are the Input/Output Controllers (IOC) which actually perform the data acquisition and control in real-time in the accelerator housing. These machines are microprocessor based platform using PCI, and VXI bus backplanes. Each device will have its own Real-time Executive, and will run a standard set of acquisition/control software. Database configurations will be downloaded in order to structure the number and type of devices and processes the controller handles.

These processor crates will normally be operated remotely over the Ethernet from the Control Room, but can be operated locally from a terminal for diagnostic purposes. Long-haul communications will be handled by the FDDI backbones with bridges to distribute Ethernet connections to micro crates, local control X-terms, and other Ethernet devices.

14.4.5 Data Acquisition Crates

Data Acquisition Crates will be a mixture of VME and VXI crates containing the analog and digital input/output cards used to make measurements and define set points. Low-cost VME crates will be used for most of the modules, along with some industrial equipment for the low accuracy systems. VXI will be used for the precision measurements and high-frequency rf and timing modules.

14.4.6 Instrumentation Modules (VXI & GPIB)

There will be a number of modules for high frequency monitoring and control which will emanate from VXI and GPIB controlled instruments. The VXI crates will be controlled over Ethernet through the controller module in the crate. These crates can control locally positioned GPIB instruments as well.

14.4.7 Radio Frequency Control, Phasing & Feedback

The rf requirements are covered in Chapter 8; we will not repeat those here.

14.4.8 Dedicated Control Networks

These networks are dedicated point-to-point networks to connect specific pieces of equipment. This implementation is in use at SLC has been chosen as perhaps the only one available today to get high performance and dedicated functionality.

These links will be found between control crates and large power supplies with self-contained controllers. There will also be a link or links between the control system and the detector data acquisition in order to transmit, for example, tuning information (detector to accelerator) and beam quality information (accelerator to detector). There will be point-to-point links between machine stabilizing feedbacks and feed forwards.

This approach (point-to-point links) has serious drawbacks for feedback: it does not scale well and lacks flexibility. Alternatives should be investigated as they become technologically feasible.

Special 1553 links have been employed between elements of the SLC machine protection system for security and speed.

14.4.9 Timing and Beam Rate Control

The purpose of the scheduling system is to provide control of the paths to be followed by beam pulses more or less in real time. Several pulsed beam dumpers and diagnostic stations will be installed throughout the NLC; before and after the damping rings, in the injector and positron system, after the bunch compressor, at several places throughout the linac and on either side of the big bend and collimation systems. The scheduling system will control the firing of these dumper magnets and synchronize the data acquisition on the dumped pulses. It will also be used for the machine protection system (see Chapter 16).

The scheduling system is modeled on the one currently in use at SLC. It will be used to control the injection, extraction and storage time in each of the three rings. Using it, the control system will be able to program the storage time, and therefore the output emittance, of each of the three rings. This feature will be built into the system through a linking, or pointer-based, structure that will be used to track the progress of a bunch or train of bunches from it's inception through to a dump. Another requirement of the scheduling system is to provide control of synchronized data acquisition and sequencing of pulses when a fast pulsed device, such as a kicker magnet or a pulsed phase shifter, is being adjusted. Such acquisitions will be used to quickly scan a beam across a beam size monitor and for certain classes of beam optimization. They are also required for optimizations which involve measuring derivatives such as maximizations or minimizations or maintaining a specified phase with respect to the rf. The technique to be used is a synchronous detection scheme with a sub-tolerance dither.

Accelerator Beam Rate will be controlled by a dedicated special purpose processor running custom software to support flexible rate control in both accelerators. This Master Pattern Generator will be wired into the machine protection systems to handle both rate limiting and machine protection shutdown. Rate information will be disseminated to appropriate components of the detector data acquisition as well as as to front-end computers, etc., within the control system.

Flexible rate control allows controlling the amount of beam energy transported throughout the machine. This will facilitate the alteration of beam parameters for experiments which require unique beam characteristics or timing.

This flexible beam rate control also allows the accelerator to be rate limited by classes of machine protection problems so that problems can be located and identified with low rate beams with reduced risk of accelerator or equipment damage. Beams will automatically rate limit back to designated rates as problems are resolved. Rate control will be exercised on a pulse-to-pulse basis.

14.4.10 Machine Protection Systems

The purpose of the machine protection system (MPS), which is described in Chapter 16, is to prevent damage to machine system components in the event of a routine failure. The system is not intended to provide comprehensive protection against any possible failure. One of its main functions is to automatically

provide a sequence of beam pulses that can be used as effective diagnostic tools during a startup or fault period.

The machine protection system has four logical layers: 1) mechanical, 2) device controllers, 3) power monitoring and 4) beam scheduling and control. The layers provide a graded approach that allows the production of a nominal intensity single bunch beam for diagnostic purposes. All mechanical systems should be capable of withstanding the impact of a single pulse of such a beam. They should also be able to survive another strike in the same location. It may not be possible to develop structures that can stand a single pulse strike from a nominal single bunch at nominal emittance, especially for the higher intensity versions of this design. In this case, an emittance enlarging system must be integrated with the MPS so that proper operation is ensured. Once operation is checked with single bunch beams, the repetition rate may be increased, the emittance brought to nominal and the number of bunches brought to its full value. This sequence must be applied in this order. The only viable way to transport full-intensity beams is to make sure that the transverse deflecting forces acting on the full-power beam cannot change enough in the interval between pulses to target the beam cleanly on a beamline element. One consequence of this is that low-repetition rate, full-train intensity operation, is not possible.

14.4.11 Equipment & Tunnel Access Control

The control system will monitor the status of equipment and the state of Accelerator access, but the actual control of Personal Protection, Machine, protection, and Hazards will be handled by dedicated hardware and Programmable Logic Controllers.

14.4.12 Application Software

Application packages will reside and run on the console processors and on a separate applications processor in the control cluster. This extra applications processor will take on large resource computing loads that would not run well on console machines. Other applications or analysis programs will be off-loaded to user machines via self-describing data files.

Applications packages will include Accelerator diagnostic packages, measurement packages for things like emittance and chromaticity, energy management (LEM), power steering with machine optics models, correlation plots, simulations, machine models, multiknob control, and data archiving.

Measurement packages may be operated remotely at lower priority than control room activities. Analysis of data may be run on remote hosts or the application processor in the cluster. Operation of applications which operate accelerator equipment or change machine configurations or settings will execute from the control room only.

The applications environment will be structured to enhance the ability of the Laboratory to use software developed at other Laboratories or purchased commercially.

Included applications: archiving, correlation plots, steering, LEM, models, emittance, logging.

14.4.13 Software Application Bus

Applications which run on the Console or Cluster Computers will run on a software layer which will isolate them from the complexities of where data comes from, how it is stored, and how it is transported. Hidden facilities will deal with correlation plot data which has to be correlated in time and take into account measurements taken in different parts of the accelerator.

This software structure will make available common measurement and data collection facilities which may be required by application or display processes. Data will be presented and exported in self-describing formats compatible with application packages available at other laboratories [Watson 1995].

Similarly, data files will also be available in Matlab format for local or remote analysis.

14.4.14 Software Development Environment

Software development will be accomplished on workstations similar to those used in the production environment, however, they will be configured to run a parallel but separate control environment so that actual production equipment will not be controlled by accident. With the exception of the separate environments, the software environment will be identical to the production environment. Some special hardware will be developed which will help simulate accelerator operation for software evaluation and testing. Additional software will be required to compile code, to control progressive versions of software, and database facilities to build run-time databases and configure equipment.

Diagnostic systems will include remote diagnostic and debug capability for all networked microprocessor systems (not including embedded systems). Isolated or low device count GPIB instruments will be controlled via Ethernet by GPIB network control boxes placed in the locus of the GPIB instrumentation.

14.4.15 Industry Standards

The Control System will utilize as much commercial equipment and software as practicable. Industry standard equipment and facilities will be utilized to reduce cost and improve maintenance and reliability. To the extent possible, the control system will utilize electronic modules available commercially and utilized in other laboratories to reduce cost and resources involved in hardware development and in writing low-level software drivers.

References

[Humphrey 1992] R. Humphrey, "Lessons From the SLC for Future LC Control Systems," *Proc. of the 1992 Int. Conf. on Acc. and Large Exp. Phys. Control Systems*, KEK, Tsukuba, Japan, KEK Proceedings 92-15 (1992).

[Watson 1995] C. Watson *et al.*, "cdev, a Common Device API," *Proc. of the 1995 Int. Conf. on Acc. and Large Exp. Phys. Control Systems*, Fermilab, Batavia, IL (1995).

Contributors

- Joanne Bogart
- Spencer Clark
- Rusty Humphrey
- Nan Phinney
- Marc Ross
- Hamid Shoaee

15

Instrumentation

Instrumentation performs a critical role in the operation of a linear collider. New acquisition and data processing techniques are required for feedback, tuning procedures, and performance monitoring. For example, many collider systems are initially tuned using complex bootstrap procedures whose convergence rate will depend on the speed and performance of several instrumentation systems. Furthermore, mechanical and electrical tolerances are computed assuming the success of this process.

The next leap in electron-positron accelerator performance will result in part from improvements in instrumentation technology. The latest generation of accelerators, from high-current synchrotron light machines to B-Factories and linear colliders require feedback control loops that are greater both in number and complexity than more conventional machines. As a result, the instrument is no longer a diagnostic tool, intended for use only in cases of sub-standard performance, but a truly integrated accelerator component. This has obvious implications for the instrumentation-system designer, among which is that the system must have the integrity required of other accelerator systems, such as the power conversion and vacuum systems.

Linear colliders represent the most extreme application of this philosophy. The lack of closed, equilibrium conditions that maintain stability in the machine, forces the use of several layers of sophisticated feedback loops. The underlying reason for this requirement is the tolerances that must be applied for the correct transport of low-emittance beams. In some extreme cases, initial bootstrap procedures are required before any beam can be transported through the system. Tight mechanical and rf system tolerances will not only require special systems to address them directly, but will also demand beam-based feedback and tuning procedures. For example, in the X-band linacs and the beam-delivery sections, the magnet alignment is continuously monitored and adjusted using beam-based techniques that rely on high-resolution Beam Position Monitors (BPMs).

Perhaps the most important improvements in instrumentation technology will not come from the harnessing of fundamentally new physical processes to better the performance of beam position or size monitors. Instead, they will come from the integration of existing instrument beam sensors with more powerful controls. Very strong integration with the control system is needed to provide the robust, high data-processing bandwidth needed for higher level control.

An important aspect of the shift in the role of instrumentation will be its use in general optimization systems that will ultimately change the character of the control room operator's task. Traditional applications of instrumentation systems in colliding-beam accelerators have required heavy involvement of the operator. In storage rings, for example, operator technique in optimizing injection and luminosity has proven to be a key factor in long-term performance. In a heavily feedback- and optimization-control-laden system, the operator's task becomes the more complex one of controlling and monitoring the performance of these automated tasks.

Details of the instrumentation design and requirements are distributed through the preceding chapters of this document. Many of the concepts needed for the high-resolution systems have already been tested. For example, the Final Focus Test Beam (FFTB) at SLAC utilizes stripline BPMs with 1-μm resolutions and a beam size monitor that is capable of measuring 40 nm spot sizes. In addition, rf BPMs were installed and measured to have a resolution less than the required 100 nm. Other elements will be tested in the near

future. This includes the rf structure BPMs that are needed to align the accelerator structures, a laser wire system similar to those needed to measure the beam emittances in the linacs and final foci, and the PEP-II button BPM system that is similar to those needed in the damping rings.

16

Machine Protection Systems

Contents

16.1 Introduction **890**

16.2 Single Pulse Induced Failure **890**

16.2.1 Diagnostic Pulse Protection 892

16.2.2 Transition between diagnostic pulses and full beams 897

16.2.3 Controlling the Interpulse Difference (MAID) 899

16.1 Introduction

One of the most serious operation issues that any future linear collider will face is that of the Machine Protection System (MPS). To produce useful luminosity the beam power and the beam densities must be very high. Unfortunately, these beams will almost certainly damage any material that is intercepted unless extreme care, such as that in the collimation sections, is taken. For example, in the 1 TeV NLC design, the beam power is over 8 MW and a single errant bunch train in the linacs would be sufficient to damage many unprotected accelerator structures. Obviously, this has severe implications on the beam operation during normal running as well as during tuning and commissioning.

The purpose of the MPS is to prevent damage to the collider components in the event of a routine failure or mistake. In addition, it should automatically provide a sequence of beam pulses that can be used as effective diagnostic tools during a startup or a fault period. Furthermore, after a fault, the MPS should be optimized to recover luminosity as quickly as possible in order to minimize the lost time. Of course, the system is *not* intended to provide comprehensive protection against any possible failure; the complexity of the MPS must be balanced against the cost, difficulty, and time for repair of the systems it protects.

The most serious challenge in the NLC is the prevention of 'single pulse induced failure' (SPIF). This is component failure that occurs from an aberrant single beam pulse. Because it is impossible to know the precise trajectory of the upcoming pulse, the MPS must: 1) provide pulses that cannot cause SPIF for tuning and diagnostics and 2) insure that the difference between the upcoming pulse and the one that preceded it is within some limit, known as the 'maximum allowable interpulse difference' (MAID), during normal operation. These two criteria form the basis of the NLC single pulse induced failure machine protection system.

Multi-pulse or 'average power' induced failure is component failure that occurs after a succession of pulses deposit excessive energy on a given component. This type of failure is more familiar from SLC operation and is controlled in a similar fashion, *i.e.,* by using ion chambers, thermocouples, etc., to monitor beam loss.

In the next sections, we will outline the methods that are used to protect against SPIF—as noted the multi-pulse failure mode is protected with a more standard MPS and thus will not be discussed further. At this time, we only have a conceptual description; in the future, we will need to have a detailed solution on an element-by-element basis with greater margins than is outlined in this section. Furthermore, we have only considered protection in the main linac and downstream; we have not considered the MPS issues in the damping rings or bunch compressors although the principles will undoubtedly be similar. Finally, additional detail on the MPS can be found in Sections 7.8, 8.6, 8.7, and 9.2.2.

16.2 Single Pulse Induced Failure

The strategy to protect against SPIF is to use single bunch, nominal intensity, pulses for most diagnostic purposes and at any time the interpulse difference might be outside the MAID. This strategy results from the extreme energy density of the full intensity, multibunch, NLC beam. It is not practical to build mechanical systems that can withstand the nominal NLC beam except in isolated cases such as the collimation region. It is, however, possible to develop structures that can withstand the impact of single bunches, albeit with somewhat increased emittance. Once such structures are realized, the problem becomes one of ensuring that successive pulses are alike. Thus, the single-pulse protection (SPIF) can be subdivided into protection against

a diagnostic bunch, a full current, high emittance single-bunch beam, protection against a multibunch bunch train, and the method of transitioning between the diagnostic and normal operating modes.

The protection against the diagnostic bunches is based upon a passive system consisting of thin spoilers which will increase the beam angular divergence so that, by the time the diagnostic beam strikes another component, it will not cause any damage; to prevent damage to the spoilers themselves, the emittances of the diagnostic beam must be increased by a factor of ten from the nominal beam emittance. Thus, the spoilers will allow the beamline to survive the transport of a single diagnostic bunch without regard to the state of the beamline hardware. All preliminary beam-based alignment, tuning, and diagnostics will be performed using a diagnostic bunch. Of course, if beam is being lost during the transport, these operations would need to be performed at low repetition rates to reduce the average power deposition.

After a diagnostic beam can be transported to the beam dumps without difficulty, *i.e.*, after establishing the initial beam-based alignment, the beam trajectory, feedback setpoints, and energy profile, the repetition rate can be increased to the nominal 120–180 Hz. At this point, the beam emittance can be decreased to nominal and additional bunches can be added to the bunch trains. To verify the beam loading compensation is set properly, the train length will be increased in steps. All subsequent tuning must be performed at the high repetition rate; only the diagnostic beam can be transported at low repetition rate.

To prevent the high rate beams from striking accelerator components, a trajectory window of roughly $\pm 200\,\mu$m and an energy window of $\pm 10\%$ are established about the nominal values; this is the MAID. If the beam deviates beyond these limits in any single pulse, the collider is returned to the high-emittance single-bunch diagnostic mode while the source of the problem is diagnosed from data that was taken during the errant pulse. Extensive logic will be used to prevent erroneous MPS faults due to bad BPMs readings.

This system relies on the fact that there are no transverse deflecting fields that can change sufficiently, within a single interpulse period (roughly 8 ms), to deflect the beams from the operational trajectory window into the accelerator structures or beamline elements. In most cases, this is attained by limiting the strength of all fast correctors and limiting the decay time of the quadrupole and bending magnet fields by using solid core magnets and thick conducting vacuum chambers. The few DC magnets whose fields could change too quickly will have to be interlocked directly to the MPS using a Hall probe or similar diagnostic—separate diagnostics, such as precharge monitor, will be needed for strong pulsed kicker magnets like the damping ring kickers.

In addition, the MPS must verify that the rf systems are operational and correctly phased before the beams are launched into the linacs. However, fairly large energy deviations can be tolerated. For example, a 20% energy deviation in combination with 100-μm random quadrupole misalignments, which are well in excess of what we expect, would only cause 1-mm orbit offsets in the linac.

Thus, the MPS system must only verify that 85% of the rf systems are operational and correctly phased. To this end, all modulators will be polled roughly 100-μs before beam time. At the same time, the klystron timing and phase information will be checked. If there are a sufficient number of failures, the beams will be aborted downstream of the damping rings; this verification procedure is described in greater detail in Chapter 8. Finally, there will be sacrificial spoilers and dumps located in the linac diagnostic station chicanes to prevent energy errors larger than 25% from propagating further down the linac; these are described in Chapter 7.

In the next section, we will describe the mechanical protection system and the diagnostic pulse generation. After this, we will discuss issues associated with the transition from the single diagnostic bunch mode to the full bunch train and then we will describe controlling the MAID.

	Be	*C*	*Al*	*Ti*	*Cu*	*Fe*
Radiation Length [cm]	35.7	21.7	9.0	3.7	1.4	1.8
dE/dx_{min} [MeV/cm]	3.1	3.6	4.4	7.2	12.8	11.6
Specific Heat [J/(cm^3 °C)]	3.3	1.9	2.5	2.4	3.5	3.8
Melting Point [°C]	1280	3600	660	1800	1080	1530
Stress Limit [°C]	150	2500	140	770	180	135
ΔT [°C]	3240	6530	6060	10330	12100	10520
$\Delta T/T_{melt}$	2.5	1.8	9.2	5.7	11.2	6.9

Table 16-1. *Spoiler material properties and temperature rise due to a single bunch of* 1.3×10^{10} *having emittances of* $\gamma\epsilon_x = 3 \times 10^{-6}$ *m-rad and* $\gamma\epsilon_y = 3 \times 10^{-8}$ *m-rad at end of the 500 GeV linacs; these emittances are the smallest values possible and correspond to an rms beam size of* $\sqrt{\sigma_x \sigma_y} = 3.1\ \mu m$.

16.2.1 Diagnostic Pulse Protection

Spoiler Materials

The purpose of the spoiler system is to allow a nominal intensity, single bunch, beam to be transported throughout all systems at low repetition rate without concern for damage. As stated, this will be performed using thin spoilers. If the beam is steered sufficiently far off-axis to intercept an accelerator element, it must first pass through one or more spoilers. These will increase the beam angular divergence so that, by the time the single bunch beam strikes another component, it will not cause any damage.

The length of the spoilers is a trade-off between the increase in the multiple scattering and the heating due to the electromagnetic shower in the spoiler. The projected angular distribution of the beam after a spoiler can be described by a gaussian distribution [Particle Data Book]

$$f(\theta_{x,y})d\theta_{x,y} = \frac{1}{\sqrt{2\pi}\theta_0}\exp^{-\theta_{x,y}^2/2\theta_0^2} \qquad \theta_0 = \frac{13.6\,\mathrm{MeV}}{E}\sqrt{t}(1+0.038\ln(t)) \tag{16.1}$$

where t is the length of material in units of the radiation length. Thus, we expect a 0.2 radiation length spoiler to increase the beam angular divergence of a 500 GeV beam by 11.4μr; this is hundreds of times larger than the incoming angular divergence.

Table 16-1 lists properties for a number of different materials including the dE/dx_{min}, the specific heat at room temperature, the melting point, and the stress limit. The stress limit is based on the tensile strength, the modulus of elasticity, and the coefficient of thermal expansion for the material. When a beam strikes the material, there is a sudden local temperature rise that may create local thermal stresses. If the temperature rise exceeds the stress limit, micro-fractures can develop in the material. In addition, it has been observed in experiments, that if the local temperature rise exceeds four times the stress limit, the shock wave due to the thermal rise will cause the material at the surface to fail completely or "delaminate" [Walz 1973, Walz 1996].

In the spoilers, we are not actually concerned by micro-fractures or deformations that might develop when the temperature exceeds the stress limit. These will not degrade the performance of the spoilers and would likely be partly re-annealed with further heating. Thus, the allowed temperature rise is limited by either the melting point of the material or four times the stress limit at which point the material will fail catastrophically.

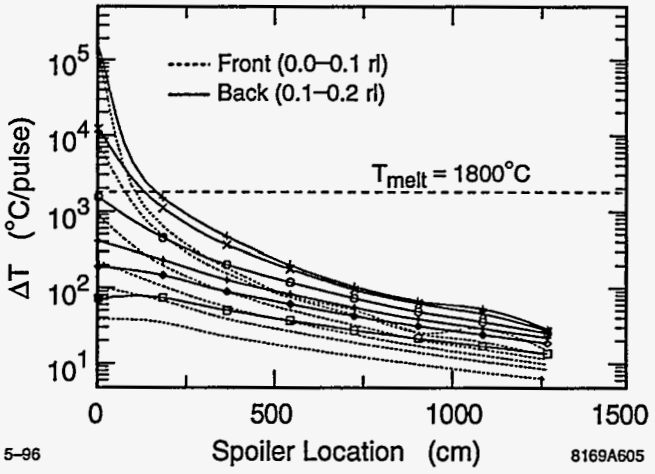


Figure 16-1. *EGS simulations of temperature rise in 0.2 radiation length long Ti spoilers with a single bunch of* 1.3×10^{10}. *The dotted and solid lines show the temperature rise in the front and back halves of the spoilers for six different incoming beam sizes:* $\sigma_r = 0, 3.1, 10, 20, 30,$ *and* $50\,\mu m$.

Table 16-1 also lists an estimate of the temperature rise ΔT assuming a single bunch of 1.31×10^{10} particles with a beam size of $\sqrt{\sigma_x \sigma_y} = 3.1\,\mu$m. This beam size corresponds to emittances of $\gamma\epsilon_x = 3 \times 10^{-6}$ m-rad and $\gamma\epsilon_y = 3 \times 10^{-8}$ m-rad at the end of the 500 GeV linac which are the smallest emittances that could be delivered from the damping rings and assumes that there are no emittance dilutions through the end of the linacs. The temperature rise is calculated from a simple analytic model:

$$\Delta T = \frac{N}{2\pi\sigma_x\sigma_y}\frac{dE/dx_{min}}{\text{Spec. Heat}} \tag{16.2}$$

which ignores the effect of the shower buildup and the variation of the specific heat with temperature.

In both the linacs and the beam delivery, the spoilers are constructed from titanium. Although other materials such as beryllium or graphite would be able to handle higher beam densities, titanium is the most practical choice, balancing the spoiler survival against the length of spoiler required and the vacuum and handling properties. Unfortunately, the surface temperature rise due to this low emittance beam is over five times the melting temperature of *Ti*. Thus, to prevent damage to the spoilers themselves, the emittance of the diagnostic beam must be increased significantly.

These analytic calculations have been supplemented with EGS simulations. In Figure 16-1, the maximum temperature rise is plotted in each of eight linac spoilers separated by 1.8-m for round gaussian beams having sizes of $\sigma_r = 0, 3.1, 10, 20, 30,$ and $50\,\mu$m; it is thought that the round beam case will model a flat beam with similar density although it may slightly overestimate the temperature rise. The dotted and solid lines show the temperature in the front and back halves of the spoilers and the dashed horizontal line shows the melting temperature of 1800°C. Notice that the temperature rise in the back half of the spoilers is roughly twice that in the front half due to the buildup of the electromagnetic shower.

Clearly, all spoilers downstream of the first will survive a single pulse of any incoming beam size. But, the first spoiler will be damaged unless the incident beam has an rms size greater than $\sqrt{\sigma_x\sigma_y} > 10\,\mu$m. Thus, to prevent damage to the spoilers, either both the horizontal and vertical emittances of the diagnostic beam must be increased by a factor of ten or the vertical emittance could be increases by a factor of 100 to yield the required beam size.

Spoiler Placement

In the linacs, the spoilers are primarily needed to protect the accelerator structures since these have the smallest aperture while, in the beam delivery, the spoilers are needed to protect the vacuum chamber and magnets. Unlike the spoilers, it is important to limit the temperature rise of the other accelerator elements to a value below the stress limit listed in Table 16-1. Deformations or micro-fractures in the accelerator structure irises would probably increase the multipactoring and the dark current from the structures and could distort the acceleration field patterns. Similarly, deformations of the magnet poles or coils could lead to large multipole fields or shorted coils.

We can get a first estimate of the requirements by looking at Table 16-1. Here, the temperature rise was calculated assuming a diagnostic pulse with an rms beam size of 3.1 μm. In this case, the temperature rise in Cu is estimated to be over 12000°C while the stress limit is 180°C. Thus, to decrease the temperature to a more reasonable value, the incident beam density must be decreased by a factor of 70. Actually, the requirements are much greater because of the electromagnetic shower. In practice, we need to decrease the beam density of a 500 GeV beam by over a factor of 3000.

Now, to determine the spoiler placement, we need to determine the failure scenarios. Ultimately, we will have to consider the failure modes on an element-by-element basis but at this time we will only consider three global scenarios:

1. First, as a "worst case scenario," we consider a FODO channel where a large deflection arises at the focusing quadrupole; a FODO array describes most of the linac beamline and much of the beam delivery beamline. Such a deflection could arise if one or more poles of the quadrupole magnet becomes shorted or if the magnet mover runs away to an extreme value, typically limited to ±1 mm. This deflection will offset the beam in the next defocusing magnet which then deflects it further and possibly into an accelerator element.
2. Most other scenarios such as large energy errors or smaller amplitude deflections will cause the beam to oscillate at large amplitudes before being lost into an element.
3. Finally, there are a number of special situations which are not covered by either of these two cases. In particular, failures of the bending magnets or some specialized strong quadrupoles could directly drive the beam into an element with a very large deflection angle.

Unfortunately, there is no way to provide passive protection against this third case. Thus, these components will have to be directly monitored and interlocked to the MPS. In contrast, the second case is straightforward to protect against by placing a number of spoilers along the beamline with apertures smaller than that of all other elements. This leaves the first case which is more difficult to passively protect than the second case because there is relatively little distance in which the beam size can be increased before it intercepts an accelerator element. Of course, we could adopt the solution used for the third case, *i.e.*, direct monitoring, but this is less desirable than a passive system because of the large number of elements that would have to be monitored and the potential reliability problems.

To evaluate the requirements to protect the first case, we can calculate the distance from the defocusing quadrupole to the position at which the beam strikes the accelerator structure or the vacuum chamber ΔL:

$$\Delta L = \frac{L_c}{2}\,\frac{(r_a/r_{QD}-1)}{1+2\sin\psi/2}\,. \tag{16.3}$$

Here, ψ_c is the phase advance per FODO cell, L_c is the FODO cell length, and r_a and r_{QD} are the aperture of the chamber/structures and the trajectory offset in the defocusing magnet.

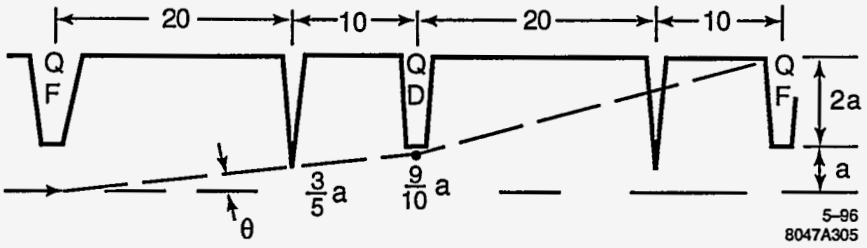


Figure 16-2. *Possible trajectories of a single bunch through a beam line as a result of completely or partially shorted quadrupole legs. Every trajectory encounters a spoiler at least 10 m before reaching the wall or another element. The quadrupoles are assumed to be spaced at 30 m and have a focal length of 20 m in this example.*

Assuming a phase advance of 100° per cell, we find that, to shadow all the downstream elements, spoilers located at the quadrupoles would have to have a radius roughly $\frac{1}{4}$ the minimum radius along the beamline. Although this may be the simplest solution, it implies very small aperture spoilers with correspondingly large wakefields; the transverse wakefield of a single spoiler would be comparable to that of an entire 1.8-m accelerator structure.

This becomes simpler in the beam delivery beamline where the vacuum chamber aperture can be increased to relatively large values between the magnets. In this case, the spoilers can also have relatively large radii which reduces the wakefields. Then, the placement of the spoilers must be chosen so they are located sufficiently far from the magnets, where the vacuum chamber constricts again, so that the spoiled beam will not damage the elements. This is illustrated schematically in Figure 16-2 where 0.25 radiation length Ti spoilers have been placed 10 m from the magnets to increase the beam size to roughly $\sqrt{\sigma_x \sigma_y} \sim 140\,\mu$m. The magnets and vacuum chamber will be further protected with sacrificial absorbers located immediately upstream of the chamber constriction. Both the spoiler placement and the absorbers are discussed further in Chapter 9.

The linac is more difficult to protect and the solution is different for the beginning and end of the linac. At the beginning, the low energy end, the quadrupoles are only separated by a few accelerator structures. In this case, a spoiler with an aperture radius of 1 mm, $\frac{1}{4}$ of the iris radii, would increase the total transverse wakefield seen be the beam an unacceptable amount; as stated, each spoiler would have a transverse wakefield comparable to that of an entire accelerator structure.

Fortunately, at low beam energy the angular divergence due to the spoilers is relatively large, θ_0 scales inversely with the beam energy, and the required spot size to prevent damage to the structures is relatively small. Thus, the spoilers can be placed close to the location that the beam first strikes a structure. This is illustrated in Figure 16-3 which shows the temperature rise in the irises of an accelerator structure due a 30 GeV single bunch with 1.3×10^{10}. The four different curves correspond to initial beam sizes of $\sigma_r = 10$, 20, 30, and 50 μm. Notice that, with an incoming 30 μm spot size, the temperature remains less than the stress limit of 180°C. A 30 μm spot size can be attained at 30 GeV by passing through a 0.2 radiation length spoiler 15 cm before striking the structure.

The picture is completed by looking at the failure modes. The magnet movers have a maximum range of ± 1 mm while in typical quadrupole designs, if a pole is shorted, the field on axis is roughly $B_{pole}/5$ and causes a 45° angle deflection of the beam.

Assuming a 100° cells, which is stronger focusing than in most of the linacs, a 1 mm offset of the first quadrupole would cause a 1.5 mm offset in the second quadrupole. In the linac, this beam would then strike

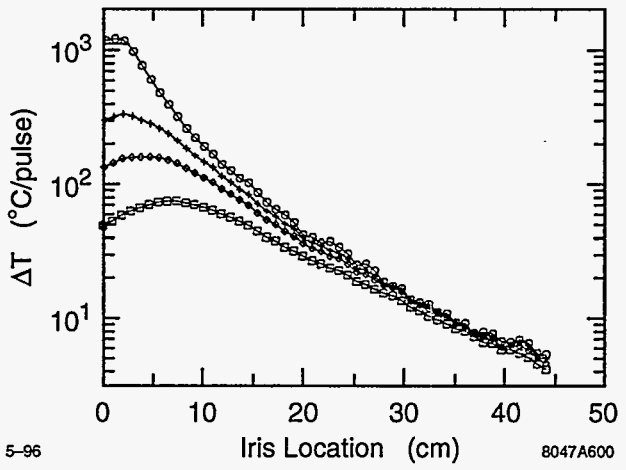


Figure 16-3. *EGS simulations of the temperature rise in the accelerator structure irises due a 30 GeV single bunch of* 1.3×10^{10}*; the different curves correspond to initial beam sizes of* $\sigma_r = 10, 20, 30,$ *and* $50\,\mu$m.

the accelerator structure irises, having a 3.9 mm radius, after 60% of the distance to the next quadrupole. Similarly, assuming a single shorted pole of a linac quadrupole with a 7 mm bore radius, the beam is also offset in the next quadrupole by 1.5 mm in both the horizontal and vertical planes and again the beam will strike the structures after 60% of the distance to the next quadrupole.

Both of these cases, are protected against by placing 0.2 radiation length spoilers with 2 mm radii at the end of every accelerator structure. As discussed in Chapter 7, the transverse wakefield deflection of the spoiler is roughly 10% of that due to an accelerator structure. Provided that the spoilers could accurately mounted to the structure ends, they would be aligned along with the structures as discussed in Chapter 7.

This solution works well until the beam energy is above roughly 400 GeV. At this point, the scattering due to the spoilers is small and a large amount of energy is contained in the electromagnetic shower. Thus, to be effective, the beam must intercept the spoilers a long distance before striking the irises. An example is illustrated in Figure 16-4 which shows the temperature rise in the irises due to a 500 GeV bunch of 1.3×10^{10} that has passed through eight 0.2 radiation length spoilers separated by 1.8 m. The four curves correspond to initial beam sizes of $\sigma_r = 10, 20, 30,$ and $50\,\mu$m.

Unfortunately, in both cases that were considered for the beginning of the linac, a single shorted pole or a 1-mm offset, the beam would only intercept six spoilers before striking an accelerator structure; this would lead to a temperature that is roughly a factor of two above the stress limit.

There are two solutions: first, we could adopt solution (3), direct monitoring of the magnets, or, second, we could use a single spoiler with a 1 mm radius located at the quadrupoles. Although the transverse wakefield of a spoiler with a 1-mm radius is roughly five times greater than that with a 2-mm radius, the single small radius spoiler would replace ten of the larger radius spoilers since there are ten accelerator structures between the quadrupoles at the end of the linac. Thus, the integrated wakefield is the same. Additional benefit is gained by placing these elements before the quadrupole magnets since a spoiler will induce a large energy spread in the beam which leads to a large decrease in the beam density after the beam is deflected by the subsequent quadrupole. Finally, the spoiler could be directly mounted on and pre-aligned to the quadrupoles; thus, they would be aligned to the beam when the quadrupole beam-based alignment is performed. At this time, we have not determined the best option and will reserve that decision for the future.

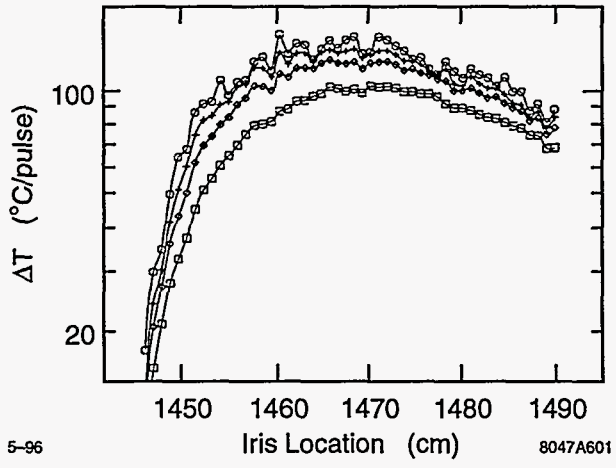


Figure 16-4. *EGS simulations of the temperature rise in the accelerator structure irises at the end of the 500 GeV linac due a single bunch of* 1.3×10^{10} *that has passed through eight spoilers separated by 1.8 m; the four curves correspond to initial beam sizes of* $\sigma_r = 10, 20, 30,$ *and* $50\,\mu m$.

Diagnostic Pulse Emittance Enlargement

As noted, to prevent destruction of the protection spoilers, the diagnostic beam emittance product $\epsilon_x\epsilon_y$ must be increased by a factor of 100. This enlargement must be turned-on and -off at the full repetition rate of 180 Hz without significantly steering the beam. The logical place to perform the enlargement is either in or just after the damping ring. At this time, we have considered two solutions: first, we could induce a large horizontal and/or vertical dispersion oscillation in the damping ring wigglers using pulsed quadrupoles. These would be located in the dispersion suppressor sections at the end of the arcs. As an example, to increase the vertical emittance by a factor of 100, skew quadrupoles can be used with an integrated strength of 7 kGauss—these change the tunes by less than 0.01 while increasing $\gamma\epsilon_y$ to 2.5×10^{-6} m-rad. Second, we could use a pulsed chicane, at the exit of the damping ring, to direct the beams through a gas filled chamber. Passing the beam through 1-m of *Ar* at a pressure of 1 Torr will increase the emittance product by 100. Alternately, one could use a thin foil ($50\,\mu m$) of *Be* although in this case the power density may be a problem.

16.2.2 Transition between diagnostic pulses and full beams

As discussed, a diagnostic beam, a high-emittance full-current single bunch beam, will be used to recover from a fault or to start up after an off period. To then establish full current operation requires a sequence of steps that include ramping the repetition rate, the emittance, and the number of bunches. Specifically, after the diagnostic pulse can be cleanly transported to the beam dump, the repetition rate must be increased to the nominal full rate. This is necessary since the MAID is only guaranteed for short periods. Next, the beam emittance can be decreased to the nominal and finally the bunch train can be lengthened. It is this last step that is probably the most difficult because of the changes in beam current and beam loading which can have significant dynamical and thermal effects in the damping rings and linacs. A single bunch diagnostic pulse was chosen since it's transport dynamics are, in many ways, similar to that of the full bunch train. Nevertheless, the transition from the diagnostic pulse and full intensity operation must be automatically checked at several steps along the way.

Linac Operations

When transitioning from the single bunch beam to the full bunch train care will be needed to verify the beam loading compensation. For the full bunch trains, the beam loading is roughly 25% in the main linacs. Thus, if the loading compensation is not set, the trailing bunch will have an energy roughly 25% too low.

To verify the loading, we will increase the number of bunches per train in steps. At this time, it seems reasonable to increase the number of bunches from a single bunch to 10 bunches per train. In this case the maximum energy deviation should be less the 3%, even if the loading compensation is not set. At this point, the loading profile can be accurately determined and thus it seems reasonable to increase to a full train in the next step; details of the beam loading compensation and control can be found in Chapter 8.

At each stage during the ramp the collider should operate for enough time to allow the beam-based feedback systems to stabilize. The time required should be less than a second (100 pulses).

In addition, active feedforward will be required to control the rf power delivered to the accelerator structures. Without changing the temporal profile of the rf power to the structures, they will cool significantly during full bunch train operation and will heat up after an MPS fault when operating with a single bunch. This arises because the full train absorbs roughly 25% of the rf power. As described in Section 8.7, the solution to this problem is to adjust the klystron phase profile so that the additional power from the rf pulse compressors is sent to the loads and is not delivered to the structures.

Damping Ring Operation

The damping ring rf system has been designed to accept full current trains at the maximum repetition frequency but it is heavily beam loaded. Thus, changes in the total current must be accommodated with the appropriate beam scheduling information which includes the repetition frequency and a measure of the incoming charge which is required by the rf feedforward algorithm.

There are two sequences that can be used to ramp the rings from zero to full current. The first is used after an extended off period or for initial checkout at reduced current. The second is used for ramping to full current quickly as required, for example, following a spurious MPS ion chamber trip. In either case, the number of bunches per train will be increased in steps from 1 bunch per train (bpt) to 10 bpt and then to 90 bpt.

In addition to the feedforward algorithm to adjust the rf systems for changes in the beam current, feedforward will be needed for thermal regulation. The damping rings emit an enormous amount of power as synchrotron radiation and higher-order modes which is removed by the cooling systems. Thus, the feedforward system, likely consisting of both flow control and heating elements, will be required to prevent significant thermal fluctuations during changes in the stored beam current.

Recovery procedure after extended off period For single bunch operation, beam loading of the rf cavities in the damping ring may be neglected. However, the cavity loading angle must be adjusted to at least partially minimize the reflected power during extended periods of low current operation. Because the reflected power depends only weakly on loading angle with 1 bpt, the tuners positions are moved to minimize reflected power at 10 bpt. This allows for efficient ramping of the current from 1 to 10 bpt. The injection procedure is given below:

1. With the beam off and the extraction kicker deactivated, the tuners are adjusted to $\phi_{z,10}{}^*$, which is the tuning angle for which the reflected power is minimized at 10 bpt.
2. Since feedforward for reduced beam loading is not required, the tuner loops are left closed and the ring is sequentially filled with 4 trains of 1 bpt.
3. The extraction kicker is activated and the ring is operated for as long as desired with single bunches using single-turn extraction and injection.
4. The repetition frequency is then increased, if desired.
5. The number of bunches per train is then increased to 10 bpt. The principle is the same as that for full-current (90 bpt) injection, with the exception that the voltage and phase offsets for feed-forward are adjusted for 10 bpt using Eqs. 4.64 and 4.65. (If the current of the incoming beam is known, then the offsets are automatically adjusted by the feedforward controller.)
6. To resume operation at 90 bpt, the beam is turned off and the tuner setpoints are adjusted to $\phi_{z,90}{}^*$ and the rapid recovery procedure is followed. (Here $\phi_{z,90}{}^*$ is the tuning angle for which the reflected power is minimized at full current—Section 4.5.3.)

Rapid recovery procedure The fastest possible recovery from no beam to full current operation requires a minimum of 12 cycles (or 66.7 ms at 180 Hz); the actual process will be slower to allow the feedback systems sufficient time to stabilize between changes. The procedure is as follows:

1. With the tuners fixed at $\phi_{z,90}{}^*$, the extraction kicker deactivated, and feedforward deactivated, inject 4 trains of 1 bpt.
2. Activate the extraction kicker and feed-forward. Begin single-turn extraction and injection. Inject 4 trains of 10 bpt.
3. Inject 4 trains of 90 bpt. (The presence of the 10 bpt trains in the ring is taken in into account in the feedforward algorithm.)
4. Close tuner loops and proceed with nominal operation.

16.2.3 Controlling the Interpulse Difference (MAID)

When operating with more than one bunch or when operating with the nominal beam emittances, the MPS must guarantee that the trajectory of the upcoming pulse is within the MAID. In order to do this, every device that can change the beam energy or trajectory by more than the MAID in the interval between pulses must be checked before allowing the permit to the scheduler. For transverse deflecting magnetic fields, the best way to do this is to actually prevent those fields from making large changes during the interpulse period. This is practical for all but the strongest magnets since the typical magnet L/R decay times can be made ~ 200 ms. High bandwidth magnets, such as linac fast feedback correctors, must not have enough strength to exceed the MAID. One consequence of this is that low repetition rate, full train intensity operation, is not possible.

Fast, powerful, pulsed dumper magnet and extraction kicker systems must provide an electronic warning of their behavior prior to each pulse. This warning must be timely enough to stop beam extraction from the damping ring (in the case of the linac) or to fire protection dumper magnets.

The linac klystrons must also provide a summary status signal that can be used for MPS. Since the NLC linac MAID requires the energy difference to be less than 10completion of these tests for each klystron is not required. The klystron pre-pulse checking system must be most effective for warnings about common mode failures that affect many klystrons. On a pulse-to-pulse basis, a klystron's amplitude, phase or trigger timing can vary outside nominal limits and both must be checked. The modulator switch high voltage will have a charging step in it's cycle that takes place about $100\,\mu s$ before beam time. A comparison of that amplitude against the expected value will be used as an input to the MPS permit. The RF drive will also be tested for amplitude at about that time. The phase is more difficult to measure and will require that a phase measurement be made using an independent reference system.

To verify the beam loading compensation when going to the longer bunch trains, the train length will be increased in steps. Presently, we believe that the loading could be checked using only one intermediate train length, namely, going from one bunch per train to 10 bunches per train and then to a full 90 bunches per train; this needs further verification and increasing the steps has minimal implications for the rest of the systems. In addition, to prevent thermal changes of the accelerator structures when changing from low repetition rate to high rate and from the single- to the multibunch modes, the klystron phases are varied so that the addition power which normally would accelerate the missing bunches is dumped into loads rather than into the accelerator structures; this is discussed further in Chapter 8.

References

[Particle Data Book] Particle Data Group, "Particle Properties Data Book", *Phys. Rev. D* , **50**, 1173 (1994).

[Walz 1973] D. Walz, D. Busick, T. Constant, K. Crook, D. Fryberger, G. Gilbert, J. Jasberg, L. Keller, J. Murray, E. Seppi, and R. Vetterlein, "Tests and Description of Beam Containment Devices and Instrumentation–A New Dimension in Safety Problems", SLAC–PUB–1223 (1973).

[Walz 1996] D. Walz, "Justification for Temperature Rise and Thermal Stress Limits", NLC–Note–22 (1996).

Contributors

- Vinod Bharadwaj
- John Irwin
- Michiko Minty
- Ralph Nelson
- Tor Raubenheimer
- Sayed Rokni
- Marc Ross
- Dieter Walz

17

NLC Reliability Considerations

Contents

17.1	Goals	904
17.2	Reliability and Availability	904
17.3	Target NLC Availability	905
17.4	NLC Machine Availability and System Reliability	906
17.5	A Formal Solution	908
17.6	Three Examples: Klystrons, Power Supplies, and Motors	911
17.7	Summary	911

17.1 Goals

The SLC operates with an overall accelerator availability of about 80% [Erickson 1995]. The NLC will be nearly ten times as large and consume approximately six times the power of the SLC. Simple scaling of the SLC fault rates to the NLC results in an NLC which is effectively never operational. It is important that the issues of NLC availability be addressed from the onset of the design and engineering process so that the required component and system reliabilities are achieved. The goals of this chapter are threefold:

1. Establish an availability/reliability specification for the NLC on a machine basis (e^- injector, damping ring, main linac, etc.) and on a system basis (power supplies, magnets, klystrons, etc.). These specifications are arbitrary by nature but are to be compared, as far as possible, with the operational experience of existing accelerator complexes. An availability target of 85% for the full NLC has been adopted.
2. Develop a formal solution to the problem of how availability/reliability is to be accomplished. This requires shifting the responsibility for availability/reliability from a separate and detached upper-level oversight management team to those who are responsible for system development, engineering, implementation, and maintenance. In order to succeed, the concepts of reliability and availability need to be integral to the systems development and must receive necessary resources through a bottoms-up approach with top-down support and review.
3. Identify where reliability engineering effort should be initiated because of discrepancies between performance requirements and known behavior, in those areas where information is lacking, and where exorbitant costs are projected.

17.2 Reliability and Availability

Reliability is the probability that an item or system will perform the necessary function without failure for a given period of time. Reliability, $R(t)$, is characterized by the mean time to failure, $MTTF$. For a system of N_s identical components, the $MTTF$ of the system is taken to be $MTTF_i/N_s$ where $MTTF_i$ is the mean time to failure of an individual component. For the case of constant failure rate λ,

$$\lambda = (MTTF)^{-1} \tag{17.1}$$

and

$$R(t) = e^{-\lambda t} \tag{17.2}$$

A more reliable system thus lasts longer between repairs than a less reliable system.

Availability is the probability that a repairable system will be available for use when required. Availability for the NLC is defined as A,

$$A = 1 - MTTR/MTTF \tag{17.3}$$

wherein $MTTR$ is the mean time to recover which is the average repair time plus accelerator operations recovery time. In general, system availability is enhanced by high reliability and short repair and recovery times. The definition 17.3 is adopted for the NLC accelerator as representative of a complex system in which additional components continue to fail during the time in which recovery is being made for a previous fault. To show how availability of the NLC can vary with respect to that of the SLC, Eq. 17.3 is plotted in

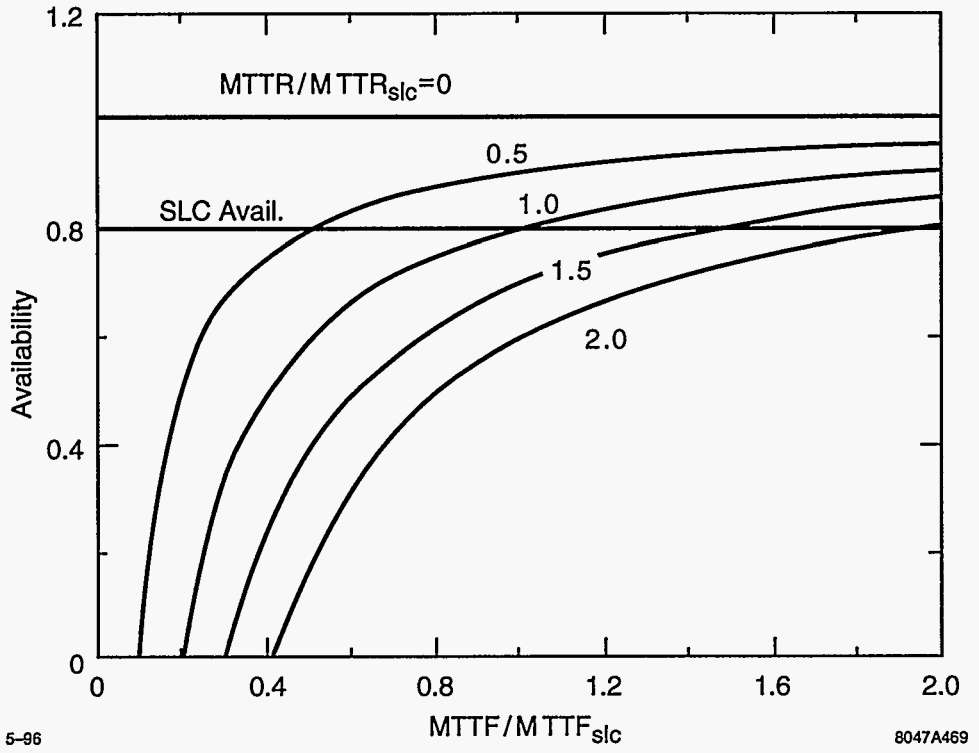


Figure 17-1. *Availability vs. $MTTF/MTTF_{slc}$ for various values of $MTTR/MTTR_{slc}$.*

Figure 17-1 over a range of $MTTF$, normalized by the $MTTF$ of the SLC ($MTTF/MTTF_{slc}$), for several values of $MTTR$, again normalized by the $MTTR$ of the SLC ($MTTR/MTTR_{slc}$). In Figure 17-1 it is seen that the availability is increased by reducing the $MTTR$ or alternatively increasing the $MTTF$ for a given $MTTR$. If the $MTTF$ in the NLC equals that of the SLC (albeit the increased number of NLC components) while the $MTTR$ increases by a factor of two due to say travel time, the availability drops to 60% compared with the value of about 80% for the SLC. Also from Figure 17-1 it is seen that if the $MTTR$ exceeds the $MTTF$, availability drops to zero. This has been the experience with SLC operations.

When discussing system performance, reliability is most often used as a figure of merit. For the machines which comprise the systems, availability is the appropriate figure of merit. As noted above, availability can be enhanced through high reliability. Fortunately, availability can also be improved through reduced repair and recovery times and through component redundancy.

17.3 Target NLC Availability

A target of 85% NLC availability over a scheduled running cycle of 6500 hours is assumed. This is a running period of nine months on and three months off in a calendar year. Numerous short maintenance and repair periods erode the time allocated for machine operations. One-shift-per-week maintenance is a 5% cost to operation. In a strict accounting view, one shift per week of scheduled maintenance during the nine-month cycle leaves only 10% of the time to be allocated to unscheduled outage, from all causes. Experience at SLAC

indicates that with the exception of some of the utility installations, very few of the accelerator components have a preventive maintenance program which require scheduled outage during a nine-month cycle. Most, if not all, scheduled maintenance tasks can be accomplished during an annual three-month down. The bulk of the eight-hour scheduled outages taken during a running cycle are used to accomplish remedial repair tasks which have accumulated during the period since the previous outage. These are most appropriately charged to unscheduled downtime.

It appears to be a straightforward task to design an NLC which has minimum maintenance requiring scheduled outages. The incremental cost to accomplish this is minimal given that so little of the present accelerator systems (both at SLAC and at other accelerator laboratories) have scheduled maintenance requirements. It is important, however, to identify those components which presently require periodic outages for maintenance and to reduce such requirements through judicious design and configuration modifications.

17.4 NLC Machine Availability and System Reliability

A proposed NLC availability specification has been developed for the NLC machines and systems for 85% availability over 6500 hours per year of scheduled operations. To develop this specification, the NLC has been divided into 12 machines: e^- source and linac, e^- damping ring and first compressor, e^- booster linac and second-stage compression, e^- main linac, e^- final focus and dumpline, and an identical breakout for the positron complex with the addition of the e^+ source and linac and e^+ pre-damping ring. Similarly, the NLC has also been divided into eight categories of systems: power supplies, magnets, klystrons, modulators, etc. When divided in the same fashion, the SLC consists of six distinct machines, each of similar complexity as an NLC counterpart; the SLC has the same eight categories of systems but with fewer components per system. The overall product of NLC machine availabilities is 85%; the overall product of the NLC system availabilities is 85%. A mean time to recover ($MTTR_s$) of one hour has been chosen for the systems. Equal weighting for each of the machines has been assumed except for the cases of the main linacs which are each given three times the weighting of the other machines. Table 17-1 lists the proposed availability specification for the various NLC machines and the assumed weighting factors. In Table 17-1 the listed availability is simply A_m,

$$A_m = 0.85^{w_m/16} \quad , \tag{17.4}$$

wherein w_m is the weight factor for a given machine, 16 is the sum of the 12 weight factors, and 0.85 is the target availability for the full NLC. Table 17-2 lists the proposed availability specifications for NLC systems. For the noted assumed $MTTR_s$, the required $MTTF_s$ for the system as a whole is given by

$$MTTF_s = MTTR_s/(1 - A_s) \tag{17.5}$$

where A_s is the listed system availability. For Tables 17-1 and 17-2, the allowed unscheduled outage is based on an assumed 6500 hours per cycle of scheduled operating time. Given the $MTTF_s$ for a system, the corresponding required $MTTF_i$ for an individual component is noted in Eq. 17.5.

Each of the 12 machines must be available 99% of the time (97% for the main linacs) in order to achieve the 85% availability goal. For a scheduled operating cycle of 6500 hours this allows for 66 hours of outage per machine per cycle (195 hours for each of the main linacs) The subtotal outage for the e^+ machines is greater than that of the e^- machines because of the added complexity of a positron production system and pre-damping ring.

Scheduled Operating Hours: 6500			
	Weight	Availability	Unscheduled Outage (hours)
e^- Inj, Source and Linac	1	0.99	66
e^- DR and Compressor 1	1	0.99	66
e^- Booster Linac and Comp. 2	1	0.99	66
e^- Main Linac	3	0.97	195
e^- Final Focus and Dumpline	1	0.99	66
Subtotal e^- machines:	**7**	**1**	**458**
e^- Inj, Source and Linac	1	0.99	66
e^+ Source and Linac	1	0.99	66
e^+ Pre-damping Ring	1	0.99	66
e^+ DR and Compressor 1	1	0.99	66
e^+ Booster Linac and Comp. 2	1	0.99	66
e^+ Main Linac	3	0.97	66
e^+ Final Focus and Dumpline	1	0.99	66
Subtotal e^+ machines:	**9**	**1**	**589**
Totals:	**16**	**0.85**	**1047**

Table 17-1. *Availability specifications for the NLC machines.*

NLC Systems	Availability	$MTTR_s$ (hours)	$MTTF_s$ (hours)	Unscheduled Outage (hours)
Power Supplies	0.975	1	40	163
Magnets	0.975	1	40	163
RF Systems	0.950	1	20	325
Motors	0.975	1	40	163
BPMs	0.990	1	100	65
Controls	0.985	1	67	98
Utilities	0.995	12	2400	33
Miscellaneous	0.995	1	200	33
Totals:	**0.85**			**1040**

Table 17-2. *Availability specification for the NLC systems.*

A preliminary specification of NLC component reliability has been developed. The minimum $MTTF_i$ of the components which is needed to achieve the system availability specification is given by

$$MTTF_i = N_s MTTR_s/(1 - A_s) \tag{17.6}$$

wherein N_s is the number of identical components in a system, $MTTR_s$ is the mean time to recover of the particular system, and A_s is the specified availability for the system. As an example, for $N_s = 1500$, $MTTR_s = 1$ hour, and $A_s = 0.995$, the required $MTTF_i = 300{,}000$ hours.

SLC machine and system availabilities and component $MTTF_i$ have been compiled for the 1992, 1993, and 1994/1995 SLC operating cycles. Operating experience of SLAC systems compares favorably with experience at Fermilab, CERN, KEK Photon Factory, Cornell, APS, and AGS. Table 17-3 lists the accelerator availabilities for physics of these various laboratories; the running cycles are noted. In general it was found that the same sorts of problems exist at all the labs. When the lengths of the running cycles are considered along with the sizes of the various machines and the peculiarities of the various accounting methods, the performance of the different accelerators are quite similar. Some labs do better with certain technologies than others but there are no clear differences on the whole. Because of the apparent similarities between the labs, it has been decided to base NLC technology expectations on SLAC experience, since the details of the SLAC data are more readily available at SLAC. It is important however to make comparisons with the other labs on a case-by-case basis when anomalies or uncertainties occur. On average, the six SLC machines (injector, two damping rings with compressor systems, e^+ source, linac, and arcs and final focus) each had an availability of approximately 97%.

Table 17-4 lists a preliminary parts count for the NLC. This information was taken from the NLC ZDR WBS [NLC WBS 1996]. For comparison purposes, Table 17-5 lists a parts count for the SLC. The data in Table 17-5 was gathered by counting entries in the SLC control system database. Initial counts of the numbers of NLC components indicate that there is about a factor of ten more components of all types in the NLC compared to a similar count of SLC components. Attention must be paid to improving the performance of NLC systems over that which is being achieved in existing systems of similar complexity.

17.5 A Formal Solution

Achievement of the specified NLC availability comes through the integration of the system and machine availability/reliability specifications into the component, system, and machine-functional specifications at the onset of the engineering design phase. Performance specifications of individual components will include the specification of reliability. The design review process must include attention to the availability/reliability requirements. A precision supply that never works is no better than an out-of-tolerance supply that never fails. Within a machine the availability budget must be respected. This task is best done at the engineering level but must be managed in the same fashion and at the same time that the more familiar performance criteria are managed.

Reliability engineering is a recognized discipline which plays an important role in all technologically-oriented industries (*e.g.*, semiconductor, aeronautics and astronautics, automotive, telecommunications, and power industries). There are a number of professional societies dedicated to developing the techniques and methodologies of reliability (*e.g.*, IEEE Reliability Society, Society of Automotive Engineers, Society of Reliability Engineers, Society of Logistic Engineers, American Institute of Aeronautics and Astronautics [RS IEEE, SAE, SRE, SLE, AIAA]). There are numerous annual meetings of these societies wherein tutorials on these methods are given in addition to the familiar conference presentations of topical issues (the Annual Reliability and Maintainability Symposium [ARMS 1996], for example). There are a large number of textbooks and

Laboratory	Availability	Reference
ANL (APS) 95	68.30%	Argonne National Lab., Private Communication, Site Visit – R. Gerig, D. Ciarlette
CERN (SPS) 94	69.30%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 93	72.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 92	74.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 91	72.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 90	74.00%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
CERN (SPS) 89	71.20%	1994 SPS & LEP Machine Statistics CERN SL / Note 95–15 (OP) M. Colin, G. Cultrut and B. Desforges
Fermi 91	72.64%	Fermi Accelerator System Tally Sheets, Site Visit – R. Mau
Fermi 92	65.86%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
Fermi 93–94	63.71%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
Fermi 93–94	63.71%	Fermi Acclerator System Tally Sheets, Site Visit – R. Mau
SLAC (SLC) 92	81.00%	1992 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (SLC) 93	84.53%	1993 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (SLC) 95	80.87%	1994/95 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 92	87.01%	1992 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 93	93.25%	1993 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC (ESA) 94	93.33%	1994 SLC Revealed Failure Tables, Internal SLAC Memo – W. Linebarger
SLAC SSRL 94	97.04%	SSRL, Private Communication, Site Visit – E. Guerra
SLAC SSRL 95	96.60%	SSRL, Private Communication, Site Visit – E. Guerra
AGS, FY95Q3	86.30%	Brookhaven National Lab, FY 95 3rd Qtr. Report – F. Weng
AGS, FY94Q4	86.70%	Brookhaven National Lab, FY 94 4th Qtr. Report – F. Weng
Cornell 91–92	74.10%	CESR Reliability Summary FY 1992–FY 1994 – D. Rice
Cornell 92–93	77.90%	CESR Reliability Summary FY 1993–FY 1994 – D. Rice
Cornell 93–94	84.00%	CESR Reliability Summary FY 1994–FY 1994 – D. Rice
KEK Photon Factory Linac 10/92–9/93	98.70%	KEK Operations Report FY 1992–FY 1993
KEK Photon Factory Linac 10/91–9/92	98.40%	KEK Operations Report FY 1991–FY 1992
KEK Photon Factory Linac 10/90–9/91	97.70%	KEK Operations Report FY 1990–FY 1991

Table 17-3. *Availabilities of several accelerator laboratories.*

	Pwr sup	Magnets	Klystrons	Modulators	Motors	BPMs	Sys. Total
e^- Inj, Source and Linac	245	229	16	16	0	381	**887**
e^- DR and Compressor 1	817	709	5	5	300	555	**2391**
e^- Booster Linac and Comp. 2	452	482	116	116	1077	291	**2534**
e^- Main Linac	736	756	2264	1132	14643	5300	**24831**
e^- Final Focus and Dumpline	871	1466	1	1	1344	472	**4155**
e^- Inj, Source and Linac	244	229	40	40	0	381	**934**
e^+ Source and Linac	236	241	32	32	0	81	**622**
e^+ Pre-damping Ring	700	700	2	2	300	300	**2004**
e^+ DR and Compressor 1	817	709	5	5	300	555	**2391**
e^+ Booster Linac and Comp. 2	452	482	116	116	1077	291	**2534**
e^+ Main Linac	736	756	2264	1132	14643	5300	**24831**
e^+ Final Focus and Dumpline	871	1466	1	1	1344	472	**4155**
NLC Total	**7177**	**8225**	**4862**	**2598**	**35028**	**14379**	**72269**

Table 17-4. *Preliminary NLC parts count for several systems.*

	Pwr sup	Magnets	Klystrons	Modulators	Motors	BPMs	Sys. Total
e^- Inj, Source and Linac	249	247	16	16	10	37	**575**
e^- and e^+ DRs and Compressors	40	456	5	5	6	199	**711**
e^+ Source and Linac	30	452	2	2	5	204	**695**
e^- Main Linac	608	608	242	242	22	283	**2005**
SLC Arcs	119	1000	0	0	912	978	**3009**
SLC Final Focus	192	192	0	0	23	59	**466**
SLC Total	**1238**	**2955**	**265**	**265**	**978**	**1760**	**7461**

Table 17-5. *SLC parts count for several systems.*

courses on availability [Lewis 1996, O'Connor 1985]. It is important to take advantage of the tools developed and to apply them to the issues of NLC reliability. It is also necessary to understand the lessons learned in areas other than accelerators and to apply these lessons to the problems facing NLC construction. In many cases, the detailed solutions of how reliability in a Boeing 777 is achieved are not directly applicable to the NLC, but the thought processes going into developing a Boeing 777 are identical to what is required to successfully meet the NLC reliability goals.

For the NLC CDR, it is important that the issues associated with component reliability and system availability be fully integrated into the component and system engineering. Segregation of the discussion of availability into a separate chapter (in the CDR) will not fulfill the need to infuse the requirement for reliability beginning at the most basic levels of NLC design. If availability is to be achieved for a system which is nearly ten times larger than what has previously been achieved by the accelerator community, reliability must be fully accepted by the engineering and fully supported by the management.

Availability of the systems is based on the reliability of the individual components in concert with component configurations which include considerations of system repairability and redundancy. The solutions are specific to the particular systems; redundancy in the rf systems is a straightforward cost-effective solution, whereas component reliability combined with ease of changeability appears to be the proper solution for many of the magnet power supply applications.

17.6 Three Examples: Klystrons, Power Supplies, and Motors

In the main linacs, the expected $MTTF_i$ of the klystrons is 20,000 hours [Caryotakis 1995] and the $MTTF_i$ of the thyratrons is 10,000 hours [Wait 1996] Given an estimated count of 4000 klystrons and 2000 thyratrons in the NLC, approximately 1300 of each will fail and need replacement every cycle; this is a combined failure rate of one klystron or modulator every 2.5 hours. In order to operate the machines, on-line redundancy is required. By necessity, the repair rate must be equal to or faster than the failure rate. Therefore, the availability for the rf system is simply A_{rf}

$$A_{rf} = 1 - e^{-1}/n! \tag{17.7}$$

where n is the number of redundant rf modules available for use when needed. For $n = 6$, $A_{rf} = 0.9995$. Present plans call for 3% redundancy in the number of rf modules which is quite sufficient. The rf systems are an operating cost issue but not so much one of availability. It is important to work to extend the $MTTF_i$ of the klystrons and thyratrons so as to reduce the cost of these consumables. It is worth noting, that effort must go into developing reliable waveguide valves to permit changing to klystrons during accelerator operations and to design the modulators such that the thyratrons can be easily changed.

There are approximately 750 quadrupoles per main linacs. The power supplies for these magnets are expected to be in the power range of a few kilowatts each. For the pair of linacs, the $MTTF_i$ of the power supplies is 300,000 hours to give system availability of 0.995, assuming the nominal one-hour $MTTR_s$. Should the $MTTR_s$ increase to two hours due to travel time or complexity of changing, the $MTTF_i$ increases to 600,000 hours. Rack-mounted power supplies in this power range used at SLAC have an $MTTF_i$ of about 300,000 hours [Donaldson 1996] and an $MTTR_i$ of about 1.5 hours. Whereas the present performance of similar power supplies meet the NLC goals, care must be taken to keep the $MTTR_s$ of less than one hour.

There are approximately 35,000 motors in the quadrupole and structure mover systems of the two NLC linacs. Since a stuck mover is a "soft" failure that contributes to emittance growth but does not stop the machine dead, it has been decided to allow 1% of the motors to fail each month before stopping to fix the accumulated failures. A failure rate of 1% per month corresponds to a $MTTF_i$ of 8.3 years. Motor manufacturers claim $MTTF_i$s of five to seven years for 100% duty factor usage and seven-to-ten-year $MTTF_i$ for 50% duty factor usage [Parker 1996, Warner 1996]. SLC experience has been quite good with motors. However, it will be important to design the movers with motor replaceability in mind since 1% per month failure rate is 3500 failures per year and the $MTTR_i$ needs to be small (on average 350 motors need to be replaced each month during a "short" machine access).

17.7 Summary

Simple scaling of the SLC fault rates to the NLC results in an NLC which is not operational. Reliability and availability need to be fully integrated into the functional requirements of the NLC. Reliability and availability must be explicit at the component, system, and machine levels in the CDR as a natural and normal part of the accelerator design. Real consideration and effort must be dedicated to defining and solving the reliability issues. The solutions to these issues necessarily arise from the engineering teams charged with building the systems. There exist significant engineering disciplines dedicated to addressing the issues, but care needs to be taken such that the correct solutions are properly applied to the relevant problems.

References

[AIAA] American Institute of Aeronautics and Astronautics, 370 L'Enfant Promenade, SW, Washington, DC 20024-2518 USA.

[ARMS 1996] *1996 Proceedings* and *1996 Tutorial Notes*, Annual Reliability and Maintainability Symposium, c/o Evans Associates, 804 Vickers Ave., Durham NC 27701-3143 USA.

[Caryotakis 1995] G. Caryotakis, personal communication (1995).

[Donaldson 1996] A.R. Donaldson, personal communication (1996).

[Erickson 1995] R. Erickson, C.W. Allen, T.K. Inman, W. Linebarger, M. Stanek, SLAC–PUB–95–6895, in "SLAC Accelerator Operations Report: 1992–1995", *Proc. of the 1995 Part. Acc. Conf.*, Dallas TX (1995).

[Lewis 1996] E.E. Lewis, *Introduction to Reliability Engineering* 2nd ed., (John Wiley & Sons, Inc., New York, 1996).

[NLC WBS 1996] NLC Project WBS and Outline, Version 3–20–96 (1996).

[O'Connor 1985] P.D.T. O'Connor, *Practical Reliability Engineering*, 2nd ed., (John Wiley & Sons, Inc., New York, 1985).

[Parker 1996] Parker Hannifin Corp., Compumotor Div., 5500 Business Park Dr., Rohnert Park CA 94928 USA, personal communication (1996).

[RS IEEE] Reliability Society IEEE, Institute of Electrical and Electronic Engineers, PO Box 1331, Piscataway NJ 08855-1331 USA.

[SAE] Society of Automotive Engineers Inc., 400 Commonwealth Dr., Warrendale PA 15096-001 USA.

[SLE] Society of Logistic Engineers, 8100 Professional Place, Suite 211, Hyattsville MD 20785-2225 USA.

[SRE] Society of Reliability Engineers, 2002 Harpoon Dr., Stafford VA 22554 USA.

[Wait 1996] G.D. Wait and M.J. Barnes, "Thyratron Lifetimes, A Brief Review", *Proc. of the 1995 Second Modulator-Klystron Workshop*, SLAC (1995); SLAC Report 481 (1996).

[Warner 1996] Warner Electric, Linear Motion and Electrical Div., Superior Electric, 383 Middle St., Bristol CT 06010 USA, personal communication (1996).

Contributors

- C.W. Corvin
- A.R. Donaldson
- J.W. Humphrey
- W.A. Linebarger
- M.H. Munro
- A. Saab
- J.C. Sheppard

18

NLC Conventional Facilities

Contents

18.1	Introduction	916
18.2	Site	916
18.3	Campus	917
18.4	Injectors	917
18.5	Linac	917
18.6	Detectors	918

18.1 Introduction

The NLC conventional facilities integrate the geotechnical and topological conditions of a generic site with a sub surface linac housing and klystron gallery structure. Conventional above-grade infrastructure and support facilities generally found in high-energy physics laboratories are included along with other necessary items associated with such a very large project. A non-specific site is assumed as no site has yet been selected. Many detailed specifics regarding the conventional facilities are not as yet addressed. The goal for this report has been to quantify the general ideas to a degree adequate to define the starting point for a detailed bottoms-up conceptual design report effort. Figures 18-1 through 18-5 show details of the NLC facility and are found at the end of this chapter.

18.2 Site

The overall NLC site is approximately 32 km in length and two km in width with four distinct functional areas; detectors, campus, injectors and linac. The site is assumed to be located such that reasonable access to adequate power, water, transportation and housing is available. The center of the site would be the location of the interaction point and the detector facilities. This central location might also be the location of the campus area, however the best campus area will very likely be selected with local access and transportation considerations being the driving factors. Opposite ends of the site would each have an injector area with a damping ring. At just one of the injectors a positron target and a pre-damping ring would be added. Between the detector area and each of the two injector areas would be a linac housing and a klystron gallery.

The linac areas with their associated utility support structures would make up the largest single portion of the overall site conventional facilities

The NLC site geology and topology should be such that both tunneling and cut-and-cover methods of construction are practical and utilized. A site having a high proportion of competent rock would be ideal and is the preferred site considered for the NLC. Excellent work with respect to geology, topography and site selection was done for the SSC project and has been used to advantage for the NLC. Attachment C to the SSC Conceptual Design Report, dated March 1986, documents that earlier work.

The NLC is estimated to require 300 MW of electric power to be supplied from two 230-kV transmission lines running parallel to the NLC. Three main substations are planned, one near each end of the site and one in the center. The primary distribution voltage from the main substations along the linac would be 34.5 kV. Two hundred small unit substations, adjacent to or in utility clusters along the klystron gallery, would supply the appropriate utilization equipment distribution power. The site water-cooling system would consist of approximately eight forty-MW cooling towers distributing water to utility cluster heat exchangers that transfer heat from ten-megohm low-conductivity water systems. Total site water demand would be about 3,000 gallons per minute and would come from a combination of existing offsite sources and those developed onsite. Conventional site facilities would include roads and parking for about 1400 vehicles as well as fire protection, water and waste treatment, communications, cable plant, and construction-related mobilization and site preparation. The primary roadway to the site would be four 12-ft lanes to support heavy assembly deliveries. The secondary linac roadway would have two 12-ft lanes with shoulders adequate to stabilize the roadway base.

For the major infrastructure construction phase of the project a concrete batch plant is assumed to be provided on site with a rail siding extension to bring in the associated bulk material needed for concrete. This

rail extension would be of use later for locating modular installation and maintenance clusters associated with NLC operations. These clusters would be 40 and 20-ft transportainers loaded with tools, stored components, field offices, and mini shops. The modular clusters would be relocated off-shift to follow and support on-shift operations, installation and maintenance.

18.3 Campus

The campus area is planned to house and support about 2500 persons and provide the needed facilities to build, operate and maintain the NLC. It consists of 16 buildings, the largest being a four-story main laboratory building of 350,000 square feet. This building would include a 1000-seat auditorium, a cafeteria, conference rooms, offices and light electronics laboratories as well as a central computer facility. This building would house the central administrative area for the laboratory.

Six heavy-fabrication buildings with high bays, office mezzanines, light and medium cranes, would be needed in the industrial area of the campus together with three smaller shops buildings. The shop buildings would be used for various machining and assembly operations. These industrial buildings would total roughly 210,000 square feet and have an adjacent paved outdoor staging area for larger assembly and preparation. Two conventional warehouses will be needed during and after construction for storing the various materials and assemblies that will make up the NLC. These structures would total about 80,000 square feet in area. The campus area would also have a vehicle service area, water and sewage treatment facilities and an emergency rescue and services facility.

18.4 Injectors

Each electron injector facility will include an injector with a short linac section, a 715-ft-circumference damping ring, and a compressor section with a second linac section. The positron injector will also have a pre-damping ring and a positron target. These facilities will be rigid-frame box structures of cast-in-place concrete. They will be positioned at a depth to align with the arcs at the ends of the main linac beam tunnels. The roof spans in the damping rings are substantially greater than for the bored tunnels. A hard rock site, mined with a horizontal cutter wheel machine, would be ideal for the damping-ring enclosures. Tunnel sections between the injector and damping-ring enclosures may be bored. Structures for the supporting injector utility clusters would be somewhat similar to those used for the main klystron gallery and linac. Should the campus area be located adjacent to one of the injector areas, then the supporting utility, power supply and instrumentation structures could be installed at the surface. The injectors will have local control rooms that slave to the main NLC control room.

18.5 Linac

The linac housing tunnel is planned to be bored with a tunnel boring machine at an average depth of forty ft. Its total length is a maximum of about twenty miles, less sections where the beam passes through cast buildings and enclosures such as the detector area and the damping rings. The finished inside diameter is planned to be 12 ft, with a concrete or shotcrete liner and a concrete floor or invert. A hard rock site

might possibly require less concrete depending on rock strata and water migration. Utility chases from the linac housing to the klystron gallery and utility cluster alcove above will be installed at 156-ft intervals along the linac to provide water cooling and other utilities. The minimum earth cover between the ceiling of the linac beam housing and the floor of the klystron gallery above is 20 ft for adequate personnel radiation shielding. The klystron gallery above will enclose klystrons, modulators, power supplies, instrumentation and controls. It will extend over most but not all of the beam housing tunnel below, as klystrons and related rf penetrations are not required along the entire length of the beam housing tunnel. A minimum of twelve miles of klystron gallery is planned for the rf systems. It is planned to be 30 ft in width. The utility cluster alcoves, at 156-ft intervals, will enclose motors, pumps, cooling and power system panels and controls.

The klystron gallery and utility cluster alcoves would be constructed using cut-and-cover methods and be below the surface at a depth sufficient to distribute surface traffic loads without the need for an excessively thick roof structure. Besides the utility chase housing penetrations, the housing will have 24-in inch diameter rf waveguide penetrations at six-m intervals along the beam housing. The utility chases will contain all utilities, including low conductivity water and cables, but would exclude rf waveguide. Housing and gallery entrances for vehicles and personnel are planned to be spaced at one-km intervals with adequate access ventilation and lock-up baffles to restrict air movement during linac operations.

The final-focus areas will extend from the detector building to the interaction point switch wyes at the ends of the linac where the beam is turned gradually to allow for two IPs and two detectors. The wyes are approximately two km from the interaction points. The two interaction points will be 40-m apart in the detector building making a single building feasible to house two detectors. A total of eight muon spoilers will be distributed in the four final-focus sections. Sizable structures to support and install these will be required. Their nominal dimensions fill the 12-ft inside diameter of the bored linac tunnel for three m along the beam z axis.

18.6 Detectors

The detector building facility will require about 100,000 square feet of total area. It will have a high bay area with a heavy bridge crane common to the entire facility. Two pit areas will match the two detector bore elevations to the final-focus beam line elevations. The pit will have light concrete sidewalls with earth tiebacks for support. The two pits will be structurally isolated from each other to segregate respective motion and vibration. Mezzanines will be included for computer, control room and office facilities. Areas for essential power conversion, low conductivity water, and cryogenic equipment will also be included both within and about the detector building. Shop areas will be provided for welding, metal fabrication, and instrument testing and calibration.

The two detectors will weigh in the order of 50,000 tons each and will require a precision steel and concrete floor pinned to bedrock for support. The detectors are assumed to be constructed in sub-assemblies that are skid, jacked and hoisted into place. The design of the detector building is driven by the design of the detectors themselves as the orientation and sequencing necessary for detector assembly must be consistent with the building layout. The detector control room will operate both detectors along with the instrumentation and data acquisition necessary for the relevant experimentation. Detector operation and control will be both local and remote to the NLC main control room. Kitchen, shower and rest facilities will be included to support personnel around the clock. Table 18-1 lists attributes of the interaction building, as well as those of the balance of the NLC conventional facilities.

General Site			
Area:	16,000 acres	2 km × 32 km	64 square km
Orientation:	North-South	Adjacent power R/W	Two circuits @ 230 kV
Population:	2500 persons	Peak occupancy	400 vehicles
Power:	300 MW	3 230-kV substations	@ 100 MW each
Water:	3000 gpm	Base + peak blowdown	@ 300 MW
Cooling:	320 MW	8 cooling towers	@ 40 MW each

Site Components		
Campus:	800,000 sq. ft.	Total square feet
Laboratory Bldg.	350k sq. ft.	One 4-story office-lab
Fabrication Bldgs.	181k sq. ft.	2 hvy., 4 med., w/ cranes
Shop Bldgs.	37k sq. ft.	Three light machine shops
Warehouses	80k sq. ft.	Two @ 40k sq. ft. each
Miscellaneous Bldgs.	152k sq. ft.	Small light buildings
Structures:		
Klystron Gallery	18.4 km × 9.25 m	1,809,000 sq. ft.
Beam Housing	31 km × 3.7 m ID	14 ft. TBM bore @ −40 ft. el.
I&C, Utility Alcoves	650, two per sector	400 sq. ft. each
Utility Chases	650, two per sector	6-ft dia. × 25 ft
rf Penetrations	2318 at 26-ft intervals	2-ft dia. × 25 ft
Housing & KG Access	34 with labyrinths	Two tiered ramps @ 1 km
Damping Rings	Two e^+ and One e^-	715-ft cir., 50k sq. ft. ea.
Injectors	Two housings	23k sq. ft. each
Position Target	One dual housing	28k sq. ft. for 2 targets
Final Focus	Four @ 2 km each	w/muon spoilers & dumps
Interaction Bldg.	One bldg., pit & crane	100k sq. ft. for 2 detectors
Utilities:		
Electrical	34.5-kV ring bus & ckts.	200 double-end unit subs
Communications	Phone, computer, radio	w/ fibre net, SCADA, video
Site H_2O + Waste	Water-waste, processing	Fire, reclaimed water lagoon
Roads & Parking	54 km roads, 1400 spaces	34 km 2-lane, 20 km 4-lane
Alcove Utilities	650 system clusters	Accel, klys, mag LCW; CA, HVAC
Cable Plant	Wire, tray, racks	337,750 cables; 6,825 racks
Alignment Network	Surface, X-fer, tunnel	Laser trackerbot, barcode index

Table 18-1. *Next Linear Collider zero-order design: Facility criteria.*

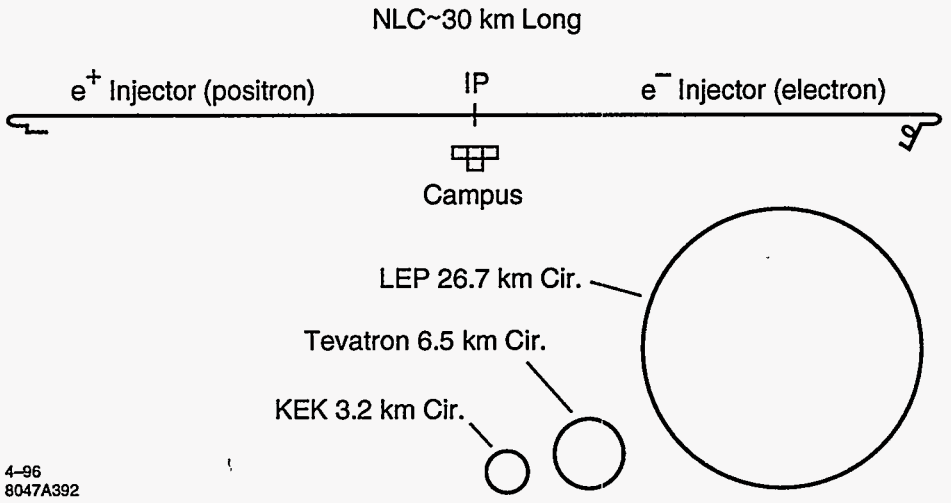


Figure 18-1. *NLC site relative proportions.*

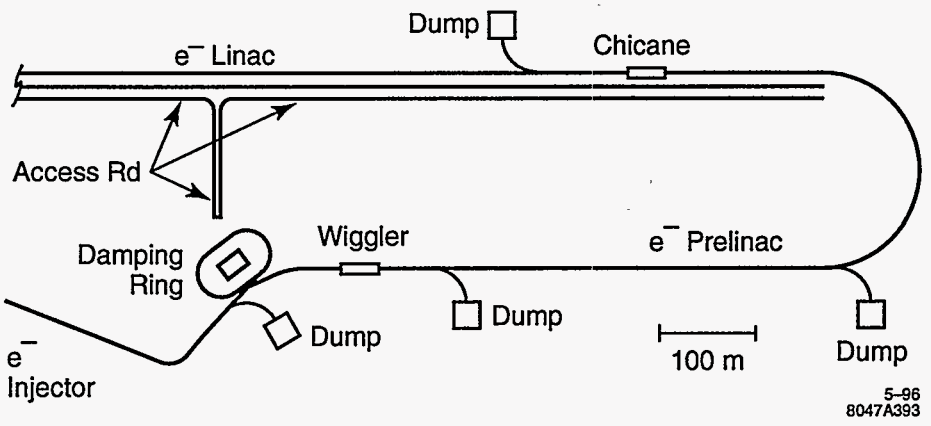


Figure 18-2. *NLC e^- injector and damping ring.*

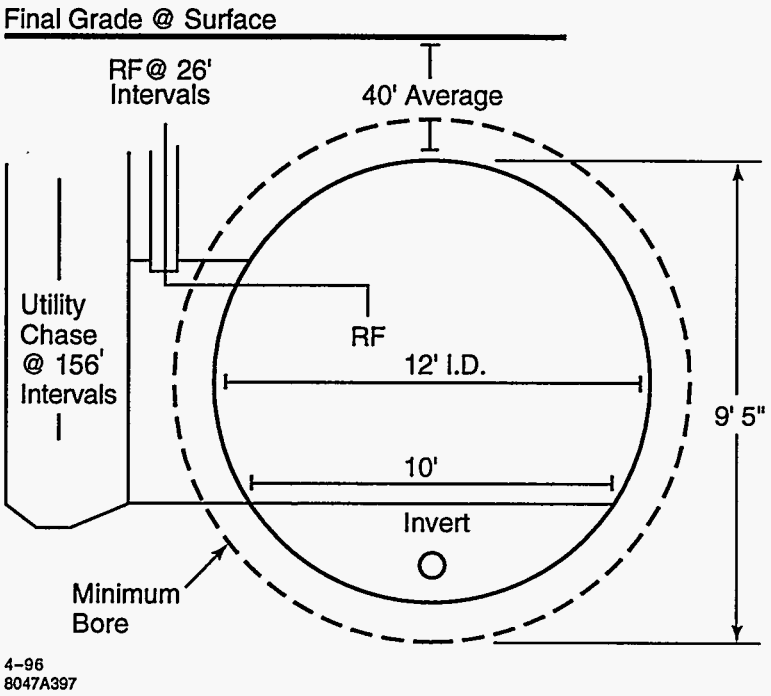


Figure 18-3. *NLC beam tunnel housing.*

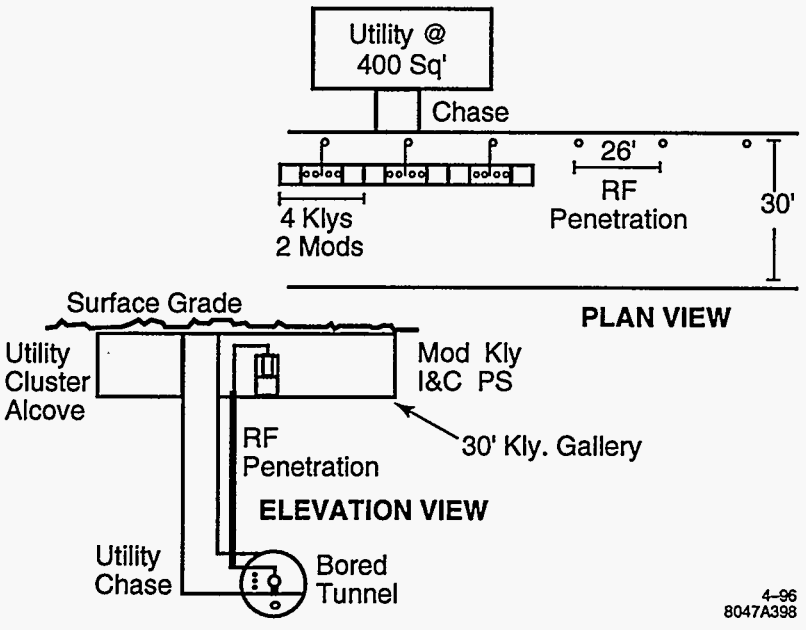


Figure 18-4. *NLC half-sector linac layout.*

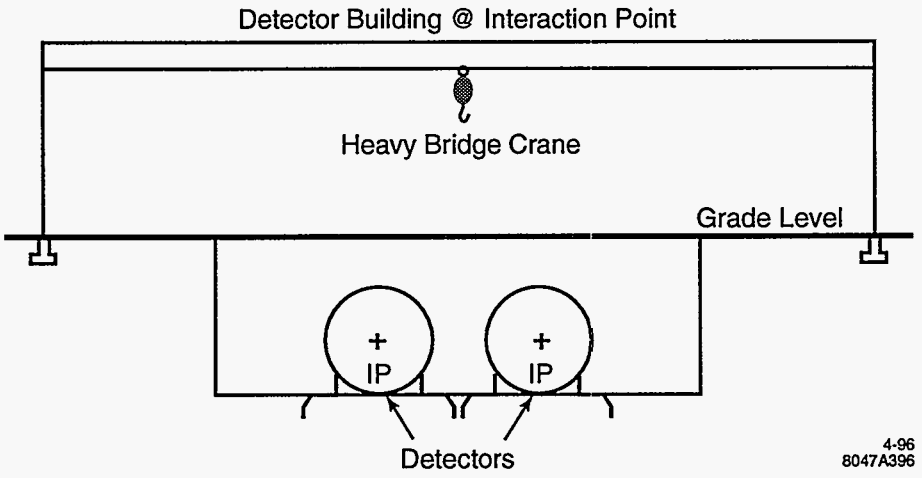


Figure 18-5. *NLC interaction point.*

Contributors

- C. Adolphsen
- C. Corvin
- T. Elioff
- R. Fuller
- D. Hopkins
- P. Kaul
- R. Ruland
- P. Rodriguez
- B. Schmidt
- S. Virostek

A

An RF Power Source Upgrade to the NLC Based on the Relativistic-Klystron Two-Beam-Accelerator Concept

Contents

A.1 Introduction 926
A.2 A Design for an RK-TBA-Based rf Power Source 927
A.2.1 Power Source Requirements 927
A.2.2 TBNLC Architecture 928
A.2.3 Main RK 929
A.2.4 Front End and Back End 935
A.2.5 The Key Ideas 936
A.3 TBNLC Physics Studies 937
A.3.1 Longitudinal Beam Dynamics 937
A.3.2 Transverse Beam Dynamics 940
A.3.3 RF Extraction Cavity Design 943
A.3.4 Induction Cavity Design 945
A.4 TBNLC Engineering Design 947
A.4.1 TBNLC Systems Study Approach 947
A.4.2 Electrical Systems 947
A.4.3 Mechanical Systems 953
A.5 RTA Test Facility 956
A.5.1 Induction Cores and Pulsed-Power System 957
A.5.2 Injector: Gun and Accelerator Sections 958
A.5.3 Chopper: Beam Modulation 958
A.5.4 Adiabatic Compressor 960
A.5.5 RF Power Extraction 960
A.5.6 Beam Dynamics Issues 962
A.6 Conclusions 964

A.1 Introduction

There are a number of possible routes to upgrade the center-of-mass energy to 1.5 TeV. The most straightforward is to use the same accelerating gradient as the 1-TeV design, 85 MV/m, and increase the length of the X-band linear accelerators by 50%. This would require extending the 10-GeV trombone arm to incorporate the extra length. In addition, magnets in the final focus and collimation regions would have to be rearranged, but the lengths of these regions have already been sized for the 1.5-TeV upgrade.

Unfortunately, the present power sources which would be used to attain the 1-TeV-c.m. energy, *i.e.,* the 50-MW or 75-MW klystrons and SLED-II pulse compression system, are too inefficient for the 1.5-TeV design; they would require an AC power in excess of 350 MW. There are a number of possible future power sources that might be utilized for the 1.5-TeV NLC such as the sheet beam and cluster klystrons, the Binary Pulse Compression system, and the Two-Beam Accelerator. In this appendix, a more detailed description of a Two-Beam Accelerator source, which has been designed at LBNL and LLNL, is described.

As an rf power source candidate for linear colliders, two-beam accelerators (TBA) [Sessler 1982, Sessler 1987] have the inherent advantage of very high efficiency for power conversion from drive beam to rf. In addition, induction-linac-based TBAs have favorable scalings with high frequencies ($\geq$11.4 GHz) and high accelerating gradients ($\geq$100 MV/m). Conversion of high-current electron beam power to rf power has been demonstrated at the gigawatt level at 34 GHz in free-electron laser experiments [Orzechowski 1986], and at several hundred megawatt levels at 11.4 GHz in relativistic klystron experiments [Allen 1989]. Reacceleration experiments [Westenskow 1994] have successfully demonstrated bunched beam transport through two reacceleration induction cells and three traveling-wave extraction cavities for a total rf output of over 200 MW. The phase and amplitude were shown to be stable over a significant portion of the beam pulse.

The technical challenges for making TBAs into realizable power sources lie in the dynamics of the drive-beam which must propagate over long distances. In particular, the beam breakup (BBU) instability through a long, multicavity, relativistic klystron (RK-TBA) is known to be severe. While BBU suppression techniques have been successfully demonstrated for a few cavities [Haimson 1992, Houck 1992a], a scenario with acceptable BBU control over many traveling-wave cavities must be constructed. Similarly, the longitudinal stability of the rf bunches over a multicavity TBA must be demonstrated. For rf phase and amplitude stability, the induction machine must produce and maintain a beam with constant current and energy over the duration of the pulse. In addition to technical feasibility, a case for economic attractiveness is no less essential for the viability of the TBA. Cost and overall system efficiency are essential elements of an acceptable linear-collider power source.

With these general considerations in mind, a systems study, including physics and engineering designs, as well as bottom-up costing, was conducted for a point design using the RK-TBA concept as a power source for the upgraded Next Linear Collider (NLC) [Siemann 1993]. We refer to this design as the TBNLC. In the following we present the results of the TBNLC system study and our current estimate of total system efficiency of 50%. We also describe an experimental program to reduce the risk of implementing the RK-TBA concept to a large collider. Demonstrations of the key concepts of the TBNLC design are expected to be completed by 2002 at which time the TBNLC could be seriously considered for a 1.5-TeV NLC upgrade. A schematic of a portion of the low-energy e^- side of the TBNLC is shown in Figure A-1.

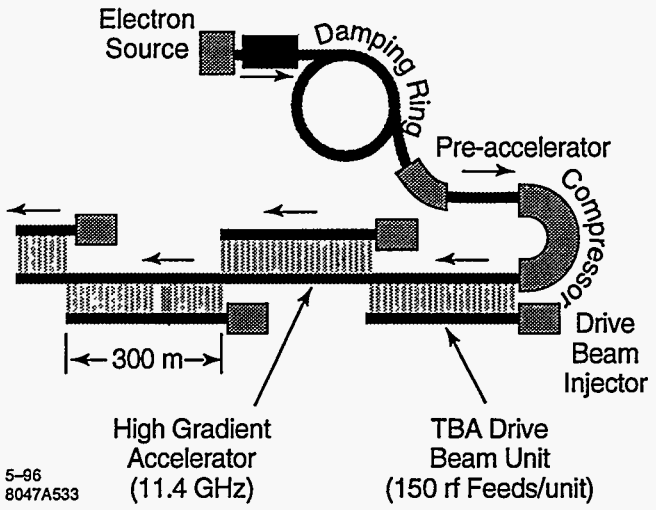


Figure A-1. *Low-energy portion of the e^- side of a linear e^+e^- collider using an RK-TBA design for the rf power source. Our design requires 76 TBA units to provide 1.5 TeV of energy gain.*

A.2 A Design for an RK-TBA-Based rf Power Source

A.2.1 Power Source Requirements

The objective of the TBNLC system study was to construct a conceptual design of the power source for the NLC. Our philosophy was to stay within the design constraints of the NLC and to have a power source system that matches the high-gradient structures being considered for the NLC. However, since the parameters of the NLC were not finalized at the time of the study, we made some choices for power source requirements so that the design and costing studies could be concrete. The design goals for the study were as follows:

Rf frequency	11.424 GHz
Repetition rate	120 Hz
Peak power/structure	360 MW
Distance between extraction structures	2 m
Pulse length (flattop)	200 ns
Pulse rise time	125 ns
Center-of-mass energy	1 TeV

The peak power of 360 MW at 11.424 GHz corresponds to an unloaded gradient of 100 MV/m in the NLCTA high-gradient structures presently tested at SLAC. This power is generated from an extraction cavity in the RK every 2 m, to match the 1.8-m accelerator sections, and to provide adequate spacing in between the HGSs for input and output couplers, focusing, etc.. In the study, we costed the power source for a total of 15 km of this high-gradient structure, 7.5 km for each arm of the collider. There is a total of 7500 rf extraction cavities, with a total peak power of 2.7 TW. The unloaded energy gain from each arm of the collider is 675 GeV. Assuming a loaded gradient that is 75% of the unloaded, final energies of the electron and positron bunches are 506 GeV each.

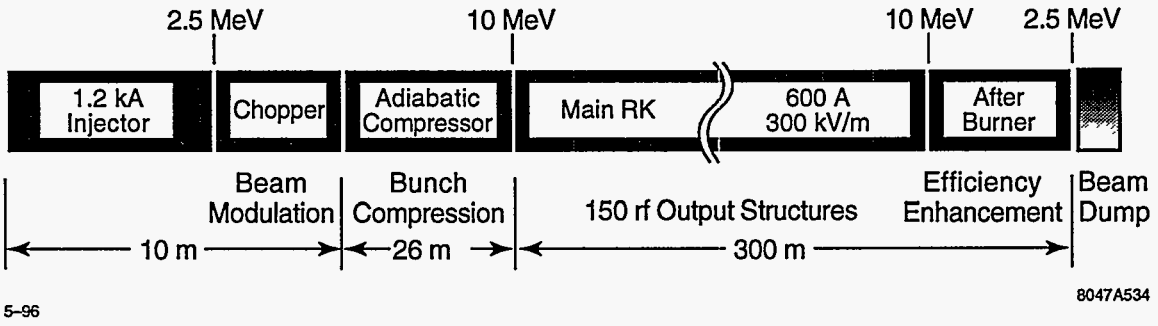


Figure A-2. *Schematic showing major components of the RK-TBA unit.*

The pulse length was determined by the fill time of the structure plus the length of the multi-bunch particle train. To maintain constant particle energy from bunch-to-bunch, the required electric field should grow linearly in the first 100 ns, and remain constant over the remainder of the beam-on time. The flattop is specified as 200 ns, approximately twice the flattop pulse length of the present parameter list (~100 ns flattop), which allows us to accelerate pulse trains of 143 bunches and leads to a corresponding increase in luminosity. This is possible for our TBA design at only modest cost increase, and is a natural extension of the NLC parameter set to take advantage of the RK-TBA structure. Because of the intrinsically high-efficiency of the RK-TBA, the average power required stays at about 200 MW.

The required electric field at the front of the 100-ns rise is roughly 25% of the asymptotic value. Hence, we specify our input pulse to rise linearly (in field) from zero to its full value in 125 ns. The power increases quadratically over the rise time. This pulse shape can be generated in an RK by a corresponding linear rise in the current waveform of the drive beam, and a linear rise in the reacceleration voltage as well. Further refinement of the pulse shape to match dispersion in the high-gradient structure and to improve energy flatness can be achieved by appropriate shaping of the drive beam current and voltage, but were not considered in the TBNLC study.

A.2.2 TBNLC Architecture

To provide rf power for a 1.5-TeV high-gradient linear accelerator, the TBNLC design requires 76 independent RK-TBA units. Each RK-TBA unit is about 300-m long, and has 150 extraction cavities (Figure A-2). To replenish the 360 MW generated from the extraction cavity every 2 m, the 10-MeV drive beam with average current of 600 A is reaccelerated at 300 kV/m. The design current could be raised slightly to compensate for rf losses in the induction cells (*e.g.*, 625 A for 4% loss). The main body of the RK consists of identical 2-m modules each of which has six 100 kV induction cells and one extraction cavity. Both the drive beam current and reacceleration voltage have a rise time of 125 ns and a 200-ns flattop, with a falltime that is comparable to the rise time.

The front end of each RK consists of an injector in which a 1.5-kA beam of electrons is generated and accelerated to 2.5 MeV. A 5.7-GHz chopper is placed at this point to generate a bunched beam at 11.4 GHz. See Section A.5.3 and Figure A-20 for additional information on the modulation of the beam. Chopping reduces the DC current from 1.5 kA to 600 A. This bunched beam at 2.5 MeV is then accelerated to 10 MeV in an adiabatic compressor section. Bunching cavities in this section further reduce the length of the bunches,

and prepare the beam with the right phase space for injection into the main RK. At the end of the RK, there is an afterburner section in which rf extraction continues, but reacceleration is absent. The afterburner allows us to further extract rf power from the bunched beam, and thus enhances the overall efficiency of the RK. At the end of the afterburner, the spent beam (at 2 to 3 MeV) is collected at a beam dump. The adiabatic compressor section is 26-m long, while the main RK has 138 extraction cavities over 276 m. The afterburner consists of 12 rf cavities, making a total of 150 extraction cavities.

The overall length of a RK-TBA unit is determined by a balance of two opposing considerations. The longer the RK-TBA unit, the higher the overall efficiency, as the overhead losses from the front and the back ends become a smaller fraction of the total power. On the other hand, the control of beam instabilities and beam degradation effects become increasingly more difficult as the overall length is increased. Longitudinal and transverse beam dynamics simulations indicate that the stability of the bunched beam can be maintained in the proposed RK-TBA unit.

The efficiency for conversion of power in the drive beam to rf power is easily evaluated for the proposed RK system. Allowing for some rf loss in the induction cavities, the overall efficiency is estimated to be 90%. The 10% loss is shared among the beam loss on the chopper (3.7%), beam dump (2.8%), and rf into induction cavities (3.6%). The efficiency for wall plug to drive beam is 55%. Hence the total efficiency from wall plug to rf is 50%.

The rf power requirement of 360 MW/2 m determines the product of the accelerating gradient and beam current in the drive beam. Our particular choice of 300 kV/m and 600 A is again based on a balance of two considerations. With higher current and lower gradient, the volume of magnetic material required is reduced and the efficiency is increased. However, beam transport becomes more difficult with increasing current.

The linear rise in the drive beam current assures that the extracted rf field has the right waveform for beam loading in the HGS, while the linear rise in the voltage of the reacceleration cells assures that the entire drive beam stays at 10 MeV from head-to-tail. This is important for both beam transport and for rf phase stability. Hence, the rising portion of the current and voltage generate useful rf power, but the fall portion is wasted. The corresponding loss in efficiency due to finite falltime is included when we consider the efficiency of the pulse power system.

A.2.3 Main RK

The main RK consists of identical 2-m modules. The key elements of a module are the permanent quadrupole magnets for focusing of the drive beam, the induction cells for reacceleration, and one rf extraction cavity. Design considerations for each of these major components is discussed below. The key issues in the design are related to the physics of rf extraction, reacceleration, and transport. However, attention was paid to making the module compact and efficient. The resulting unit has a diameter of 35 cm, which is quite small compared to existing induction accelerators. A schematic of the 2-m module is shown in Figure A-3.

Pipe size

In the design of induction accelerators, the size of the beam tube is usually determined by BBU considerations, since the transverse impedance from the acceleration gap is inversely proportional to the square of the pipe radius. In our RK design, the low-frequency BBU arising from the induction gaps is ameliorated by the low-beam current of 600 A and Landau damping associated with the energy spread inherent in the rf buckets.

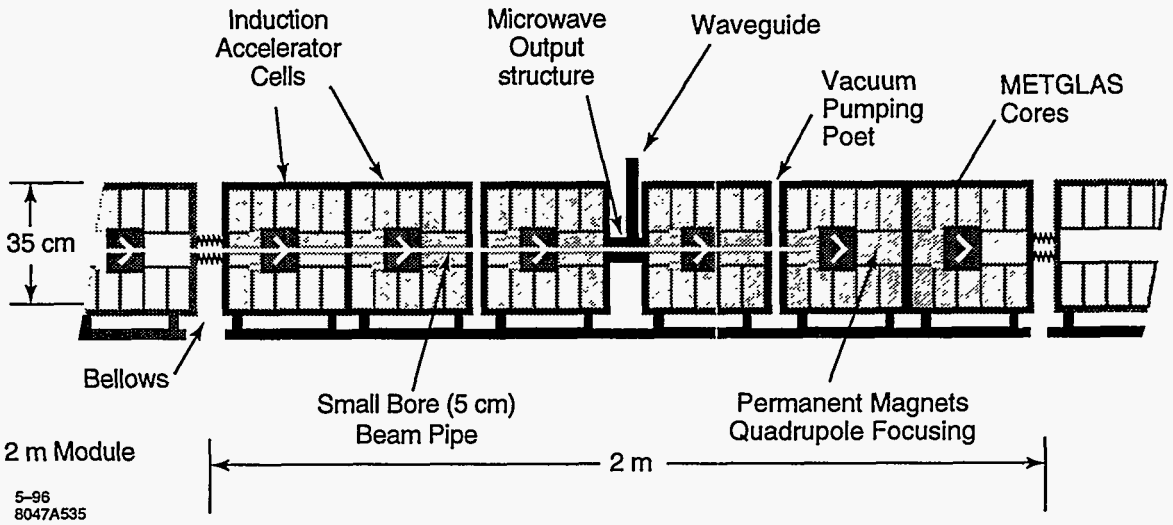


Figure A-3. *Schematic of 2-m RK Module*

With proper attention to the gap geometry, we were able to design a pipe with a radius of 2.5 cm, and predict acceptable BBU growth.

Permanent magnet focusing

The small pipe radius allows us to take full advantage of permanent magnets, which have significant cost advantages when the focusing systems are small. Another major reason for using permanent magnets is the associated efficiency, as they eliminate the need for power supplies.

Several basic considerations govern the design of the FODO lattice. The foremost requirement is associated with a technique concocted to minimize the high-frequency BBU growth arising from the HEM_{11} mode in the rf extraction cavities. To suppress the instability growth, we demand that the betatron period be equal to the distance between adjacent extraction cavities,*i.e.,* 2 m. In addition, the phase advance per lattice period must be less than 90° to ensure beam stability. Finally, the focusing strength must be sufficient to keep the beam envelope small enough to stay within the beam tube.

The zeroth-order design equation is given by the thin lens approximation

$$\sigma_0 = \frac{\eta B L^2}{R[B\rho]} \tag{A.1}$$

where σ^0 is the phase advance per lattice period, and B is the quadrupole field strength at pole-tip with radial position R. L is the half-lattice period, and η the occupancy factor for the quadrupole. The rigidity of the electron beam is given by

$$[B\rho] = \frac{\beta\gamma mc}{e} \tag{A.2}$$

where $\beta\gamma mc$ is the momentum of the relativistic beam, and $[B\rho]$ has the value of 0.035 T-m for 10-MeV electrons.

A first-order correction for deviations from the thin-lens approximation is obtained by replacing the phase advance and occupancy factor with scaled variables given by [Barnard 1991]

$$\hat{\sigma}_0 = \sqrt{2(1-\cos\sigma_0)} \quad \text{and} \quad \hat{\eta} = \eta\sqrt{2\left(1-\frac{2\eta}{3}\right)} \tag{A.3}$$

In our design, we use σ_0 of 60°. There are 12 half-lattice periods in 2 m. The physical occupancy factor is 0.48. The modified thin lens formula then gives a B-field at pole tip of 800 Gauss. This estimate is within a few % of the actual B-field required to give a 2-m betatron period, when we include realistic spatial profile of the magnetic field, as well as the energy variations of the drive beam as it is accelerated over the 2-m module.

A preliminary quadrupole design was constructed by Klaus Halbach [Halbach priv, Halbach 1980]. It consists of four rectangular blocks of ferrites with residual field B_r of 3.2 kg. The dimensions of the blocks are 1.1 cm × 3.22 cm × 8 cm. This design has as the first nonzero harmonic $n = 10$, with 8.6% strength of the quadrupole field at aperture radius. Finite differential permeability effects have not been included in this preliminary design.

A basic requirement of the transport channel is that the focusing strength must be sufficiently strong to keep the beam from hitting the beam tube. The beam envelope in the focusing channel can be estimated from the equilibrium formula

$$\frac{\varepsilon^2}{a^2} = \sigma_0^2\left(\frac{a}{L}\right)^2 \tag{A.4}$$

where $\varepsilon = \varepsilon_n/\beta\gamma$ and a are respectively the unnormalized edge emittance and the edge radius of the beam envelope. With a normalized edge emittance of 600 mm-mr, (which is achievable with a good gun design), we obtain an edge radius of about 2 mm for the beam.

Steering and Focusing Corrections

Steering and focusing corrections are achieved in the design by means of low-field (~10 Gauss) DC coils located in the region immediately outside of the permanent magnets, at radial positions of about 4 cm from axis. These correction coils can be manufactured inexpensively on printed circuit boards.

Since the beam apertures are small, particularly around the extraction cavities, beam centroid displacements must be kept to a minimum. Quadrupole misalignments, when combined with energy variations from head to tail, can lead to beam offsets that change over the length of the pulse, *i.e.*, the corkscrew phenomenon [Chen 1990]. Our strategy for minimizing beam displacements and associated corkscrew is to impose misalignment tolerances that are not excessively tight, design for energy flatness of ≤1%, and to correct alignment errors with closely spaced dipole-steering coils. Detailed steering algorithms have not been worked out yet, but for the purpose of costing, we provide three sets of steering coils and two beam position monitors per betatron period.

A key ingredient for the suppression of high-frequency BBU is that the betatron period be equal to the distance between adjacent cavities. To ensure that this requirement is obeyed, a feedback system with two correction quadrupoles per betatron period is incorporated into the design. A possible monitor for the betatron motion is to use the rf output from the extraction cavities at the dipole frequency. Whether the signal sensitivity is adequate for this purpose needs further study.

Induction Cores

The design of induction cells for reacceleration is based on the relation

$$V \cdot (\Delta t) = A \cdot F_p \cdot (\Delta B) \quad , \tag{A.5}$$

where the core size (with magnetic flux over area A and F_p is the packing factor, or the fraction of the total core that is occupied by magnetic material) is determined by the required voltage V and pulse length Δt. The allowed flux swing ΔB is different for different magnetic material. Our design uses METGLAS®[1] with ΔB of ~1.3 T. Flux swings for METGLAS are several times larger than ferrites, for example, which are commonly used in short-pulse induction machines such as the Advanced Test Accelerator (ATA) [Kapetanakos 1985] and the Experimental Test Accelerator II (ETA II) [Clark 1988]. The cores are correspondingly more compact. Furthermore, METGLAS is quite inexpensive, particularly when bought in large quantities.

To determine the transverse dimensions of the core, (with outer radius R_o, and inner radius R_i), we recast Eq. A.5 in terms of the accelerating gradient G and core occupancy factor η_c (fraction of the axial length occupied by induction cores) as follows:

$$R_o - R_i = G\Delta t/(\Delta B \eta_c F_p) \tag{A.6}$$

where F_p has a value of 0.65 to 0.75 for typical METGLAS cores. Since the accelerating gradient and pulse length are fixed by overall system requirements, our design philosophy was to maximize axial core occupancy. The proposed design has η_c of 0.75 and core thickness $R_o - R_i$ of 10 cm.

Compact cores reduce cost as well as energy loss. Empirical measurements of core losses can be parameterized to give the following phenomenological formula for the core current (in Amps):

$$I_c = 360(R_o + R_i) \cdot (\Delta B/\Delta t) \tag{A.7}$$

where the radial dimensions are in meters, ΔB in tesla, and Δt in microseconds. The fraction of energy lost in the core (I_c/I_{total}) is proportional to the size of the core. Our design parameters give a core current of about 117 A. With a 600-A beam, the core efficiency is 80%. See Section A.4.2, "Induction Accelerator Module", for results of core testing.

The magnetic material in a 2-m module is packaged into six independent 100-kV induction cells. Each cell in turn consists of five 20-kV cores. Packaging into small induction cores provides a natural match to the low-voltage pulse power system. Figure A-4 is an illustration of the proposed induction cell showing the five cores.

Pulsed Power

The desired voltage pulse from each core has a 125-ns linear rise, followed by a 200-ns flattop at 20 kV, followed by a fast fall. This voltage waveform is generated with pulse forming networks (PFNs) with tapered impedance to match the induction core. Power input to the PFNs consists of a DC power supply and a Command Resonant Charging unit. The entire pulse power train is at low voltage, and no step-up transformers are needed. Hence, high efficiency and low cost is possible.

[1] METGLAS® is a registered trade name of Allied-Signal.

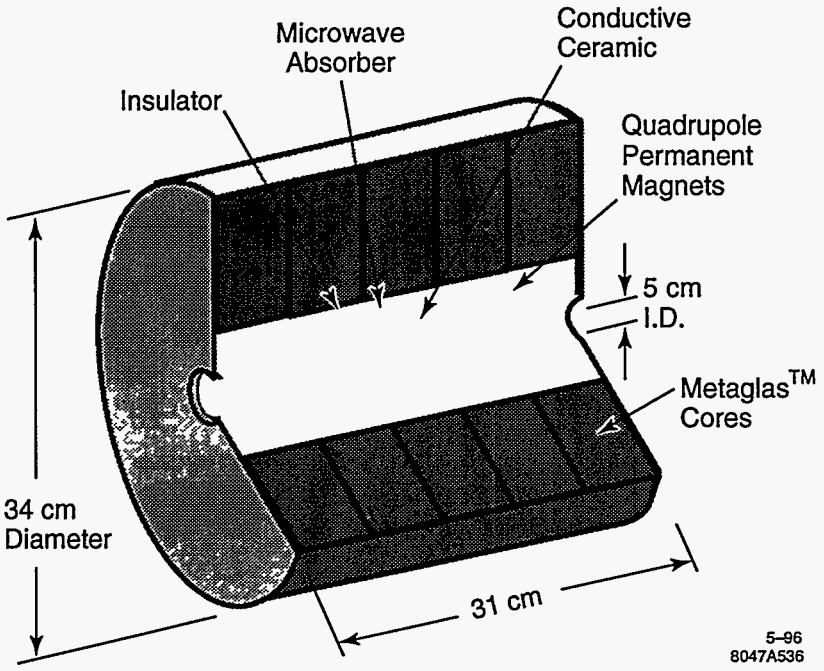


Figure A-4. *Proposed RK induction cell design.*

Acceleration Gap

Each 100-kV induction cell has an induction gap, which is one of the more critical components of the design. Ongoing design efforts evolve around the following key issues:

- **Reduction of high voltage breakdown risks.** To have an induction gap that is safe from breakdown, the gradient across the insulator must be sufficiently low (<30 kV/cm), sufficiently low fields on the metal surfaces (<100 kV/cm), proper shielding of the insulators from secondary electrons and X-rays generated by the beam, and proper design of the triple-point (the interface between insulator, metal, and vacuum).
- **Suppression of low-frequency BBU.** Dipole modes associated with the gap (at a few GHz) can lead to severe problems if not carefully damped. The required transverse impedance was achieved with heavy de$Q$ing by placing microwave-absorbing material at critical locations around the gap.
- **Reduction of transverse and longitudinal impedances at high frequencies.** While the high-frequency BBU (HEM_{11} mode at ~14 GHz) is generated primarily in the rf extraction cavities, and the major BBU suppression activities center around them, one must be careful to ensure that the transverse impedance contribution from the induction cavities are indeed negligible. In the present designs, the transverse impedance around 14 GHz can be made to be lower than half an ohm, and its effect on the high frequency BBU is insignificant. The longitudinal impedance at 11.424 GHz must be low to minimize the microwave power loss. Our design goal is to maintain the induction gap loss to less than 4%, which requires that the longitudinal impedance be less than 2.4 Ω. Present designs are approaching the required impedance.

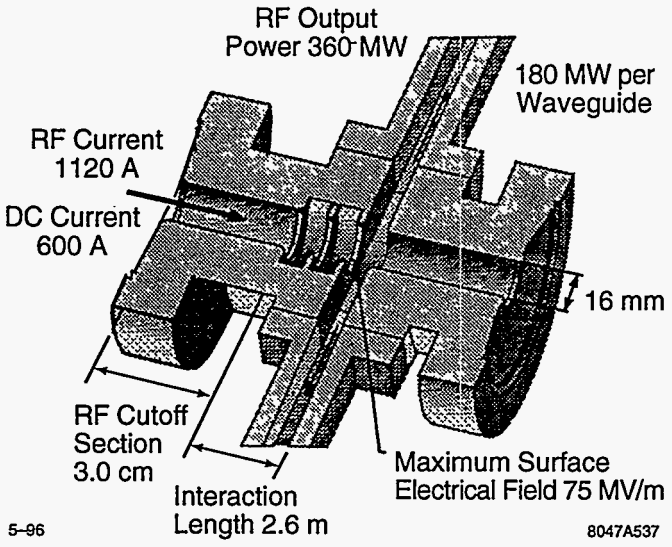


Figure A-5. *Illustration of an rf extraction structure.*

RF Extraction Cavities

Design of the rf extraction cavity is a major ongoing activity in theory and simulations. Present designs evolve around traveling-wave structures with three cells of eight-mm inner radius. The rf output is extracted through two separate ports in the third cell, with 180 MW each transported through separate waveguides, and fed directly into the two input couplers of the high-gradient structure. An illustration of a proposed design is shown in Figure A-5. Detailed design of the cavities centers around five key issues:

- **Required output power.** The extraction cavity must have the right effective impedance of about 540 Ω to extract 360 MW from a highly bunched beam (70° microbunch length) whose first harmonic current at 11.424 GHz is about 1.15 kA.
- **Minimal breakdown risks.** Our present designs use traveling-wave structures with three cells to reduce the surface fields. Surface-field gradients of less than 70 MV/m appear possible.
- **Inductive detuning.** In addition to generating the right amount of power, the rf cavities provide continuous longitudinal bunching for the drive beam. This is accomplished by inductively-detuned traveling-wave-structures.
- **Dipole de*Q*ing.** The rf cavities must have low-transverse impedances for the suppression of BBU. Fortunately, there is a natural de*Q*ing mechanism, as the required extraction cavities have inner radii of around 8 mm, and the dipole modes in the cavity couple to the TE_{11} mode in the pipe. Simulations to date suggest that the resulting impedances are sufficiently low for BBU suppression provided that the Betatron Node Scheme is used.
- **Transverse focusing.** The rf cavities have transverse fields associated with the fundamental mode. They can degrade the beam envelope if they are sufficiently strong and/or not properly corrected. Evaluation of these transverse forces from theory and simulations is ongoing.

A.2.4 Front End and Back End

Each RK unit has a front end that consists of an injector, a chopper, and an adiabatic compressor section. At the end of each RK unit, there is an afterburner, followed by a final beam-dump. These five elements are described below.

Pre-chopper

The injector is basically a short-induction accelerator to generate an unbunched 300-ns electron beam at 2.5 MeV. This section may consist of an induction injector (at 1–1.5 MeV), followed by a short-induction accelerator section that takes the beam to 2.5 MeV. The required current from the injector is 1–1.5 kA, depending on the efficiency of the chopper which is required to produce at its exit a bunched beam with a DC current of 600 A.

The current waveform required consists of a linear rise over 100 ns, followed by 200 ns of flattop. The beam energy should be constant over the entire 300 ns, including the front portion with the rising current. To generate the linearly rising current, the voltage pulse in the 1-MeV injector will also, by the Child-Langmuir Law, have a (nonlinear) 100-ns rise time. The PFN in the subsequent short accelerator section must be arranged to produce a higher voltage at the beam front, thereby compensating for the low-energy at the injector exit.

Since the electron beam will be transported over 300 m of narrow pipes, caution must be taken to minimize head-to-tail energy variations, as well as transverse beam offsets and beam temperature. The injector design should produce a bright beam with a normalized edge emittance of 600 mm-mr or less, and transverse displacement of 200 microns or less.

Chopper

The purpose of the chopper is to generate cleanly-separated microbunches at 11.4 GHz. The basic scheme follows the design of Haimson Research Corporation's Choppertron [Haimson 1989], a 11.4-GHz microwave generator that has been deployed at LLNL. A subharmonic dipole deflecting cavity at 5.7 GHz causes the electron beam to oscillate about a limiting aperture, leading to chopped bunches at twice the oscillating frequency. A similar device can serve as the TBNLC front-end chopper, except that much more effective heat dissipation is required for the 120-Hz operation of the NLC upgrade.

Since a substantial amount of energy is lost on the chopper, there is incentive to make it more efficient. One idea is to precede the subharmonic deflecting cavity with conventional bunching cavities at 11.4 GHz. The role of the chopper then is primarily to clean up the particles with the wrong phase. This requires straightforward phasing of the 5.7-GHz input power to the deflecting cavity relative to the incoming prebunched beam.

Adiabatic Compressor

The exiting beam from the chopper has microbunch lengths equivalent to 180° or greater in longitudinal phase space. In the adiabatic compressor region, the microbunch lengths are further reduced to 70°, which is the needed bunch length for long-distance propagation in the main RK. This is accomplished by a number of idler cavities that are more inductive than the rf extraction cavities in the main RK.

In addition to microbunch sharpening, this section also serves to provide the energy transition from 2.5 MeV to 10 MeV. The induction core structures are very similar to the main RK. However, the voltage waveform should be flat over the 300 ns of beam-on time, so that the head-to-tail energy flatness required for phase stability could be maintained at 10 MeV.

The quadrupole magnets are weaker at the lower energies, and continues to increase with increasing energy, so that the betatron wavelength is kept fixed at 2 m. Structurally, the adiabatic compressor section looks very similar to the main RK, except that no power is extracted from the idler rf cavities.

Afterburner

At the end of the RK is an afterburner section, the primary purpose of which is to increase overall system efficiency by extracting more power out of the bunched beam at the end of the main RK. This section has a number of rf extraction cavities, permanent magnets for focusing, but no reacceleration cells. The spacing of the rf cavities is changed as the average energy of the beam continues to decrease. The impedance of the cavities is also modified to compensate for changes in the rf bucket. The TBNLC design has 12 cavities in the afterburner section, each generating 360 MW, and together covering a total of 12 m.

Beam dump

As the spent electron beam reaches the final beam dump, its average energy is less than 3 MeV. The design of this component is straightforward.

A.2.5 The Key Ideas

A few key ideas underlie the whole design of the relativistic klystron for the TBNLC and are summarized as follows:

- **Betatron Node Scheme for high frequency BBU control.** The most severe BBU instability is associated with the HEM_{11} transverse mode in the rf extraction cavities. Strong suppression of this mode is achieved by introducing a Betatron Node Scheme [Li 1994] in which adjacent rf extraction cavities are placed exactly one betatron period apart. This scheme minimizes beam centroid displacement which excites the transverse mode, and alters the nature of the instability from exponential to a slow secular growth.
- **Landau damping for low-frequency BBU suppression.** Quite apart from the high-frequency dipole mode associated with the rf cavities, there is a low-frequency (a few GHz) dipole mode associated with the induction reacceleration gaps. This BBU instability is ameliorated by Landau damping due to a large energy spread inherent in the rf buckets of the bunched drive beam. Combining this feature with dipole mode suppression measures in the induction gap design, the calculated low-frequency BBU growth is minimal.
- **Inductively detuned cavities for longitudinal beam stability.** To maintain tight rf bunches over long distances with multiple extraction cavities, the rf output structures are inductively detuned. While the concept of inductive detuning is not new, the theoretical framework has to be developed and implemented in simulation codes for traveling wave output structures. Cavity parameters required for stable beam propagation through multiple structures were determined in the TBNLC system study.

- **Small low-cost induction cells.** The induction cells in the TBNLC design are about one foot in diameter, much smaller than most of the existing induction machines. This design was possible because of three key features:
  - Use of METGLAS for the induction cores. This is a low-cost magnetic material that can accommodate a large flux swing.
  - Use of low-field permanent magnet quadrupoles. The relatively high drive beam energy of 10 MeV (compared to klystron beams) is a natural match to a strong focusing transport system with low-cost ferrites. An additional advantage of permanent magnets is, of course, the elimination of power supplies.
  - Narrow beam pipes—the design pipe diameter of 5 cm is much smaller than other existing high current machines. This is possible because of the relatively low current (by induction machine standards) of 600 A, and the low-frequency BBU suppression features described earlier. Large beam pipes in usual induction machines are dictated by BBU considerations.
- **Low-voltage pulse power system for efficiency and cost.** The induction cell design matches naturally to a 20-kV system with PFNs triggered by ceramic thyratrons and powered by DC power supplies and command resonant charging systems. This system does not require step-up transformers and eliminates losses associated with a high-voltage system.

A.3 TBNLC Physics Studies

The TBNLC system study addressed the issues of longitudinal and transverse stability of the drive beam and the related areas of rf extraction cavity and induction cavity designs.

A.3.1 Longitudinal Beam Dynamics

Longitudinal stability of the drive beam is among the important issues for the demonstration of technical feasibility of the RK-TBA concept. In the TBNLC design, the drive beam is required to stay bunched longitudinally over 150 extraction cavities. Space charge effects cause initially tight bunches to expand. The debunching process is further aggravated by the energy spreads along the bunches as they interact with rf fields. In usual traveling wave extraction cavities, rf waves are in synchronism with the drive beam and debunching becomes very severe after a few cavities. Beam debunching, if uncompensated, will result in reduced power extraction in subsequent cavities. To overcome this problem, we employ a scheme in which the drive beam is not synchronous with the operating wave mode; more specifically, the phase velocity of the rf field is larger than the speed of light so that the bunches always lag behind the wave. This is what we call inductive detuning and can effectively bunch the beam. The concept of inductive detuning is well known for standing wave structures (SWSs), *e.g.*, the penultimate cavity in a klystron, in which the frequency of the cavity is detuned from the resonant frequency. For the traveling wave structures (TWSs) of our RK, the approach we take is to keep the frequency ω of the operating mode unchanged, but reduce its wave number k, such that the wave field advances faster than the beam. In this case particles at the front of the bunch lose more energy and slow, while particles at the tail of the bunch lose less energy and catch up. This mechanism causes a continuous sharpening of the bunch, thus counteracting the debunching forces. The resulting longitudinal phase space continues to rotate in a stable rf bucket with a relatively stable bunch length. Thus both constant power extraction and beam bunching can be achieved simultaneously.

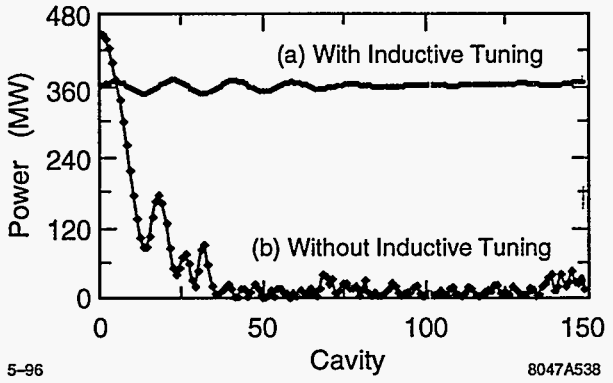


Figure A-6. *Power extraction from 150 cavities in one unit of RK-TBA: (a) With inductive detuning (Δ = 30°, phase velocity 1.33 c); (b) No inductive detuning (Δ = 0°, phase velocity 1.0 c).*

Parameters for the detuned rf output structures were estimated from analytical expressions developed during the TBNLC system study. Numerical simulations were then performed using the relativistic klystron code RKS [Ryne 1990]. This code solves self-consistently the single-particle equations of motion for the beam and the coupled circuit equations that govern the cavity excitation, and it includes the calculation of the space charge effect. It assumes a single-dominant mode and cylindrical symmetry of its fields inside the cavity.

Figure A-6 presents the output power from each of 150 TWSs in the main RK section for both a successful inductive detuning case and its corresponding synchronism case. The parameters for the inductive detuning case are given in Table A-1. For the synchronism case, the level of the extracted power P_{out} declines sharply due to the space charge debunching effect as the drive beam traverses the RK. In contrast, when the TWSs are properly detuned, the rf bucket can remain stable and output power can be sustained at the desired level (~360 MW) for the 150 extraction cavities.

The synchronism case in Figure A-6 consists of conventional three-cell TWSs operating at the $2\pi/3$ mode. In the inductively detuned case the operating detuning angle is 30°. The cavity is therefore operating in a $\pi/2$ mode (*i.e.*, phase advance of the field across one cell is 90°). The longitudinal dimension of each cell is the same in the two cases while the transverse dimensions are varied. URMEL and MAFIA codes were used for detailed cavity design.

A key feature for RK design is the cavity filling time, *i.e.*, the time it takes for the rf field in a cavity to reach equilibrium state. In Figure A-7, we present the time dependencies of output power at the 50th, 100th and 150th extraction cavities for the inductive detuning case in Figure A-6. It shows that after about 15 ns, the fields in all the cavities reach their equilibrium states. This indicates that the erosion on the beam head due to the cavity filling process is not serious. The short fill time is a result of low Q and high V_g.

Further RKS code simulation studies were conducted to examine the sensitivity of the inductive detuning scheme to the important parameters of bunch length and detuning angle. For 30° of detuning, the scheme was insensitive to variation in bunch lengths from 50° to 90°, although performance deteriorated for lengths above 90°. Similarly, the scheme was insensitive to increasing the detuning angle from 30° to 35°, but performance deteriorated for angles below 30°.

The issue of radial defocusing, or emittance growth, due to the interaction of the beam with the rf fields in the extraction structures is an active area of study. The underlying theory of the RKS code involves power balance. Detailed rf field information is less stringently modeled. This approach works satisfactorily for standing wave and synchronous traveling wave structures. However, for detuned traveling wave structures,

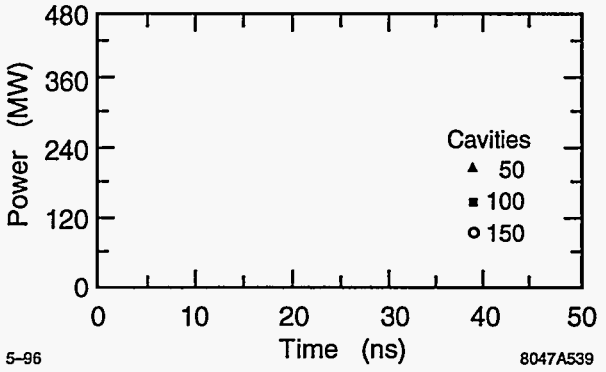


Figure A-7. *Time evolutions of output power at 50th, 100th and 150th TWSs for the nonsynchronism case in Figure A-6.*

Drive frequency	11.424 GHz
Forward traveling mode	TM_{01}
Number of cavities	3
Phase shift per cavity	90°
Wave length	2.626 cm
Phase velocity	1.33 c
Group velocity	0.28 c
Shunt impedance per cell $(R/Q)^*$	27.0 (Ω)
Eigenfrequency for the first 2 cells	11.424 GHz
Eigenfrequency for the 3rd cell	11.666 GHz
Wall-dissipation quality factor	7000
External quality factor for the 3rd cell	6.5
Aperture inner radius	8 mm
Aperture outer radius	12.5 mm
Iris thickness	2.5 mm
Longitudinal dimension of each cell	8.754 mm
Beam energy	10 MeV
Beam current (peak)	600 A
Bunch length	0.51 cm
Beam radius (rms)	2.5 mm

*Traveling-wave shunt reactance is twice the standing-wave shunt reactance.

Table A-1. *Parameters related to the inductive detuning case.*

the RKS code does not adequately account for radial forces on the beam limiting its usefulness for extended structures to 1-D simulations.

A.3.2 Transverse Beam Dynamics

The excitation of transverse beam instabilities due to higher order modes has been identified as a major issue in the design of a long multicavity RK-TBA. The narrow aperture and high average current of the RK accentuates the problem. There are two separate structural components that contribute to the transverse instability. The induction module with an aperture of 5.0 cm has a trapped dipole mode around 3 GHz and the output structure with a 1.6-cm aperture has a transverse mode near 14 GHz. Both modes can interact strongly with the beam. The effect of these structures on transverse instability can be studied separately due to the difference in resonant frequencies.

Low-Frequency BBU

There are three induction modules per meter through the RK. Although considerable damping can be accomplished by the insertion of absorbing material, design constraints imposed by maximum surface electrical fields and vacuum requirements preclude complete damping of higher order modes. Modeling the module's gap as a cylindrical resonator with an impedance boundary condition on the outer wall [Brigg 1985] and assuming an optimum design, the transverse impedance, $Z_{\perp 1}$, will be on the order of 4,000 Ω/m.

Analytical theory [Panofsky 1968] indicates that the transverse instability will grow exponentially along the length of the accelerator. Experience with ATA and ETA II indicate a growth of 4 to 5 e-folds from noise is tolerable in the transverse instability. For a 10-MeV e-beam, using a betatron wavelength of 2 m, a current of 600 A, and gap spacing of m/3, a theoretical estimate of the maximum tolerable transverse impedance of the module, $Z_{\perp 1}$, is about 573 Ω/m. Thus, additional measures are needed to suppress the transverse instability.

The most promising technique for suppressing the transverse instability is Landau damping. To maintain longitudinal equilibrium, inductively detuned rf extraction structures are used. The resulting rf buckets have an intrinsic spread in beam energy over the microbunch on the order of ±7.5%. Hose instability theory [Lee 1978, Houck 1993] can be used to estimate a maximum $Z_{\perp 1}$ of about 5,178 Ω/m for total suppression of the low-frequency BBU growth. Theoretically, we can build a suitable induction module to avoid BBU.

The OMICE code [Houck 1992b] was used to model the growth of the transverse instability. In Figure A-8, relative growth in the displacement of the beam's centroid from axis as a function of distance along the axis is shown for several different energy spreads. The finite number of beam slices used in the OMICE code did not permit the level of phase mixing available in a physical beam. Thus, the simulations presented provide a conservative upper bound. Table A-2 lists additional input parameters used in the OMICE code. The seed perturbation for exciting the instability was a sinusoidal oscillation of the centroid's transverse displacement at the instability frequency. The conclusion from the numerical modeling is that the instability growth will be acceptable for $Z_{\perp 1}$ less than or equal to 5,400 Ω/m and energy spreads of ±2.5% or greater.

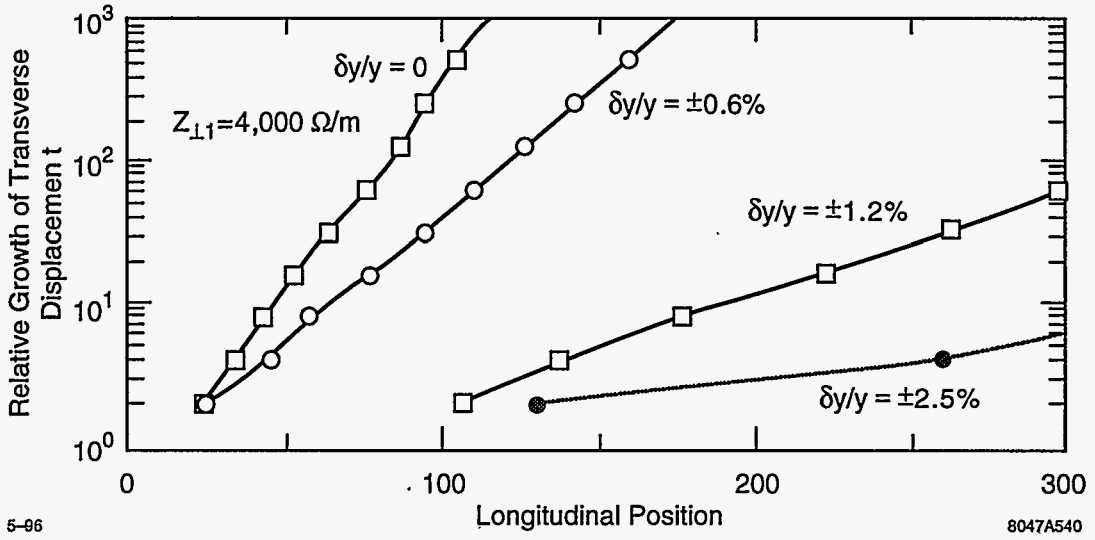


Figure A-8. *Effect of energy spread on the low-frequency instability.*

BBU frequency	3.0 GHz
$Z_{\perp 1}$ (Ω/m)	5,400
$Z_{\perp}/Q$ (Ω)	42.972
Cell Q	2
Gap width	2.5 cm
Average beam energy	10 MeV
$\Delta\gamma(\gamma_{max} - \gamma_{min})$	3
Current (DC component)	0–600 A in 100 ns
	600 A flattop for 200 ns
Focusing system	Quadrupole - 2-m period
Time step	1/600 ns

Table A-2. *Input parameters for simulated induction cell gaps.*

High-Frequency BBU

There is one three-cell, traveling-wave, output structure (TWS) every 2 m through the RK. In the context of the TBNLC design, we limit the growth of the transverse beam displacement to 4 to 5 e-folds over the 300-m system, or 150 TWSs. To keep BBU to this level, we need to minimize and/or suppress excitation of the higher order (transverse) modes in the TWSs.

The contribution of the output structures to the transverse instability of the beam is greater than that of the induction cell gaps for several reasons. The output structure has a smaller aperture to obtain the desired longitudinal shunt impedance. Damping of the higher order modes must not affect the fundamental mode used for power extraction. Also, the three cells comprising the TWS are electromagnetically coupled. Within a TWS, the regenerative BBU mechanism will increase the interaction of the higher order modes with the beam.

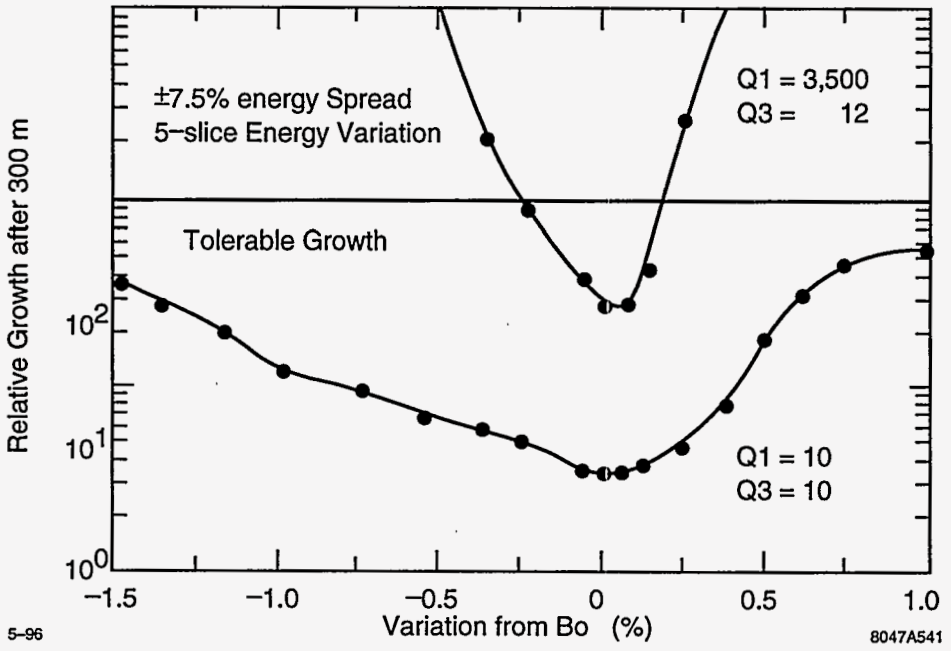


Figure A-9. *Relative growth after 150 TWSs vs. variation from the optimum focusing field. Comparison of an aggressively lowered Q to a high-Q for the first cell.*

The basic design of the output structure has a transverse shunt impedance ($Z_\perp/Q$) of about $3\,\Omega$ per cell. This is significantly lower than for the induction cell gaps. However, the Q_{wall} of the cells is high, on the order of several thousand. The third cell has two output ports that remove energy from the higher order modes. This produces an effective Q_{ext} of about 15. In addition, fields in the first and third cell couple strongly to the TE_{11} mode of the connecting pipe, thus leading to a drastic reduction of Q in both of these cavities. The growth in the transverse instability is also reduced by Landau damping due to the energy spread on the beam. However, even with these instability reduction features, the high-frequency BBU growth is still too high.

The Betatron Node Scheme [Li 1994] was used to limit the growth in the transverse instability to an acceptable level. Figure A-9 shows the large difference in BBU growth as we vary the focusing field from the optimum for the Betatron Node Scheme. This is especially true for the case with a high-Q value for the first cavity. A list of the pertinent variables, and their base values, is given in Table A-3. The Betatron Node Scheme works well in principle; the key issue is the robustness of the scheme to deviations in the betatron wavelength resulting from errors in focusing field and/or beam energy, as well as sensitivities to cavity parameters.

The main effort of the numerical studies was to delineate the acceptable parameter space. The OMICE code was used to model the time-dependent transverse behavior of a 300-ns pulse over the 300-m length of the RK. An initial constant offset of the beam centroid was used as the seed perturbation for exciting the instability. Base parameters were varied individually to characterize the system's sensitivity to different design parameters such as effective impedance, Q values of the cells, quadrupole focusing field strengths, and variation in average beam energy.

The conclusion of this parameter study was that the Betatron Node Scheme is a promising means for controlling the high-frequency instability caused by these structures. Requirements on energy flatness and

Transverse mode	HEM_{11}
Frequency of mode ω_o	14.1 GHz
Number of cells/TWS	3
Electrical length of each cell L_c	8.754 mm
Phase advance/cell ϕ	120°
Group velocity v_g	0.25c
Q_{wall} (1st and 2nd cells)	3,500
Typical Q_{damped}	15
$Z_\perp/Q$ for each cell	3 Ω
Current (DC component)	0–600 A in 100 ns 600-A flattop
Pulse length	300 ns
Average beam energy	10 MeV
Quadrupole field	0.0812 T

Table A-3. *Base parameters for simulated traveling wave structures.*

field accuracy are quite acceptable (±1%) provided that the output structures are indeed strongly de$Q$ed. Feedback schemes may further relax the requirements on field and energy accuracy.

A.3.3 RF Extraction Cavity Design

The most basic requirement for the cavity design is that it generate 360 MW of rf power for a drive beam current of 600-A DC and 1–1.5-kA rf current. The cavities must also be inductively detuned to maintain longitudinal beam stability over long distances. The surface fields of the cavity must be sufficiently low to avoid breakdown. Finally, the cavity must have low transverse shunt impedance in order to minimize the high-frequency BBU. The required cavity parameters have been discussed in Sections A.3.1 and A.3.2. In this section, we present electromagnetic calculations to determine the structure of the extraction cavities. We chose in this design effort one specific path to meet the general requirements. The procedure adopted is by no means unique, and we anticipate further optimizations and more detailed calculations in the future.

The numerical tools we have used are URMEL and SUPERFISH for 2-D frequency domain calculations, ABCI and TBCI for 2-D time-domain calculations, and MAFIA for 3-D frequency as well as time-domain calculations. The design procedure is carried out in several steps, starting with the simplest approximations, and adding more realistic features with each successive iteration. The successive approximations are summarized as follows:

- Step 1. We construct a three-cell, disk-loaded, synchronous traveling-wave structure. Desired power extraction determines $(R/Q)/v_g$. URMEL is used to study an infinitely periodic structure, the inner radius a and outer radius b of the structure is varied, and the set of solutions with $v_p = c$ is obtained following the procedure of Thompson *et al.*, [Thompson 1993].
- Step 2. The inductively-detuned structure is constructed by a variation of Step 1. We want the resonant frequency to remain unchanged, but the wavelength increased by a factor of 1.33. The R/Q and v_g for the new configuration are determined with URMEL. The geometry is adjusted to ensure that the R/Q and v_g provide the right power extraction.

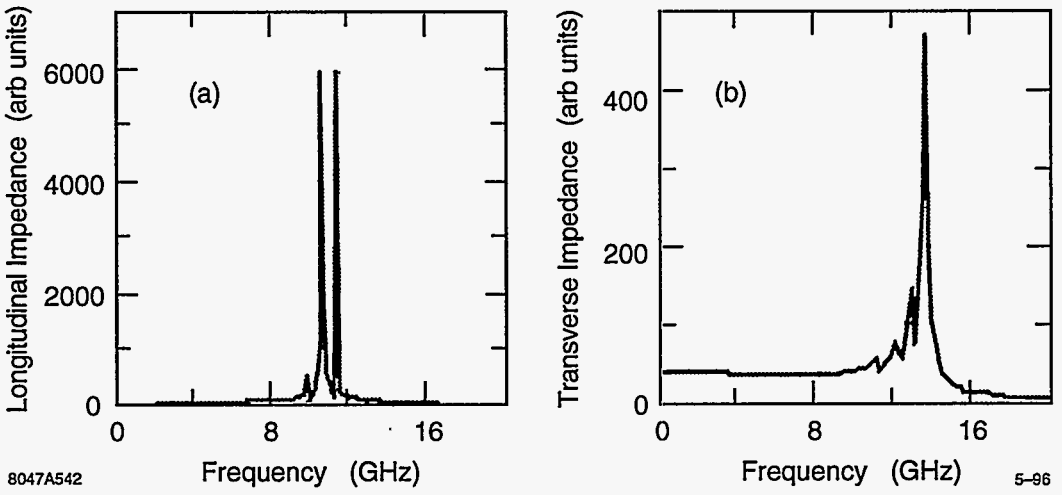


Figure A-10. *Longitudinal and transverse impedances of a three-cell structure with beam pipes.*

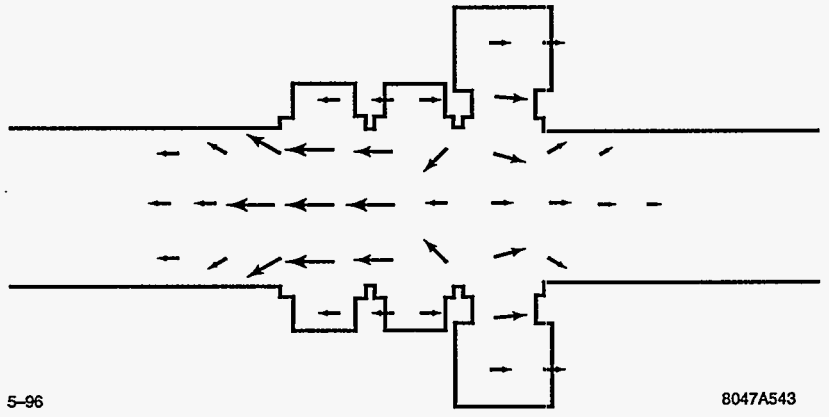


Figure A-11. *Field configuration for the $\pi/2$ mode of the 3-D cavity with output structure.*

- Step 3. Finite beam pipes are then included to model the finite cell structure. The effect of the modified geometry on the field configuration and cavity parameters is then studied.
- Step 4. The 3-D aspects of the output ports are studied using MAFIA. The geometry of the output is varied to achieve the value of Q and ω for proper matching.

The relevant dipole cavity parameters for BBU considerations and the field enhancement factor for assessing the surface field are also determined.

Longitudinal and transverse impedances for the proposed extraction structure are shown in Figure A-10. The field pattern from a MAFIA run is shown in Figure A-11. The pertinent parameters of the structure are given in Table A-3. The determination of the external Q of the 3-D cavity with output structure is based on the Kroll-Yu method [Kroll 1990]. A $Q_{\text{ext}} \approx 80$ was calculated for various waveguide iris apertures. The theory in Section A.3.1 was used to estimate the Q of individual cells. Initial estimates suggest that the transverse mode Q for the first and last cavities is as low as 10.

Material	ε'	ε''	$\mu\prime$	μ''
Ferrite[a]	13.0	0.0	1.0	9.4
TDK[b]	25	3.3	1.9	2.1
Carbon-loaded ceramic[c]	5.9	1.07	1.0	0.0

[a]Ferrite is modeled after the ferrite used in the ATA induction cells [Deford 1990a], measured at 1 GHz.

[b]TDK Electronics Co., LTD., Absorber IR-B006, measured at 2.5 GHz.

[c]University of Maryland ceramic [Calame 1991], carbon concentration 0.34% (by weight), measured at 9.9 GHz.

Table A-4. *Electrical properties of simulated microwave absorbers.*

A.3.4 Induction Cavity Design

The interaction of the induction cell gap with the beam is a critical issue for the RK. There are three major criteria that the cell gap must meet: hold-off of the applied 100-kV voltage, low transverse impedance for BBU minimization, and low longitudinal impedance at the beam modulation frequency and harmonics to minimize power loss. The gap should be as narrow as possible consistent with the maximum surface electrical fields. Larger apertures reduce the transverse impedances, but increase the core volume. An aperture of 5 cm was chosen as a base design. This allows sufficient room for focusing magnets while meeting the desired core volume.

Microwave absorbing materials can be used to damp resonant modes. These materials have a complex permittivity and/or permeability. Table A-4 lists nominal values of the permittivities and permeabilities for the absorbing materials used in the numerical modeling. The permittivity and permeability are expressed as $\varepsilon = \varepsilon_o(\varepsilon' + i\varepsilon'')$ and $\mu = \mu_o(\mu\prime + i\mu\prime\prime)$, respectively. Here ε_o and μ_o are the free space permittivity and permeability.

Several geometries were considered for the gap design. Several variations to each geometry was made to study the effect on the gap impedance. Simulated material properties of the absorbers and insulator were varied over a nominal range as well as their location and size. POISSON was used to adjust the geometry for specific designs to achieve acceptable surface electrical fields. The induction cell design code AMOS [Deford 1989, Deford 1990b], was then used to determine the rf characteristics of the gap. AMOS simulations include power loss to absorbing material. All geometries considered met the transverse impedance requirement of $Z_{\perp 1} \leq 5,400\Omega/\mathrm{m}$ for transverse stability by aggressively damping all resonant modes.

A conventional design similar to the ETA II induction cell gap was chosen as the base design for engineering and costing purposes. Advantages of the conventional design include demonstrated performance on several induction accelerators, a ceramic insulator hidden from the electron beam thus lowering the susceptibility for arcing, and relatively low-surface electrical fields. The disadvantages are a large inner radius for the core and a low-gap capacitance. A low-gap capacitance is desired for fast rise times, but tends to be related to a high-longitudinal impedance which is undesirable. Figure A-12 displays a full-scale schematic of this design. Longitudinal and transverse impedances are shown in Figure A-13.

The most difficult design problem for all geometries studied was achieving a low-longitudinal impedance. The goal is for a power loss to the induction cells of $< 4\%$ of 360 MW per two-m module. This requires a $Z_{L0} < 2.4\Omega$ at 11.424 GHz. While impedance decreases rapidly above the cutoff frequency, the lowest Z_{L0} achieved thus far is about 10Ω at 11.4 GHz. Determining the longitudinal impedance accurately at 11.424 GHz is difficult. Analytical theory is available [Chattopadhyay 1990], but it is not clear that the quantitative accuracy is adequate. Numerically, the parameters of the test charge used, fineness of the grid,

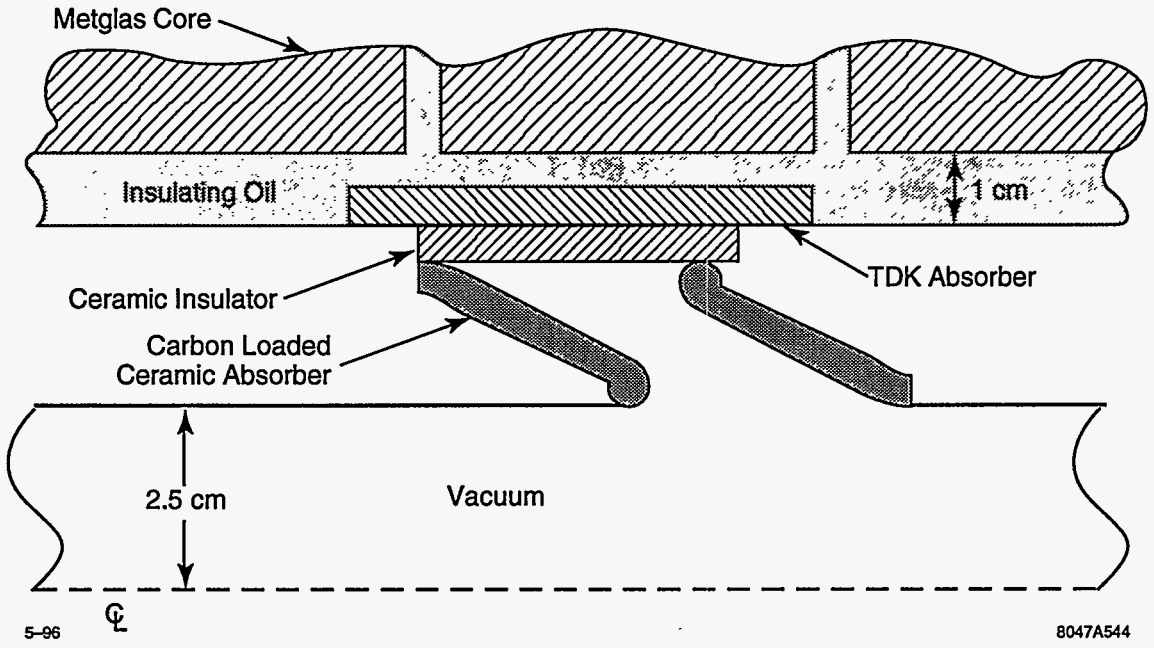


Figure A-12. *Schematic of a conventional design induction cell gap.*

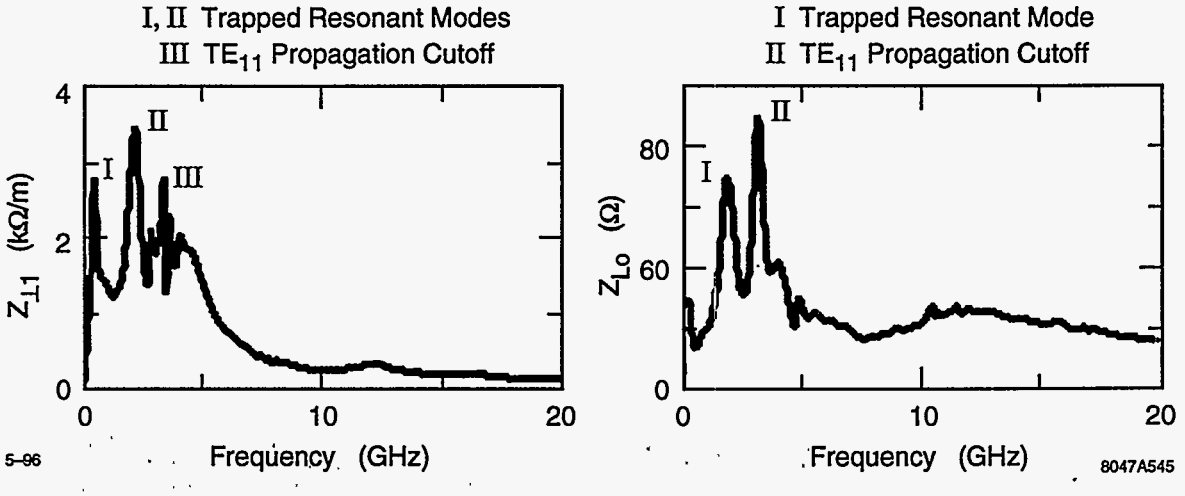


Figure A-13. *Longitudinal and transverse impedance of the induction cell gap.*

and the boundary conditions significantly impact the results. A cold test model needs to be fabricated and tested in the laboratory for a definitive answer.

A.4 TBNLC Engineering Design

A.4.1 TBNLC Systems Study Approach

The objective of the TBNLC system study was to assess the technical feasibility as well as the cost and efficiency of a TBA-based system. To achieve this goal, we felt that it was essential to be reasonably concrete. For this reason, our study specifically addressed a TBA power source for the NLC Upgrade considered by SLAC. The present thinking about the NLC is a 14.16-km active length linac with a center-of-mass energy of 500 GeV powered by X-band klystrons and rf pulse compression systems. This machine should be upgradeable to at least 1-TeV-c.m. Our study considered the TBA as a power source candidate for the 1-TeV version of the NLC. However, our modular architecture is directly applicable for a 1.5-TeV, ~22-km NLC upgrade.

The engineering and costing efforts in the study drew heavily from recent work in Heavy Ion Fusion (HIF) [Hogan 1992]. The U.S. effort in HIF is based on the induction accelerator approach. Research activities in the past few years have centered around the Induction Linac Scaled Experiments [Fessenden 1992] (ILSE). Substantial engineering activities were committed to the ILSE CDR, and a 2-MV heavy-ion injector at full-driver scale was constructed and successfully operated in 1993. An engineering and costing study was performed recently for a recirculator version of the induction fusion driver.

A first engineering and costing exercise for the full TBNLC system was performed. The electrical design included all components starting from the AC power distribution system, to the DC power supplies, command resonant charging system, PFNs, and induction cores. Racks and installation, as well as instrumentation and control, were included in this exercise. The mechanical design and costing included details of the induction cells, rf structures, vacuum system, alignment, and utilities. Important aspects of the engineering design are summarized below.

A.4.2 Electrical Systems

A significant factor in the total efficiency of the TBNLC rf power source is the conversion efficiency of wall plug power into induction beam power. Figure A-14 is a schematic of an equivalent circuit of an induction accelerator cell. I_C is the core magnetizing current, I_B is the beam current, and I_N is any compensating network current. For optimum efficiency, I_N is minimized by designing the impedance of the PFN to match the nonlinear impedance of the induction core. The efficiency of an induction accelerator can approach 100% if the beam current is much greater than the current required to magnetize the transmission line (or autotransformer) which forms the induction cell. For example, the ATA induction cell required less than 1 kA of magnetizing current while accelerating a 10-kA beam by 250 keV for 70 ns; this corresponds to an induction cell efficiency of 91%.

The efficiency of an RK-TBA induction accelerator will depend on a number of factors. Beam transport dynamics will determine the size of the beam pipe and accelerating gradient. The rf power requirement will

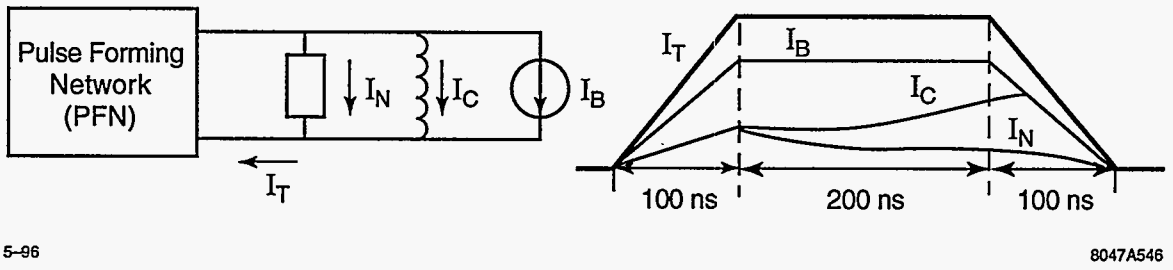


Figure A-14. *Equivalent circuit of an induction cell.*

determine the pulse duration, beam current, and repetition rate. Once these factors are set, the outer radius and flux swing of the core can be calculated from Equ. A.6.

Induction Accelerator Module

The linear induction accelerator can simply be described as multiple transmission line transformers driven in parallel by a pulse modulator which add energy to the electron beam in series as it passes through them. In order for the acceleration process to be efficient, the transformer or induction losses should be small compared to the energy imparted to the beam. Simply put, the transmission line current should be small compared to the beam current; therefore, the effective impedance of the transmission line, Z_{eff}, should be greater than the beam impedance $Z_b = V_{\text{drive}}/I_{\text{beam}}$. In previously constructed induction accelerators, the high impedance is achieved by loading the transmission line with a ferri- or ferromagnetic material which is appropriate for the pulse duration. This material increases the impedance by $\sqrt{\mu/\varepsilon}$ and the electrical length by $\sqrt{\mu\varepsilon}$.

Typically, for induction accelerators with pulse durations of less than 100 ns, Ni-Zn ferrites have been used for the magnetic material due to their relative low saturation losses at very short pulse durations. However, the flux swing for Ni-Zn ferrite is only about one-fifth of that for Ni-Fe or METGLAS, and these materials are preferred for pulse durations of several hundred nanoseconds to microseconds. The optimization process in selecting the type of material and quantity is based on pulse duration, impedance, magnetizing losses, and economics. For the TBNLC design, the parameters are as follows:

$$I_{\text{beam}} = 600\,\text{A}\,, \quad V = 300\,\text{kV}\,, \quad T = 300\,\text{ns}\quad .$$

To satisfy the requirement that the effective impedance, Z_{eff}, is greater than the beam impedance, Z_{beam} ~500 Ω, we can use the equation for a coaxial transmission line where $Z = \sqrt{\mu/\varepsilon}60ln(r_o/r_i)$. For ferrites, $200 < Z_{\text{eff}} < 1000$, and for ferromagnetic materials $500 < Z_{\text{eff}} < 5000$ depending on the type of material used and the magnetization rate. To satisfy the pulse duration requirements, the transmission line length should be about a single transit time. Since transit time is proportional to $\sqrt{\mu\varepsilon}$, for ferrites and ferromagnetic materials $T_{\text{eff}} \simeq 100\ T_{\text{vac}}$. The design of the transmission line (induction cell) must also ensure that no portion of the line is driven into saturation. This requirement dictates the outer radius of the line once the inner radius is determined by beam transport physics.

One of the most important considerations in the selection of a magnetic material for an induction cell is the actual losses during the magnetization pulse. Ferromagnetic materials such as nickel-iron have been available in thin ribbon cores since the late 50s when the first induction accelerator, the Astron, was constructed at Livermore by N.C. Christofilos. These iron-based alloys are very competitive in magnetic properties to the

amorphous materials manufactured by Allied Signal and referred to as METGLAS. The time-dependent losses in ferromagnetic ribbons are typically explained in terms of a saturation wave which encircles the tape and proceeds toward the center at a rate proportional to the applied voltage. The magnetic intensity required to change the state of magnetization is

$$H_a = H_c + (d^2/4\rho)(\Delta B/2B_s)(\Delta B/\Delta t) \quad , \tag{A.8}$$

where H_c is the DC anisotropy coercive field, d is the ribbon thickness and ρ is the resistivity. The energy density deposited in the tape during saturation is given by

$$E_L = H_c\Delta B + (d^2/4\rho)(\Delta B^2/2B_s)(\Delta B/\Delta t) \quad . \tag{A.9}$$

At high-magnetization rates, one can see that the losses are proportional to the thickness of ribbon squared and inversely to the resistivity. Since the ΔB of the iron alloys is similar to that of METGLAS, but the resistivity of METGLAS is three times that of the iron alloys, the ribbon thickness must be thinner in order for the losses to be comparable for METGLAS. The Ni-Fe alloys are manufactured by a rolling technique which yields a more uniform cross section while METGLAS alloys are manufactured by rapidly quenching a mixture of silicon and iron which remains amorphous. The quenching process leaves the ribbon surface rough compared to the rolled ribbon, hence it is more difficult to insulate and can yield an uneven cross section in the core winding process. Currently, the unannealed 20-μm METGLAS is wound into cores with 2.5-μm mylar between layers yielding a packing factor of 0.7 to 0.8 for a finished core. The iron alloys can be rolled into thinner ribbon than the METGLAS, and have produced lower overall losses. However, the manufacturing and winding process for METGLAS usually results in less expensive cores.

The induction cell will be driven directly by the modulator, *i.e.,* without step-up transformers. The modulator design is described below. To obtain 300 kV, 15 induction cells will be required. Since a standard width for the METGLAS is two inches, the TBNLC induction cell design was based on this width. The METGLAS alloy used for the design is 2714 AS. An example of a core test for the 2714 AS is shown in Figure A-15. The minimum amount of core material is realized when the flux swing approaches saturation. Minimizing the amount of magnetic material, however, does not lead to the most cost-effective design. A flux swing approaching saturation requires a very nonlinear drive current. This leads to a PFN which is difficult to design to achieve a flat voltage pulse. Furthermore, since the losses per unit volume are nearly proportional to ΔB^2, a better design is achieved by allowing a flux swing of less than one half of the saturation value. In this case, a ΔB of 1.0 T results in a dB/dt of 3.3 T/ μs and losses per unit volume of about 150 J/m^3. The inner radius of the induction cell is dictated by beam transport physics and is 4.5 cm for our TBNLC design.

Equation A.6 can be used to determine the required outer radius of the core. With $\Delta B = 1.0$ T, $\Delta t = 300$ ns, $G = 300$ kV/ m, $R_i = 4.5$ cm, core occupancy $\eta_c = 0.75$, and a packing factor $F_p = 0.75$, we find that $R_o =$ 20 cm and the volume of METGLAS is 4.5 10^{-3} m^3. The losses in joules are $E_L = (150\,\mathrm{J/m^3})$ (4.5 10^{-3} m^3) or about 0.68 J per core. The drive voltage for the induction core is 20 kV and the average current dissipated can be estimated from the losses, $E_L = VI_C\Delta t$, or from Eq. A.7. For a 300-ns pulse, the average magnetizing current or drive current of the core I_C is 117 A. The magnetizing current actually has a nonlinear component, as seen in Figure A-15, which reflects the characteristics of the B-H loop. The total drive current (I_T) from the PFN is the sum of the beam current (I_B), the core current (I_C), and the compensation network current (I_N), which for our design is approximately $I_T = I_B + I_C + I_N = (600 + 117 + 33)\,\mathrm{A} = 750$, yielding a core efficiency of 600/750 = 80%.

Line Modulator

The modest repetition rate (120 Hz) and current rise time (100 ns) envisioned for the NLC permits the use of a simple and cost-effective thyratron-driven modulator. For the TBNLC, each induction cell is comprised

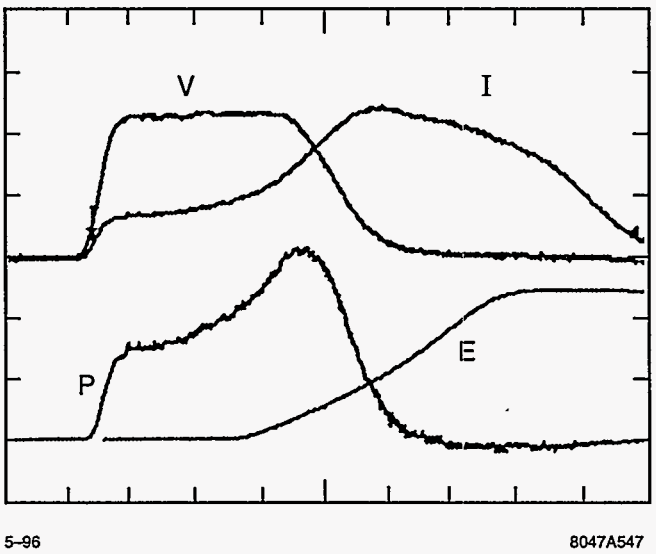


Figure A-15. *Oscilloscope traces for pulsed METGLAS 2714 AS core with cross section of 12.7 cm^2, volume of 450 cm^3, and packing factor = 0.7. Time scale = 100 ns/div. V is applied voltage (1 kV/div), I is current dissipated in core (50 A/div), P is power (VI, 50 kW/div), and E is energy (time integrated power, 20 mJ/div, 250 ns offset). Total energy dissipated = 47 mJ, or 150 J/m^3.*

of five cores individually driven at 20 kV. Driving at this voltage level avoids any step-up transformers and can be generated directly by a thyratron with a 40-kV charging voltage on the PFN. After a preliminary search, the English Electric Valve (EEV) CX1525A appears to be an excellent thyratron for our application. It is a two-gap, deuterium-filled thyratron with a voltage hold-off of 50 kV and capable of delivering 15 kA at 120-Hz repetition rate. Such a device would allow us to drive 15 cells for a total of 300 kV. A simplified schematic of the line modulator is shown in Figure A-16.

As shown in Figure A-15, the current drive to the cores is nonlinear, but a constant amplitude pulse can be generated, within bounds, simply by tapering the impedance of the PFN stages. The PFN will consist of many coupled L-C stages with impedances adjusted to temporally match the induction core impedance. The PFN charging current flows through the induction core, resetting the core prior to the next acceleration cycle. The combination of 2714AS METGLAS cores and the thyratron-driven modulators will result in a conversion efficiency of wall-plug power to cell power of 62%.

Command Resonant Charging and Core Reset

During the energy delivery cycle to the beam, the magnetic core requires a large magnetic intensity to swing 1.0 T in 300 ns. After the pulse is over, the magnetic intensity is reduced to zero, but the flux density has a remnant field, B_r. In order for the core to be useful on the next forward or energy delivery cycle, a reversed magnetic intensity or reset must be applied to the core. The reverse magnetic intensity, however, is much smaller than the forward one, since the rate of demagnetization can be much lower or $\Delta I_{\text{reset}} \ll \Delta I_{\text{forward}}$. By applying the reset current, the flux density is returned to $-B_r$, and the core is ready for the next forward or acceleration pulse. The PFN must be recharged after each forward cycle. The charging process occurs when the command resonant charging (CRC) switch is closed and delivers a half-sine current pulse from the large energy storage capacitor, C_E, to the PFN capacitors. It can be seen from Figure A-16 that the

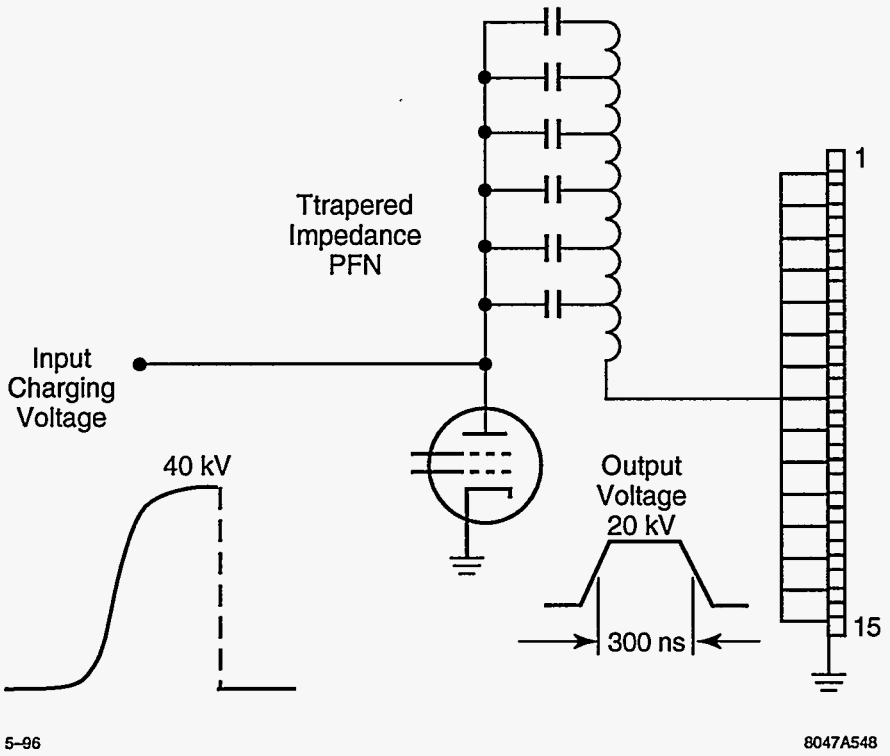


Figure A-16. *Simplified schematic of the line modulator. The PFN will have the same temporal impedance as the nonlinear magnetic core of the induction cells.*

charging current for the PFN flows through the induction cells in the proper direction to reset them. By arranging the impedance and charge time, the proper reset current can be delivered to the cores at the same time that the PFN is being recharged. The reverse magnetic intensity required for reset is about 25 A/m. This is equivalent to about 24 A per core or 360 A for a complete induction module. Since one CRC charges six line modulators, a total of 2.2 kA are required. At the charging voltage of 40 kV, the reset/recharge impedance is 40 kV/2.2 kA= 18.2 Ω. The total capacitance (C_T) of six line modulators is 0.66 μF. The resonant impedance, $Z_r = \pi\sqrt{L_T/C_T}$, requires a charging inductor $L_T = 225\ \mu H$. Hence, the reset/recharge period $T_R = \pi\sqrt{L_T C_T} = 38\ \mu s$.

Efficiency of Induction Accelerator Components

The power conditioning system for the low-energy accelerator has energy losses associated with each major component from the utilities feeding the DC power supply to the induction cell coupling to the electron beam. These losses will be described in reverse order beginning with the induction core to beam coupling as shown on Figure A-17.

The acceleration voltage pulse is shown in Figure A-16. It has a rise time and a falltime of 100 ns with 200-ns flattop. There is no useful energy generated during the falltime, hence, this is lost. The energy during the rise and fall of the pulse is:

$$E_r = \int_o^{t_1} V \cdot I \ dt \quad , \tag{A.10}$$

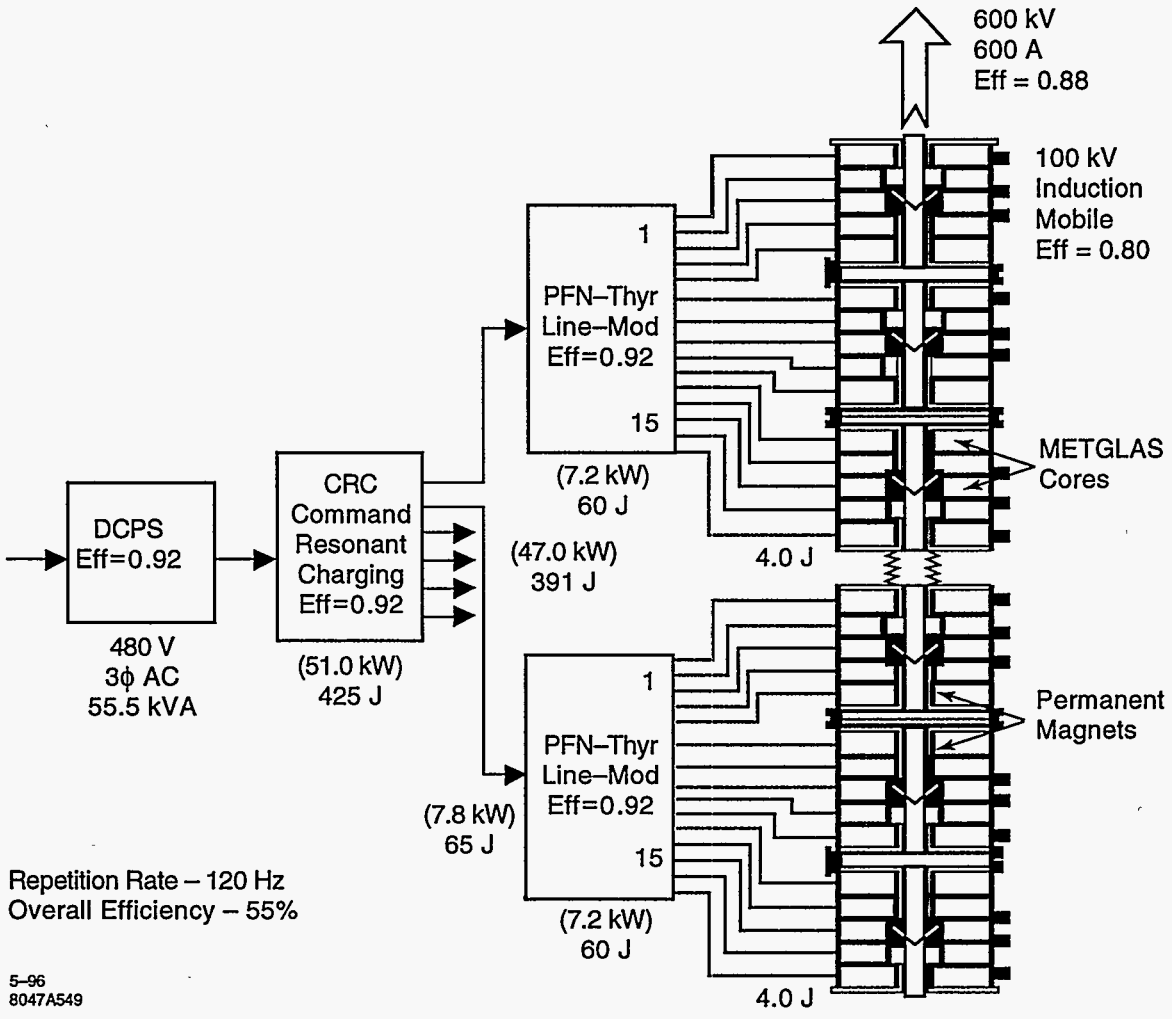


Figure A-17. *Schematic of the power conditioning system for a two-m RK module section.*

where $V = V_o(10^7 t)$ and $I = I_o(10^7 t)$. By taking the integral, we find that $E_r = 10^{14} V_o I_o (t^3/3) = 0.4$ J, while the total energy during the pulse is $E_t = 3.2$ J . The induction core to beam efficiency is then $\varepsilon_b = 2.8/3.2 = 87.5\%$.

The magnetic materials losses have been discussed above. The magnetization current for each induction core was estimated to be 117 A. Another 33 A of current was added for compensation network at the core for a total current drive of 750 A. Hence, the efficiency of the induction accelerator cell is $\varepsilon_c = 600/750 = 80\%$.

As previously discussed, the modulator consists of an impedance-tapered PFN matched to the drive required by the magnetic material, METGLAS 2714 AS. The efficiency is calculated by including the thyratron dissipation, filament, grid bias, keep-alive, and the pulse-shaping resistors losses. The modulator losses add up to nearly one kilowatt. At 120-Hz operation, the total power input has been calculated at 10.6 kW which yields an efficiency $\varepsilon_m = 92\%$.

The line modulator is charged by a solid-state CRC system. The CRC generates a $1-\cos\omega t$-voltage waveform which charges the PFN. The current in the CRC is a half-sinusoid which flows through the induction cells in the resetting direction while the PFN is being charged. The losses for the CRC include the inductor mode-damping resistors, the solid-state switch losses, the trigger generator, and the de$Q$ing or regulation system if the DC power supply has insufficient regulation. The overall efficiency of the CRC is calculated to be $\epsilon_{ch} = 92\%$.

The 20-kV DC power supply will be a conventional 60-Hz, 3-phase, full-wave rectifier with filter. It is projected that the power factor for the power supply will be 0.9 for an efficiency $\epsilon_{\rm ps} = 92\%$.

The overall efficiency of the induction accelerator from the utilities to the electron beam which drives the RK is the product of the individual component efficiencies or $\epsilon_o = 55\%$.

A.4.3 Mechanical Systems

Module Design and Fabrication

The TBA consists of essentially 50 identical RKs stacked end-to-end to form the 15-km overall length. Each of the 300-m-long RKs contain 150 identical extraction and reacceleration sections. These 2-m-long sections are referred to as RK modules. A schematic of a 2-m RK module is shown in Figure A-3. An RK module consists of six identical induction modules and one extraction cavity. Each 300-m RK contains 900 modules and the complete TBNLC will contain 45,000 identical modules.

In the design of the module a large effort was made to keep the geometry of individual parts simple and amenable to mass production. This basically requires designing so that a minimum amount of material has to be removed during part fabrication. Drilling, tapping, and machining of small slots are kept to a minimum as these operations are relatively slow operations. All metal parts are made of 304 stainless steel alloy. Advances in modern stainless steel alloys enable significantly greater cutting speeds. Metal stamping and molding of plastic parts are used to further reduce part costs.

The module assembly relies on brazing and welding. The beam pipe assembly uses six brazed joints. Use of simple fixtures to align the stacked parts and brazing in large batches keeps the time and cost per assembly very low as compared to other methods of assemblies. The final housing assembly is done with an inner and outer weld at each end plate. Automatic welding machines make this a quick and reliable operation.

Core Winding

An individual core winding is made by winding approximately 4700 turns of METGLAS ribbon onto a winding mandrel. The ribbon is 20-μm thick and 51-mm wide. A thin layer of insulation will be dipped or sprayed onto the METGLAS during the winding process to provide adequate resistance to eddy currents between turns. The average voltage between turns is 5. This insulation will replace the thin mylar ribbon presently used to provide core interturn insulation. The mylar ribbon overhangs the core edges and presents a serious problem to edge-cooling the cores. A technique being developed at LBNL uses a thin layer of epoxy, dipped or sprayed on, and then dusted with alumina grit to form a tough interlayer insulation. The epoxy additionally binds the METGLAS layers together to form a rigid self-supporting core.

Core Cooling

Heat generation within the cores will be about 300 W per core. Cooling of the core windings is accomplished by circulating oil through the module housing. Oil flows in at one end of the module housing and flows out at the other end on the opposite side. This causes the oil to flow transversely in the gaps between the cores. Spacers around the outside diameter of the cores prevent oil from flowing around the outside of the cores. Only moderate oil flow rates are required and as a result there is a low-pressure drop between supply and return manifolds. For a flow rate of 27°C oil at 1.25 gpm through the module, the maximum temperature for all the cores is about 60°C on the down-stream side. Quadrupling of the oil flow rate will drop the core temperature by about 13°C. Core temperatures in this range are acceptable; magnetic properties are not affected, thermal stresses are minimal since the core remains at a nearly uniform temperature, and temperature limits on the plastic materials within the module are not approached.

Development of an interlayer insulation to replace the presently used mylar tape is critical for the cooling of the cores. Presently used mylar tape overhangs the edges and largely blocks the transfer of heat from the METGLAS to the oil. As described in Section A.4.3, "Core Winding", alternative interlayer insulations are being developed.

Vacuum System

The accelerator will be pumped in two places in each 2-m module. Pumping will occur at the start of each module on each side of the extraction cavity. There is a 2-cm gap between modules at these points which permits good pumping conductance to a standard mounting flange. The two ports are manifolded together to a single ion vacuum pump. The total gas load per cell block is about 8×10^{-7} Torr-ℓ/s. For this geometry, the pressure between the pump and the rf output structure entrance is calculated for several pumping speeds and results are shown in Figure A-18. In this figure, Induction Gap #5 refers to the most distant induction cell from a pumping manifold. If a 20-ℓ/s pump is chosen, the pressure at the rf output structure entrance will be about 5×10^{-7} Torr.

An inline gate valve between every fourth 2-m module will enable isolation of small sections of the accelerator. In each of these four 2-m modules a right-angle gate valve will be used to enable initial pump down with a portable turbomolecular pumping station.

Moderate pressure levels within the linac will enable some of the more costly cleaning methods used on ultrahigh vacuum systems to be avoided. This will be an unbaked system with a base pressure in the 10^{-7}-Torr range. Systems of this type are generally cleaned by chemical or glow discharge cleaning techniques and perhaps a bake at low temperatures.

Assembly and Alignment

Alignment requirements for the TBA can be achieved using standard fabrication, assembly, and alignment techniques. The centerlines of the 2-m modules will be aligned to within 0.1 mm of their nearest neighbor's centerlines. A gradual accumulation of position errors between successive 2-m modules will occur, and will result in the accelerator not being straight. This is acceptable as long as these bends are not extreme and occur gradually over many half-lattice periods.

Six induction cells will be supported and aligned accurately with respect to one another on one strongback support beam. As shown in Figure A-4, each cell has two quadrupole assemblies. The permanent magnets will be sorted by field strength to optimize paring in the quadrupole assemblies. The magnets will then be

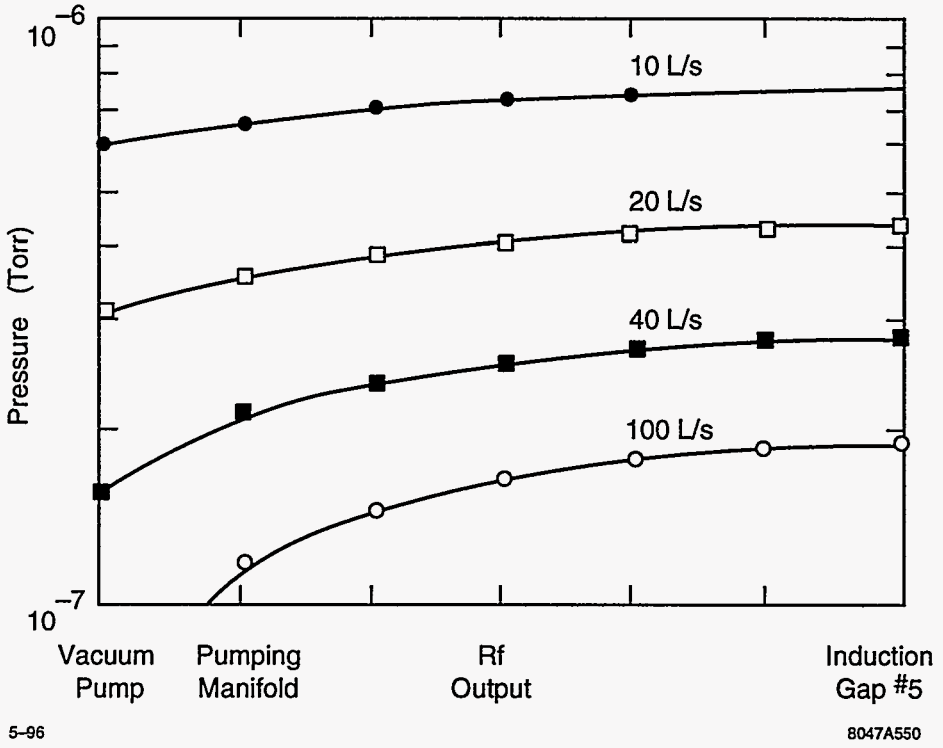


Figure A-18. *Gas pressure distribution within the 2-m module.*

bonded into the assemblies that are held in place by the cell end plates. Quadrupole correction coils will be used to minimize the field errors, instead of attempting to adjust the magnetic center of individual quadrupole assemblies. These procedures will require the permanent magnets to be positioned with a tolerance of 0.1 mm within the 2-m modules. The required tolerance for each step of the assembly will be 0.05 mm. This will ensure that the quadrupole assemblies are well centered with respect to the outside diameter of the cell housing end plates.

The support mounts for all the induction cells on each strongback beam will be positioned and aligned accurately during fabrication of the beam. Mounts for the modules will be welded, machined, and then ground to the required precision. As a result, within each 2-m module, the 12 magnetic quadrupoles will be aligned to the required precision when the modules are assembled onto the beam. There will be no provision made for adjusting the position of magnets within a 2-m module with respect to each other on the beam.

There will be an X and Y set of fiducials at each end of the beam. After the installation of the 2-m modules onto the beam line, a precise measuring fixture will be used to adjust the position of each fiducial accurately with respect to the center of the last quadrupole at each end of the beam. It is desirable to make the fiducials reflect the position of the end quadrupoles rather than some average position for each 2-m section. This will enable the last quadrupole on one beam to be aligned well to the first quadrupole on the next beam.

As successive 2-m modules are installed, they will be aligned using the fiducials to a straight-line reference system. Each module is supported and articulated using a six-strut system. This enables the section to be moved easily and with precision in all six degrees of motion. The straight-line reference system will be able to accurately detect the position of each fiducial. The 2-m module is then articulated to bring the fiducials into a straight line.

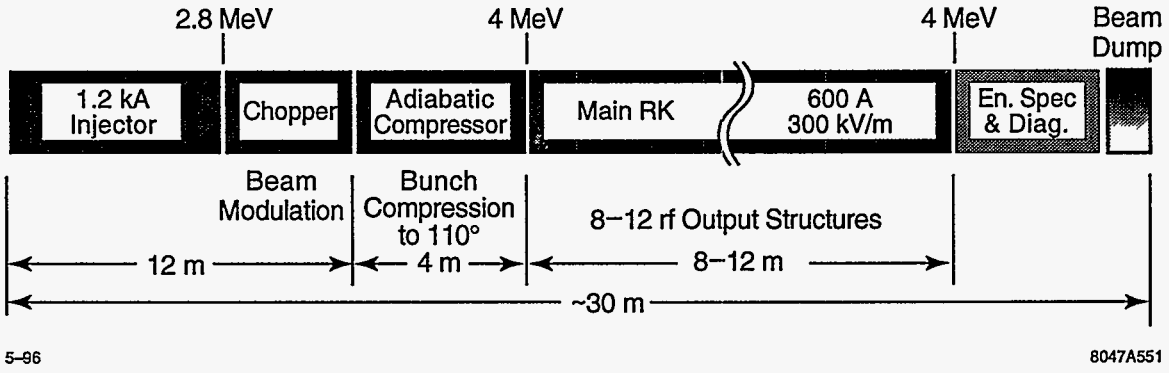


Figure A-19. *Schematic of the RTA showing major components.*

A stretched wire can be used with optical sensors to detect the position of the wire. A laser-based system can be used although a pipe under vacuum is necessary for the light. A series of overlapping straight reference lines will be needed to form the entire 300-m length.

A.5 RTA Test Facility

We are constructing a prototype RK-TBA rf power source to verify the analysis of the TBNLC system study. This prototype, called the RTA, will be located at LBNL. All major components of the TBNLC rf power source will be tested. However, due to fiscal constraints, the prototype will have only 8 rf output structures, with a possible upgrade to 12, instead of the 150 envisioned for the TBNLC. A schematic of the RTA is shown in Figure A-19. Table A-5 lists pertinent parameters for the RTA and TBNLC power source.

The more important issues to be addressed by the RTA are efficiency, longitudinal dynamics, beam stability, emittance preservation, and rf power quality. Efficiency can be separated into the conversion efficiency of wall plug power into beam power and beam power into rf power. The conversion of wall plug power into beam power is described in Section A.5.1 and can be fully measured in the RTA.

High conversion efficiency of beam-to-rf power can be obtained in a system with a large number of output structures. For the TBNLC rf power source, the number of output structures will be limited by beam stability and transport issues. The direct study of beam dynamics issues involving the beam transport through many tens of output structures will not be possible with the prototype. However, the reduced beam energy in the extraction section of the prototype permits the observation of almost an entire synchrotron period. This will be sufficient to allow the beam to approach a steady state condition that can then be extrapolated to a full-scale system with high confidence. The verification of computer simulations used to model the beam dynamics in the TBNLC system study will be a high priority. Beam dynamics issues related to transverse modulation, misalignment of magnetic focusing systems, and adiabatic compression, *e.g.*, emittance growth and corkscrew motion, can be adequately studied.

Parameter	RTA	TBNLC
Pulse		
Duration	200 ns	300 ns
Rise Time	50 ns	100 ns
Current:		
Pre-modulation	1,200 A	1,200 A
Extraction section	600 A DC	600 A DC
	1,100 A rf	1,150 A rf
Beam energy:		
Injector	1 MeV	1 MeV
Modulator	2.8 MeV	2.5 MeV
Extraction	4.0 MeV	10.0 MeV
Bunch compression	240°-110°	240°-70°
Extraction section		
PPM quadrupoles:		
Betatron period	1 m	2 m
Lattice period	20 cm	33.3 cm
Phase advance	72°	60°
Occupancy	0.5	0.48
Pole-tip field	870 G	812 G
Beam diameter	8 mm	4 mm
Rf power:		
Frequency	11.4 GHz	11.4 GHz
Power/structure	180 MW	360 MW
Structures	SW & TW	3 cell TW
Output spacing	1 m	2 m

Table A-5. *Comparison between RTA and the TBNLC.*

A.5.1 Induction Cores and Pulsed-Power System

For our prototype, we will use two different pulsed-power systems. The adiabatic compressor and extraction sections will use induction modules and a pulsed-power system very similar to those described in Section A.4.2 and the TBNLC system study. The primary difference is that each induction module is comprised of three individually driven cores. The cores are still driven at 20 kV by a thyratron charging a multistage PFN. This part of the experiment will be used to verify the efficiency, technical aspects, and cost of the induction modules in the TBNLC.

As described below, the injector is comprised of modified versions of existing equipment. Here, the issue is generating the required volt-seconds within the geometrical constraints of these components. The magnetic material will be 2605 SC METGLAS to maximize flux swing, and the cores in each module will be driven as a single unit to maximize the available cross-sectional area. Driving multiple cores will require the use of a step-up transformer to deliver the required 80–100 kV per cell.

A.5.2 Injector: Gun and Accelerator Sections

Two main goals of the injector design are minimizing electrical field stresses in the gun and realizing the lowest possible emittance growth. The gun and accelerator section comprising the injector, will be modifications of the SNOWTRON induction injector and induction cells from ETA II. Modifications include replacement of the original ferrite cores with METGLAS and a new mechanical alignment system. The present electrode package will be used during initial testing. However, a new electrode package and larger dispenser cathode are required to produce the desired 1.2-kA, 1-MeV beam. The solenoidal field configuration must be optimized for the injector to control the beam radius while minimizing emittance growth. The design goal for the experiment is for a radius < 5 mm and $\epsilon_N < 250\pi$-mm-mr at the chopper entrance.

Alignment of the focusing solenoids is critical to avoid corkscrew motion and emittance growth in the injector. A stretched wire alignment scheme [Griffith 1990] will be used to determine the offset of the solenoid's magnetic axis from a reference mechanical axis for each cell. From past experience, the resolution of this alignment scheme to offset errors is approximately ± 0.05 mm. For offset errors of more than 0.5 mm, the solenoids will be repositioned in the cells. Offset errors of less than 0.5 mm will be recorded, but the solenoids will not be repositioned. A 0.5 mm represents the maximum one induction cell can be mechanically offset from an adjacent cell to achieve magnetic alignment. Each cell also contains a steering (sine/cosine) coil to correct for tilt errors. The required tilt correction will also be determined. The entire alignment will be performed in a precision mill with tolerance on the order of tenths of a mil. A fiduciary will be placed on the outer case of the cell to permit alignment of the magnetic axis when the cells are mounted on the strongback. With this procedure, we expect solenoid offset errors of less than ± 0.08 mm and negligible tilt errors.

Experience operating the ETA II accelerator has shown that careful alignment of the solenoids is not sufficient to reduce the amplitude of the corkscrew motion [Allen 1991] to 0.5 mm desired for the RTA injector. Individual adjustments for the induction cells will permit improved solenoid alignment in the RTA. However, we anticipate using a time-independent steering algorithm [Chen 1992] developed for ETA II to control steering coils on the solenoids. This algorithm corrects for the Fourier component at the cyclotron wavelength of the field error, and led to an order of magnitude reduction in the corkscrew amplitude of the ETA II beam.

A.5.3 Chopper: Beam Modulation

A transverse chopping technique will be used to modulate the beam. The modulator section of the Choppertron, a 11.4-GHz rf generator, will be refurbished for this purpose. A schematic of the modulator is shown in Figure A-20. The solenoidal field immersed incoming electron beam is deflected in the horizontal plane by a 5.7-GHz TM_{110} deflection cavity causing the beam to describe semihelical trajectories along the drift space. The beam scans in a vertical plane across an on-axis aperture placed a quarter betatron wavelength after the deflection cavity. Thus, the 5.7-GHz spatially-modulated DC beam incident on the aperture becomes a phase-coherent, amplitude-modulated beam at 11.4 GHz.

The desired bunch length, peak current, and energy for the drive beam in the extraction section of the prototype is respectively 110°, 600 A, and 4 MeV. Such a train of bunches could be generated by directly chopping the unmodulated beam. However, considerations of efficiency (70% of the beam would be lost) and practical feasibility (we expect approximately 1.2 kA of peak current from the injector) require that we do not fully modulate the beam by chopping. Our intent is to chop the beam at an energy of 2.8 MeV into bunches of approximately 240°. An adiabatic compressor section will be used to further bunch and accelerate the beam.

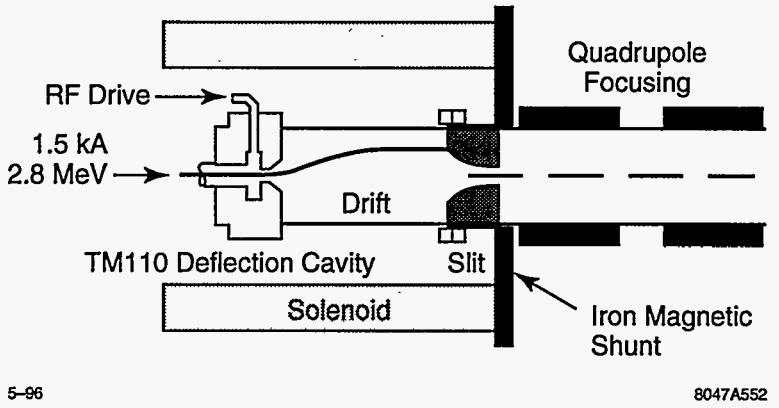


Figure A-20. *Schematic of the beam modulator (chopper).*

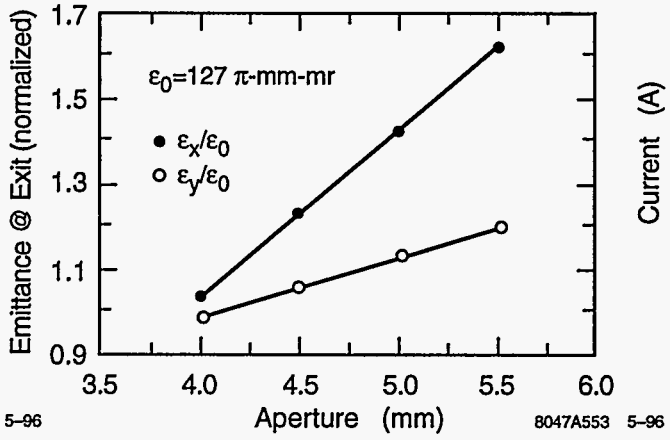


Figure A-21. *Emittance growth as a function of aperture size*

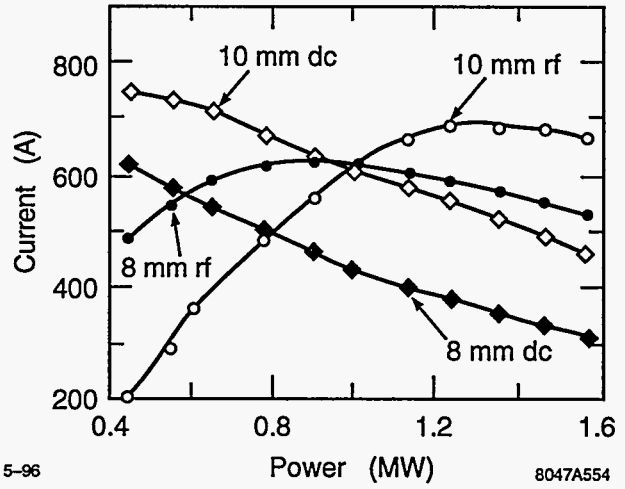


Figure A-22. *The DC and rf (11.4-GHz) current components as a function of drive power to the deflection cavity for two apertures.*

Designs of chopper systems have been extensively analyzed [Haimson 1965, Haimson 1970], and the original Choppertron was optimized for operation with the ETA II induction beam. The intent for the prototype experiment is to limit modifications of the modulator section of the Choppertron to adjustments in the drift length and the chopping slit aperture size. It is desirable for the radius of the beam to remain relatively constant in the chopper. For a given beam energy, current, and emittance, the radius is determined by the solenoidal field. Once the solenoidal field is determined for the desired beam radius, the drift section length is also fixed. The maximum deflection amplitude at the aperture is determined by the transverse momentum imparted to the beam by the deflection cavity. The deflection amplitude and chopping slit aperture determine the emittance growth and the modulated beam's characteristics.

A series of simulations were performed with the relativistic klystron design code RKS to determine the optimum deflection and aperture. Results are shown in Figures A-21 and A-22. The beam was assumed to be cylindrical with a maximum radius of 3.6 mm, waterbag distribution, and normalized, rms emittance (ε_x

and ε_y) of 127 π-mm-mr at the chopper entrance. In Figure A-21, the emittance has been normalized with respect to the initial value. The emittance represents the area in trace space for all the particles within a 360° rf bunch length. The emittance increases with the chopping slit aperture as the simulations cover the range where a substantial current loss occurs on the aperture. Thus, the beam has been deflected to fill the acceptance of the aperture. The difference between the x and y emittances is due to the deflection in a single plane. In the simulations, x represents the horizontal plane.

Considering only the increase in emittance, it would appear that the smaller aperture is better. However, as shown in Figure A-22, the amount of current contained in a 240° bunch length is substantially reduced as the chopping slit aperture is reduced. We require at least 600 A of the initial DC current to remain in this bunch length for the adiabatic compressor section. To allow for current losses during the adiabatic compression, 660 A is a practical limit. Also, the beam must be sufficiently modulated for the idler cavities in the adiabatic compressor to function. The rf current component (modulation) initially increases with drive power, as shown in Figure A-22, with a decrease in DC current. The design goal of the chopper is to generate a 240° rf bucket containing 600 A that can be captured and bunched by the adiabatic compressor, with minimum emittance growth. Maximizing the rf current at the exit of the chopper will not optimize emittance growth with respect to DC current in the rf bucket. A 10-mm aperture at a drive power of about 0.8 MW produces a satisfactory modulation for growth in emittance, although further studies are required for an optimum design. Future simulations will use the results of EGUN simulations for the initial beam characteristics.

A.5.4 Adiabatic Compressor

Adiabatic compression is used to achieve the desired beam characteristics at the entrance of the extraction section. In the adiabatic compressor, the bunch length is reduced from 240° to 110° with SW idler cavities while the beam is accelerated to an energy of 4 MeV. The accelerating gradient of the induction cells (300 kV/m) sets the minimum length of the adiabatic compressor to 4 m. Extensive 1-D numerical studies have been performed to determine the most efficient scheme for accomplishing this: The present design uses seven idler cavities appropriately spaced and detuned to progressively bunch the beam. The idler cavity has a resonant frequency higher than the drive frequency. The rf field in the cavity, excited by the beam, places itself at 90° with respect to the rf bucket. This leads to a loss of energy for electrons towards the front of the rf bucket, and a gain for those in the back.

The bunching effect of the idler cavity is illustrated in Figure A-23. As the beam drifts between the idler cavities, electrons in the tail of the rf bucket have a greater velocity leading to increased bunching. At the same time, space charge forces will increase the energy of electrons at the head of the bunch with respect to the tail slowing the bunching process. Appropriately designing and spacing the idler cavities can eventually bunch the bucket to the desired length. Table A-6 summarizes the important parameters of the adiabatic compressor, and Figure A-24 shows the overall effect on the beam.

A.5.5 RF Power Extraction

After leaving the adiabatic compressor, the beam enters the main RTA where power is extracted. Here the beam energy is periodically converted into rf energy (via output cavities) and then restored to its initial value (via induction modules). Stable propagation of the rf bucket traveling through many resonant cavities and achievable power output extraction have been studied numerically. Space-charge effects and energy

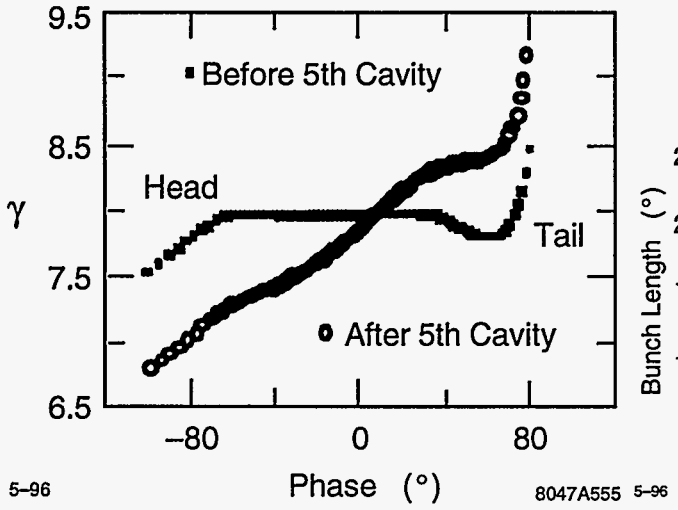


Figure A-23. *Simulations of longitudinal phase space before and after the 5th idler cavity.*

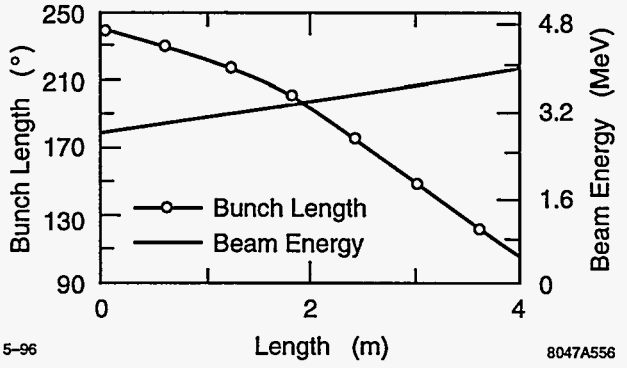


Figure A-24. *Simulations of adiabatic compressor effect on bunch compression and beam energy.*

Cavity	Position (m)	Bunch length (°)	Beam E_k (MeV)	Resonant freq. (gHz)
1	0.0	240	2.8	11.82
2	0.6	229.7	2.98	11.82
3	1.2	217.9	3.17	11.82
4	1.8	201.7	3.36	11.74
5	2.4	177.6	3.53	11.74
6	3.0	148.6	3.69	11.74
7	3.6	122.0	3.85	11.74
exit	4.0	105.5	4.00	

Table A-6. *Parameters of the adiabatic compressor.*

spreads due to rf fields in the output structures tend to debunch the beam as it traverses the main RTA. To counteract this effect, inductively detuned output structures are used.

Both traveling wave (TW) and standing wave (SW) structures are being considered for the output structure design of the RTA. The TBNLC RK design used TW structures to reduce the surface fields associated with generating 360 MW per output. RTA is designed to generate 180 MW per output. Thus, inductively detuned SW cavities are a practical alternative. Furthermore, the RKS code is believed to incorrectly describe the radial beam dynamics for inductively detuned TW structures. [Giordano 1995] We have used SW cavities in our modeling to perform complete 2-D simulations and to validate the inductive detuning concept. The detuning concept for the SW output cavities is similar to that used in the adiabatic compressor cavities. The required detuning is also affected by the finite external Q of the cavities. Different cavities can have different Q values to optimize the output power and the bunching. The detuning mechanism is required for stable rf power production through the main RTA. See the simulation results in Figure A-25 Resonant frequency of the cavities was 11.566 GHz for this simulation.

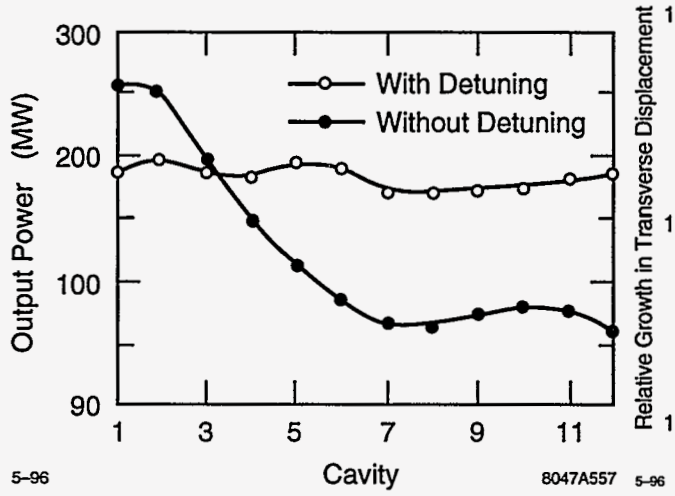


Figure A-25. *Simulations of the rf power generated per output, both with and without detuning of the output cavities.*

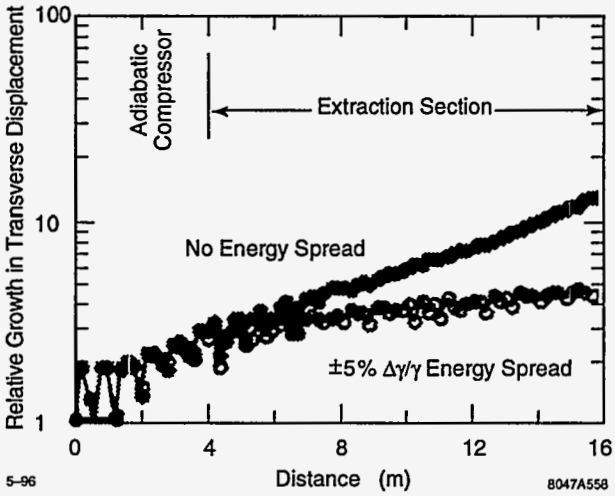


Figure A-26. *Simulations of growth in the transverse displacement of the beam centroid from deflection by HOMs in the induction cell.*

A.5.6 Beam Dynamics Issues

Transverse instability of the beam due to the excitation of higher order modes (HOM) in the rf couplers and induction cell gaps is a serious issue for a long relativistic klystron. We do not anticipate beam loss due to the HOM excitation, but the effect should be measurable. The HOMs associated with three different components will predominate in the prototype. Fortunately, the frequency ranges do not overlap so that the effects can be studied separately.

The first is the approximately 300-MHz transverse mode in the accelerator induction cells. ETA II has operated with 3 kA of current through 60 induction cells of the same geometry, but with ferrite cores. The 20 accelerator cells in the RTA should not pose an instability problem with 1.2 kA of current, but it should be possible to detect weak beam oscillations at 300 MHz by using rf probes.

Of greater concern are the 60–80 prototype induction cells of the adiabatic compressor and extraction section. The transverse impedance of a cell is roughly proportional to the cell gap divided by the square of the beam-pipe radius. The transverse impedance of the smaller prototype cells is about four times that of the larger accelerator cells. Simulations of the growth in the transverse instability were performed using the OMICE code. Results are shown in Figure A-26. For the purpose of the simulation, the prototype cells were assumed to have the same characteristics as the TBNLC induction cells ($Z_{\perp 1} = 5,400\,\Omega/m, Q = 2$, resonant frequency = 3 GHz), the current was increased from 0 to 600 A in 50 ns, total pulse duration was 200 ns, and a step function offset was used as the excitation seed. One curve represents growth for a monoenergetic beam while the second has a ±5% energy spread over the rf bucket, imposed to illustrate the effect of Landau damping. The expected exponential growth is evident in Figure A-26, but well within the design goal of less than a factor of 100.

The third and most critical instability affecting transverse motion is caused by the rf output structures in the extraction section. Despite heavy damping of the structures and the effect of Landau damping, it was necessary to use the Betatron Node Scheme to suppress the transverse instability to a tolerable level for the

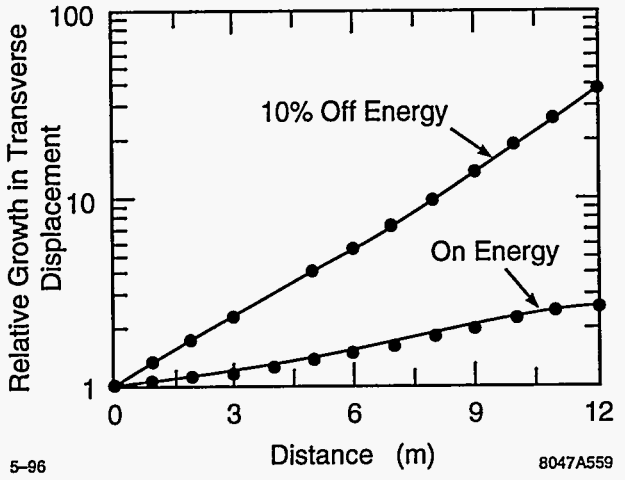


Figure A-27. *Simulations showing growth in the transverse displacement of the beam centroid from deflections by HOMs in the rf structures.*

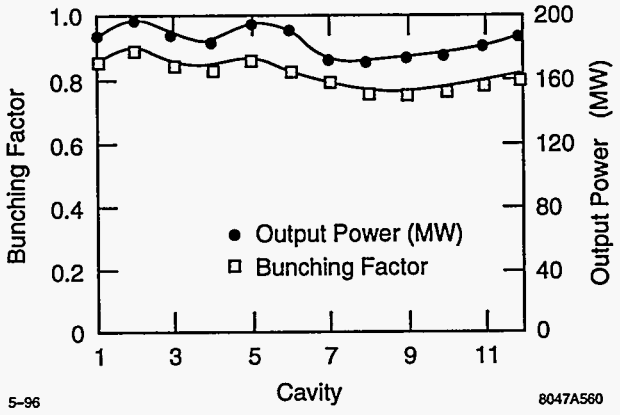


Figure A-28. *Evolution of bunching factor and power output in the main RTA with detuned rf output structures.*

TBNLC. Simulation results of the growth in the transverse instability through the extraction section are shown in Figure A-27. The rf characteristics of the three-cell traveling-wave output structures described in the TBNLC design study were used for the simulation ($Z_\perp/Q = 3\Omega$ per cell, dipole resonant frequency of 14 GHz, $Q = 10$ for the first and third cells and 3,500 for the second, and dipole phase advance of about $2\pi/3$). A monoenergetic beam with parameters the same as for the induction cell simulation was used for this simulation. A relative small change in the beam energy from that required for the Betatron Node Scheme can lead to substantial increase in the growth of the instability as indicated in Figure A-27. While the growth remains tolerable for both cases shown, the idler cavities in the adiabatic compressor require different transverse rf characteristics than the main RTA output cavities to avoid beam loss and/or adverse emittance growth. The difference in HOM power generated in the cavities is three orders of magnitude greater after 12 cavities when not operating at the energy for the Betatron Node Scheme. A measurement of the HOM component in the output power is expected to be sensitive to the effectiveness of the scheme.

Landau damping and the Betatron Node Scheme were both required for transverse stability in the TBNLC design. Neither are required for the operation of the RTA, but the effectiveness of both for a variety of operating parameters can be measured. The effect of the Betatron Node Scheme can be readily ascertained by measuring the HOM power generated in the output structure while rf loops distributed along the beam line can determine the effectiveness of Landau damping.

Longitudinal stability issues include both the rf bucket and the phase relationship between rf buckets. The rf bucket must remain appropriately bunched for stable rf current and power extraction as described in Section A.3.1. Simulation results of the rf power generated and bunching in the extraction section when the resonant frequency of all the detuned output structures has been set to 11.566 GHz is shown in Figure A-28. The bunching factor is defined as:

$$b = \left| Re\left\{ \frac{1}{N} \sum_{n=1}^{N} e_n^{j\psi} \right\} \right| \quad , \tag{A.11}$$

where the sum is over all the electrons and ψ is the phase of the electron with respect to the center of the rf bucket. The relatively constant value of the bunching factor is a good indication of long term stability of the rf bucket.

The length of RTA is adequate for a meaningful measurement of the longitudinal beam dynamics involved in the detuning of the output structures to maintain the rf bucket. The synchrotron wavelength can be expressed as:

$$\lambda_s = 2\pi \left(\frac{\omega}{c\gamma^3} \frac{d\gamma}{dz} \right)^{-\frac{1}{2}} \quad . \tag{A.12}$$

At 4 MeV, accelerating gradient of 300 kV/m, and 11.4 GHz, λ_s is $\approx$14 m. Numerical sensitivity studies indicate that rf output power is insensitive to energy variation and shows small variations ($< 4\%$) for current variations of $\pm 1\%$. The variations were imposed for the flattop portion of the beam pulse. Phase stability is not appreciably effected by current variations of less than $\pm 1\%$. However, phase sensitivity leads to a severe requirement on the average (head-to-tail) energy variation for the flat-top portion of the beam pulse. Phase variations are well modeled by the following first-order formula:

$$\frac{\Delta\psi}{\Delta z} = \frac{k}{\overline{\gamma}^3} \Delta\gamma \quad , \tag{A.13}$$

where $\Delta\psi$ is phase variation, $\Delta\gamma$ is head-to-tail energy variation over the pulse length, k is the free space wave number, and $\overline{\gamma}$ is the average beam energy. With the RTA parameters, assuming that field phase variation should not exceed $\pm 5°$, the required pulse energy flatness (flattop) is estimated to be $\pm 0.3\%$ for an 8-m extraction section.

Beam emittance is an important parameter for the RTA. After the chopper, the focusing system is comprised of permanent quadrupole magnets. The ppm quadrupole focusing is important in the TBNLC design for cost and efficiency reasons. For the RTA, mechanical design constraints and the experimental goal of studying the Betatron Node Scheme require a pole field at a radius of 2 cm for the quadrupoles of about 870 gauss, a half-lattice period of 10 cm, occupancy factor of 0.5, and a phase advance of 72°. The normalized edge emittance must be no larger than 800 π-mm-mr to meet the design goal of an average beam radius (edge) in the extraction section of 4 mm.

Our goal is to limit the emittance growth in the injector from beam optics to a factor of three times the thermal source emittance of about 80 π-mm-mr (0.1 eV) for the cathode. The chopper is expected to increase the emittance by a factor of about 1.7. Thus it is very important to minimize sources of emittance growth such as nonzero magnetic flux at the cathode, magnet misalignments, solenoid to quadrupole matching, and higher order multipoles in the quadrupole magnets. The strong focusing used in the extraction section and the large beam energy spread will convert any transverse motion of the beam, *e.g.*, corkscrew motion or transverse instabilities, into an increase in effective emittance.

A.6 Conclusions

A preliminary design of an RK-TBA based rf power source for the NLC has been presented. The TBNLC system study focused on three major areas: (1) RK-TBA physics, (2) RK-TBA engineering and (3) RK-TBA costing. Here we have described the more important results related to physics and engineering from that study. In addition, we have described an experimental program to construct and test a prototype rf power source based on this design. The description of this program emphasized a number of the RK-TBA engineering issues studied in the TBNLC system study.

The issue of an economical and efficient rf power source is central to the development of future linear colliders. While the induction linac based TBA is known to have the potential for very high efficiency, research support received to date has been somewhat limited. General concerns about the TBA concept as a realizable power source include:

- The induction linac technology is not as mature as existing rf-klystron technology.
- The TBA beam dynamics, particularly BBU, is difficult.
- Induction linacs are perceived to be very expensive.

The objectives of the RK-TBA research program at LBNL and LNLL were motivated by these concerns.

In regards to the cost issue, past cost estimates were based on induction linacs built for other applications (short pulse, high-repetition rate, high current). Our belief has been that the induction linac design ought to be optimized for the specific mission. Hence, the TBNLC system study was specifically directed towards an rf power source suitable for the NLC. Likewise, the experimental program is developing a prototype rf power source that is suitable for powering the NLCTA.

Since cost is, and will be, a key issue for future colliders, we feel that it is essential to have a believable cost estimate. The only way to have a believable cost is to carry the engineering study to sufficient details to make a bottom-up estimate possible. We have tried to estimate cost for fabrication and assembly on the basis of engineering drawings and proposed mass-production procedures. For the induction modules, a major component of the machine, we have sent out for external bids for various parts, to cross-check our own estimates. Our estimate, for approximately ten different items, agreed with external quotes to a few percent on the total. All components are based on known technology, with relatively well-known costs. Hence for the engineering design proposed, we believe that our cost estimate is realistic.

Similarly, the efficiency estimate for our pulse power design is based on well-known numbers with today's technology. We have not made any extrapolations for possible future innovations. We might note here that as we have done more testing in the experimental program, the 40% number quoted in the TBNLC system study for the efficiency of the pulse power system has been increased to 55% leading to a wall-plug-to-rf efficiency of 50%. As we learn more about the technology, we may introduce new design changes which could alter the cost and efficiency estimates. Yet, on the basis of the work that has been performed thus far, we can safely argue that the cost of upgrading the initial NLC configuration to one or more-TeV-c.m. energy is an attractive option.

In regard to the beam dynamics issues, we have offered conceptual solutions for longitudinal beam stability as well as BBU control for both the low-frequency component, associated with induction gaps, and the high-frequency component, associated with rf extraction cavities. These concepts are supported by detailed simulations. We hope to have demonstrated by these studies that the solutions offered have a reasonable chance of success.

We would like to stress that the reported work is an ongoing effort. Further theoretical and computation studies are being conducted. Specifically, cavity design work continues, in particular in relation to the issue of how to minimize the transverse defocusing on the beam envelope, and the induction cavity design is being optimized to achieve the needed longitudinal as well as transverse impedances. Feedback systems are being studied to relax the constraints on energy flatness and quadrupole field errors. Our experimental program allows for hardware development, verification of theory/simulations, and improvement of our cost estimates. At the scheduled completion of our experiments in 2002, the TBNLC could be seriously considered for a 1.5-TeV NLC upgrade.

References

[Allen 1989] M.A. Allen *et al.*, "High Gradient Electron Accelerator Powered by a Relativistic Klystron", *Phys. Rev. Lett.* **63**, 2472 (1989).

[Allen 1991] S.L. Allen *et al.*, "Measurements of Reduced Corkscrew Motion on the ETA-II Linear Induction Accelerator", *Proc. of the 1991 Part. Acc. Conf.*, 3094–3096 (1991).

[Barnard 1991] J. Barnard *et al.*, "Study of Recirculating Induction Accelerators as Drivers for Heavy Ion Fusion", UCRL–LR–108095 (1991).

[Brigg 1985] R.J. Brigg *et al.*, "Theoretical and Experimental Investigation of the Interaction Impedances and Q Values of the Accelerating Cells in the Advanced Test Accelerator", *Particle Acclerators* **18**, 41 (1985).

[Calame 1991] J.P. Calame and W.G.Lawson, "A Modified Method for Producing Carbon-Loaded Vacuum-Compatible Microwave Absorbers from a Porous Ceramic", *IEEE Trans. Electron. Devices* **38**–6, 1538–1543 (June 1991).

[Chen 1990] Y.-J. Chen, "Corkscrew Modes in Linear Accelerators", *Nucl. Instr. and Methods* **A292**, 455–464 (1990).

[Chen 1992] Y.-J. Chen, "Beam Control in the ETA-II Linear Induction Accelerator", *Proc. of the 1992 LINAC Conf.*, 540–544 (1992).

[Clark 1988] J. Clark *et al.*, "Design and Initial Operation of the ETA-II Induction Accelerator", *Proc. 14th Int'l Linac Conf.*, Williamsburg, VA, 19–23 (1988).

[Chattopadhyay 1990] S. Chattopadhyay, ed. "Impedance Beyond Cutoff", Special edition of *Particle Acclerators* **25**, Nos. 2–4 (1990).

[Deford 1989] J.F. Deford, G.D. Craig, and R.R. McLead, "The AMOS (Azimuthal Mode Simulator) Code", *Proc. of the 1989 Part. Acc. Conf.*, Chicago, IL (1989).

[Deford 1990a] J.F. Deford and G. Kamin, "Application of Linear Magnetic Loss of Ferrite to Induction Cavity Simulation", *Proc. 1990 Linear Accel. Conf.*, Albuquerque, NM, 384–386 (1990).

[Deford 1990b] J.F. Deford *et al.*, "The AMOS Wakefield Code", *Proc. Conf. on Computer Codes and the Linear Acc. Community*, Los Alamos, NM, 265–289 (1990).

[Fessenden 1992] T. Fessenden and C. Fong eds. "Induction Linac Systems Experiments Conceptual Design Report", LBL–PUB–5324 (1992).

[Giordano 1995] G. Giordano *et al.*, "Beam Dynamic Issues in an Extended Relativistic Klystron", *Proc. of the 1995 Part. Acc. Conf.* (1995).

[Griffith 1990] L.V. Griffith and F.J. Deadrick, "Progress in ETA-II Magnetic Field Alignment Using Stretched Wire and Low Energy Electron Beam Techniques", *Proc. of the 1990 Linac Conf.*, 423–425 (1990).

[Haimson 1965] J. Haimson, "Injector and Waveguide Design Parameters for a High Energy Electron-Positron Linear Accelerator", *IEEE Trans. Nucl. Sci.* **NS-12**–3, 499–507 (June 1965).

[Haimson 1970] J. Haimson, "High Duty Factor Electron LINACS", in *Linear Accelerators*, P.M. Lapostolle and A.L. Septier eds., 462–466. (North-Holland, Amsterdam, 1970).

[Haimson 1989] J. Haimson, and B. Mecklenburg, "Design and Construction of a Chopper Driven 11.4-GHz Traveling Wave RF Generator", *Proc. of the 1989 Part. Accel. Conf.*, 243–245 (1989).

[Haimson 1992] J. Haimson, and B. Mecklenburg, "Suppression of beam induced pulse shortening modes in high power RF generator and TW output structures", *Proc. SPIE Symposium on Intense Microwave and Particle Beams III*, Vol. **1629–71**, 209 (1992).

[Halbach priv] K. Halbach, private communication.

[Halbach 1980] K. Halbach, *Nucl. Instr. and Methods* **169**, 1–10 (1980).

[Hogan 1992] W.J. Hogan, R. Bangerter, and G.L. Kulcinski, "Energy from Inertial Fusion", *Physics Today* **42**–9, 42–50 (1992).

[Houck 1992a] T.L. Houck, and G.A. Westenskow, "Status of the Choppertron Experiments", *Proc. 16th Int'l Linac Conf.*, Ottawa, Ontario, Canada, 498–450 (1992).

[Houck 1992b] T.L. Houck, G.A. Westenskow, and S.S. Yu, "BBU Code Development for High-Power Microwave Generators", *Proc. 16th Int'l Linac Conf.*, Ottawa, Ontario, Canada, 495 (1992).

[Houck 1993] T.L. Houck, "Design Study of a Microwave Driver for a Relativistic Klystron Two-Beam Accelerator", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC, 2590 (1993).

[Kapetanakos 1985] C.A. Kapetanakos, and P. Sprangle, "Ultra-high-current electron induction accelerators", *Physics Today* **38**, 58 (1985).

[Kroll 1990] N.M. Kroll, and D.U.L. Yu, "Computer determination of the external Q and resonant frequency of waveguide loaded structures", *Particle Acclerators* **34**, 231–250 (1990).

[Lee 1978] E.P. Lee, "Resistive Hose Instability of a Beam with the Bennett Profile", *Physics of Fluids* **21**, 1327 (1978).

[Li 1994] H. Li, T.L. Houck, S.S. Yu, and N. Goffeney, "Design Consideration of Relativistic Klystron Two-Beam Accelerator for Suppression of Beam-Breakup", *Proc. SPIE Symposium on Intense Microwave Pulses II*, **2154**–10, 91–98 (1994).

[Orzechowski 1986] T.J. Orzechowski *et al.*, "High-Efficiency Extraction of Microwave Radiation from a Tapered-Wiggler Free-Electron Laser", *Phys. Rev. Lett.* **57**, 2172 (1986).

[Panofsky 1968] W.K.H. Panofsky and M. Bander, "Asymptotic Theory of Beam Break-up in Linear Accelerators", *Rev. Sci. Instr.* **39**, 206 (1968).

[Ryne 1990] R.D. Ryne, and S.S. Yu, "Relativistic Klystron simulations using RKTW2D", *Proc. of the 1990 Linear Acc. Conf.*, Albuquerque, New Mexico, 177–179 (1990).

[Sessler 1982] A.M. Sessler, "The free-electron-laser as a power source for a high gradient accelerating structure", Workshop on Laser Acceleration of Particles, NY, *AIP Conference Proceedings* **91**, 154 (1982).

[Sessler 1987] A.M. Sessler and S.S. Yu, "Relativistic-Klystron Two-Beam-Accelerator", *Phys. Rev. Lett.* **58**–23, 2439–2442 (1987).

[Siemann 1993] R.H. Siemann, "Overview of Linear Collider Designs", *Proc. of the 1993 Part. Acc. Conf.*, Washington, DC, 532–536 (1993).

[Thompson 1993] K.A. Thompson *et al.*, "Design and simulation of accelerating structures for future linear colliders", SLAC–PUB–6032 (Nov. 1993).

[Westenskow 1994] G.A. Westenskow and T.L. Houck, "Results of the Reacceleration Experiment: Experimental Study of the Relativistic-Klystron Two-Beam-Accelerator Concept", *Proc. of the 10th Int'l Conf. on High Power Particle Beams*, San Diego, CA (1994).

Contributors

- Fred Deadrick
- Guido Giordano
- Enrique Henestroza
- Tim Houck
- Hai Li
- Steve Lidia
- Louis Reginato
- David Vanecek
- Glen Westenskow
- Simon Yu

B

A Second Interaction Region For Gamma-Gamma, Gamma-Electron and Electron-Electron Collisions

Contents

B.1 Introduction . . . 973
B.2 Physics Opportunities at $\gamma\gamma$ Collider: The Higgs Sector and Other New Physics . . . 974
B.2.1 The Higgs $\gamma\gamma$ Partial Width . . . 974
B.2.2 Higgs CP Eigenvalue . . . 975
B.2.3 Higgs Boson Search . . . 976
B.2.4 Strongly Interacting Electroweak Sector . . . 976
B.2.5 Supersymmetry . . . 977
B.2.6 Compositeness . . . 977
B.3 Major Parameters . . . 978
B.3.1 Basic Scheme . . . 978
B.3.2 Laser Parameters . . . 978
B.3.3 Electron Beam Parameters . . . 979
B.4 CP Issues . . . 980
B.4.1 Optimization of the Laser Parameters . . . 980
B.4.2 Low-Energy Electrons Due to High-Order Multiple Scattering . . . 983
B.4.3 Compton Conversion Efficiency . . . 984
B.5 IP Issues . . . 987
B.5.1 Optimization of Spectral Luminosity . . . 987
B.5.2 Polarization . . . 987
B.5.3 Collision of the Spent Electron Beam . . . 987
B.5.4 Disruption of Low Energy Electrons . . . 988
B.6 Luminosity Calculations . . . 989
B.6.1 Simulation Code Development . . . 989
B.6.2 Simulation of the CP . . . 990
B.6.3 Telnov's Simulation Results for $\gamma\gamma$, γe^- and e^-e^- Luminosities . . . 990
B.7 Backgrounds and Other Detector Considerations . . . 999
B.7.1 Introduction . . . 999
B.7.2 Physics Requirements . . . 999
B.7.3 Backgrounds . . . 1000
B.7.4 Detector Considerations . . . 1000
B.7.5 Initial Simulations . . . 1002
B.7.6 Conclusions . . . 1003

B.8 Laser Optical Path in IR 1006
B.8.1 Single-Pass Scheme 1006
B.8.2 Optical Beam Focusing 1009
B.8.3 Laser Damage of Optics 1010
B.8.4 Ideas on Relaxing Average Laser Power Requirement Via Multipass Optics 1013
B.9 Gamma-Gamma Final Focus System 1015
B.9.1 Beam Parameters 1015
B.9.2 Final Focus Doublet 1016
B.9.3 Chromaticity Compensation 1017
B.10 Extraction and Diagnostic Line 1019
B.11 Laser Technology I: Solid State Lasers 1020
B.11.1 Laser Materials 1020
B.11.2 Chirped-Pulse Amplification 1021
B.11.3 High-intensity, Short-pulse Laser Systems 1023
B.11.4 Synchronization and Repetition Rate 1024
B.11.5 1-ps, 1-J Laser System for Nonlinear QED Experiments 1024
B.11.6 Average Power 1024
B.11.7 NLC Laser Concept 1025
B.11.8 A Ring Configuration for Multiplexing and Polarization Control 1026
B.11.9 Polarization Control at the Interaction Point 1028
B.11.10 Conclusion 1029
B.12 Free-Electron Lasers 1030
B.12.1 An FEL Scheme Using Induction Linac and Chirped Pulse Amplification Technique 1030
B.12.2 Chirping Requirement and Tolerance 1032
B.12.3 Induction Linac Driver 1032

B.1 Introduction

To maximize the accessible high-energy physics, the NLC will have two interaction regions (IRs): one will study e^+e^- collisions and the other may study $\gamma\gamma$, γe^-, and e^-e^- collisions. In this appendix, we describe the final focus and interaction region required for $\gamma\gamma$ and γe^- collisions, henceforth referred to as IR2.

For both $\gamma\gamma$ and γe^- collisions, the required high-energy photons (γ beams) are most effectively produced via Compton backscattering of focused laser beams by the high-energy electron beams. The high-energy photon beams are then brought into collision with opposing electron and photon beams for γe^- collisions and $\gamma\gamma$ collisions, respectively. This region is distinctly different from the e^+e^- final focus and interaction region in that the final focus is optimized to produce rounder beams and that the IR must contain one IP for the luminosity collision and one or two conversion points where the photon beams are generated. With suitable laser and electron beam parameters, a luminosity of γe^- or $\gamma\gamma$ collisions comparable to that of the e^+e^- collisions can be achieved. The polarization of the high-energy photons can be controlled by the polarizations of the laser and the electron beams. With high luminosity and variable polarization, the $\gamma\gamma$ and γe^- collisions at TeV energies will significantly enhance the discovery potential and analytic power of a TeV linear collider complex. A conceptual layout of the NLC including the second IR for $\gamma\gamma$ collisions is shown in Figure B-1.

A review of $\gamma\gamma$ and γe^- colliders can be found in the proceedings of a workshop at Berkeley [Berkeley 1995]. The idea of incorporating γe^- or $\gamma\gamma$ collisions in a future linear collider via Compton backscattering of a laser beam [Arutyunian 1963] has been studied conceptually, especially by scientists from Novosibirsk [Ginzburg 1981, Ginzburg 1983, Ginzburg 1984, Telnov 1990, Telnov 1991, Telnov 1995]. The nonlinear QED experiment E-144 [Heinrich 1991] on the FFTB line at SLAC may be regarded as an essential "proof-of-principle" for future $\gamma\gamma$ or γe^- colliders.

The physics opportunities for γe^- and $\gamma\gamma$ collisions at the NLC are described in Section B.2. Some examples are [Brodksy 1994, Chanowitz 1994, Ginzburg 1994]:

- A $\gamma\gamma$ collider offers a unique opportunity for measuring the two-photon decay width of the Higgs boson, providing a glimpse of the mass scale beyond the TeV range.
- A $\gamma\gamma$ collider is well suited for searching for new charged particles, such as SUSY particles, leptoquarks, excited state of electrons, etc. because photons generally couple more effectively to these particles than do electrons or positrons.
- A $\gamma\gamma$ or γe^- collider serving as a W-factory, producing 10^6–10^7 Ws/year, allowing for a precision study of gauge boson interactions and a search for their possible anomalies.
- A γe^- collider is uniquely suited to studying the photon structure functions, etc..
- e^-e^- collisions (without conversion to γ rays) are interesting by themselves.

In the following, we describe a preliminary design of the IR2 for the NLC, with the goal of obtaining $L_{\gamma\gamma}$ about $10^{33}\,\mathrm{cm}^{-2}s^{-1}$ within a 10% bandwidth or several times $10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for a broad spectrum. In this design, we chose to employ the electron beam parameters for the e^+e^- collision before the final focus system (FFS). However, the FFS for the $\gamma\gamma$ collision is modified so that $\beta_x^* = \beta_y^* < 1\,\mathrm{mm}$. An FFS satisfying the luminosity requirements is worked out with a tolerance requirement similar to that of the e^+e^- FFS. An elaborate optical mirror system in the very constrained region around the vertex detector and quadrupoles brings the laser beam into a tight focus at the conversion point (CP) located 5 mm upstream of the interaction

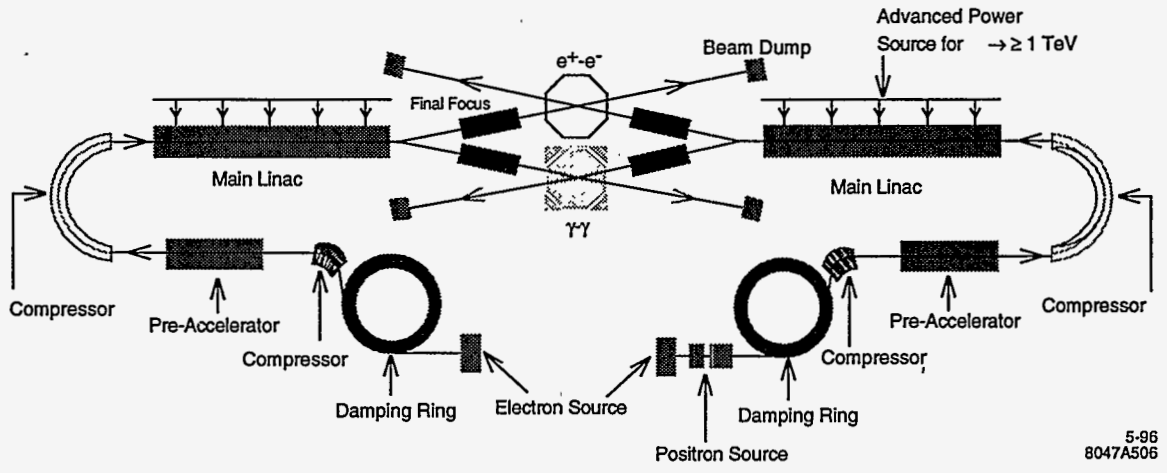


Figure B-1. *Next Linear Collider layout with a second IR for $\gamma\gamma$ collisions.*

point (IP). The laser required for the Compton conversion must have a TW of peak power and tens of kW of average power. Such a laser can be built by either combining diode pumping and chirped pulse amplification in solid-state lasers or by a free-electron laser driven by an induction linac and using chirped pulse amplification.

The phenomenon limiting the performance of the $\gamma\gamma$ collisions are different from those in the case of the e^+e^- collisions; the beamstrahlung is absent in the $\gamma\gamma$ collisions, while the e^+e^- pair creation is still important. Therefore the optimization of e-beam parameters for IR2 would be quite different from the case of IR1. Therefore it is worthwhile to revisit the damping rings and linac designs towards obtaining a smaller emittance and/or larger number of particles per bunch at a reduced pulse repetition rate if necessary. These more challenging topics are not pursued in this report.

B.2 Physics Opportunities at $\gamma\gamma$ Collider: The Higgs Sector and Other New Physics

Several review articles have been written on research that could be performed at the IR2 [Brodksy 1994, Chanowitz 1994, Ginzburg 1994]. In this section we focus on the new physics studies, especially concerning the properties of Higgs bosons. Other topics such as the study of the $t\bar{t}$ threshold region, and of the photon structure functions in the $e\gamma$ mode are omitted here.

B.2.1 The Higgs $\gamma\gamma$ Partial Width

One of the most interesting physics programs at a $\gamma\gamma$ collider is the measurement of the Higgs boson partial width into $\gamma\gamma$. This partial width is sensitive to physics beyond the Standard Model because heavy particles whose masses originate in the Higgs mechanism do not decouple in the one-loop diagram [Gunion 1993]. For

instance, the fourth generation contributes to the partial width even in the limit where its mass becomes infinity. Therefore, a study of the partial width may indicate new physics.

A plot of $\Gamma(H \to \gamma\gamma)$ with new physics contributions is shown in Figure 2.1 of Ref. [Gunion 1993].

Here we discuss measurements of the partial width $\Gamma(H \to \gamma\gamma)$ for two cases separately, if H decays predominantly into $b\bar{b}$ ($m_H \leq 150\,\text{GeV}$ in the Standard Model) or WW and ZZ (for heavier Higgs in the Standard Model).

$\gamma\gamma \to H \to b\bar{b}$

For this measurement, it is preferred to use the electron helicity to make the photon energy spectrum peaked at its high end, and fix the ee center of mass energy to put the Higgs mass at the peak of the $s_{\gamma\gamma}$ spectrum. The most dominant background is the continuum production $\gamma\gamma \to b\bar{b}$ and $c\bar{c}$. It was pointed out [Barklow 1990] that a $J = 0$ combination of the photon helicities $(++)$ or $(--)$ significantly suppresses the background as m_b^2/s (m_c^2/s) because of the fermion chirality conservation. Furthermore, the continuum production prefers small angles while the signal is isotropic. A cut on the angle, *e.g.*, $|\cos\theta_b| < 0.7$, significantly suppresses the background. For a Standard Model Higgs boson in the intermediate mass range, $80 \leq m_H \leq 150\,\text{GeV}$, the signal cross section is 300–1000 fb level and well above the background after the cuts. Statistical accuracy of $\Gamma(H \to \gamma\gamma)$ is ~5% with an integrated luminosity of 20 fb^{-1}, assuming a b-tagging efficiency of 50% and a $c\bar{c}$-to-$b\bar{b}$ acceptance ratio of 5% [Borden 1993a].

The above analysis has two limitations. One is that the resolved photon contribution to the continuum $b\bar{b}$ production may be important [Eboli 1993]. It can, however, be suppressed by choosing the center-of-mass energy such that m_H lies at the maximum $\gamma\gamma$ energy. The $b\bar{b}$ events produced by resolved photon contribution has typically much lower energy and hence can be suppressed using a visible energy cut. The other limitation is $b\bar{b}g$ or $c\bar{c}g$ final states [Borden 1994, Jikia 1994a]. They appear at higher orders in α_s, but do not have m_f^2/s suppression even in a $J = 0$ helicity combination. Still, suitable kinematical cuts eliminate most of the backgrounds from bbg and ccg [Borden 1994], requiring at least five tracks with a large impact parameter >4 sigma (with $\sigma \simeq 30\,\mu\text{m}$) to reject cs. It was also pointed out that the previous studies did not optimize the center-of-mass energy to reduce the backgrounds. By putting m_H on the top of the $s_{\gamma\gamma}$ spectrum, a measurement of $\Gamma(H \to \gamma\gamma)$ is possible with 6% accuracy 20 fb^{-1} [Watanbe 1995].

$\gamma\gamma \to H \to ZZ$

If the Higgs boson is heavier and decays predominantly into vector bosons WW or ZZ, the $b\bar{b}$ mode discussed above is not useful. One cannot use the WW mode either because of its huge tree-level production cross section of ~100 pb. Even the ZZ mode suffers from one-loop production via the W-loop [Jikia 1993], but it is manageable for $m_H \leq 350\,\text{GeV}$. One can measure the partial width $\Gamma(H \to \gamma\gamma)$ at 10% level for small m_H, but the signal is lost for $m_H < 350\,\text{GeV}$ [Borden 1993b].

B.2.2 Higgs CP Eigenvalue

A measurement of the Higgs boson property special at a $\gamma\gamma$ collider is to decide definitively whether a particular Higgs boson is CP even or odd [Grzadkowski 1992, Kramer 1994]. The basic idea is that a CP-even Higgs boson, H^0, couples to the photon with $\mathcal{L} \sim H^0(\vec{E}\cdot\vec{E} - \vec{B}\cdot\vec{B})$, while a CP-odd one, A^0, couples

with $\mathcal{L} \sim A^0 \vec{E} \cdot \vec{B}$, where $\vec{E}$ is the electric and $\vec{B}$ the magnetic field strength of photon. If the two colliding photon beams are linearly polarized, their polarizations have to be parallel to produce a CP-even state H^0 while they ought to be perpendicular to produce a CP-odd state A^0. Therefore, the asymmetry

$$A \equiv \frac{\sigma(\text{parallel}) - \sigma(\text{antiparallel})}{\sigma(\text{parallel}) + \sigma(\text{antiparallel})} \tag{B.1}$$

is $+1$ for H^0 and -1 for A^0. The studies in Refs. [Grzadkowski 1992, Kramer 1994] showed that an integrated luminosity of 100 fb^{-1} is enough to determine the CP eigenvalue if they decay dominantly into $b\bar{b}$.

B.2.3 Higgs Boson Search

An advantage of a $\gamma\gamma$ collider is that one can use full center-of-mass energy to produce Higgs bosons in s-channel, while one may need to produce them in pairs at the e^+e^- mode. For instance, the heavy CP-even Higgs H^0 and CP-odd Higgs A^0 in the minimal supersymmetric standard model (MSSM) are produced in an association $e^+e^- \to H^0A^0$, while their production with Z^0 (Z^0H^0 or Z^0A^0 final states) are suppressed if $m_{A^0} \gtrsim 300\,\text{GeV}$. On the other hand, a $\gamma\gamma$ collider can produce A^0 and H^0 states with its full center-of-mass energy, and can be used as a discovery machine even if their threshold lies beyond the e^+e^- center-of-mass energy. Using basically the same strategy in looking for the Standard Model Higgs decaying to $b\bar{b}$, one can cover a substantial region of the parameter space. The final states $t\bar{t}$ or $H^0 \to h^0h^0$ can be used as well. The $t\bar{t}$ final state suffers from continuum background, and the $t\bar{t}g$ final state has been calculated [Kamal 1995]. Even though more studies are necessary, the detection seems to be feasible.

There is a potential problem with their supersymmetric decay modes. For instance, $A^0 \to \tilde{\chi}_1^0\tilde{\chi}_1^0$ may be open and dominate the decay branching ratio, which does not leave any visible signature [Gunion 1995].

For a light Standard Model Higgs boson decaying into $b\bar{b}$, the high-energy part of the broad-band photon energy spectrum from the γ-conversion does an excellent job for the discovery. With $\sqrt{s_{ee}} = 500\,\text{GeV}$, 10 fb^{-1} and broad-band spectrum, one can observe Higgs bosons for $m_H = 110$–$140\,\text{GeV}$ [Baillargeon 1995]. Of course, with lower center-of-mass energy (*e.g.,* 350 GeV), discovery reach extends to lower mass (90 GeV). This capability is desired especially when the $\gamma\gamma$ collision operates at the second collision point of an e^+e^- collider and one cannot vary the center-of-mass energy freely.

For heavy Standard Model Higgs bosons decaying predominantly into WW and ZZ, one needs to go to WWH final state, and it requires a large luminosity. For instance, with $\sqrt{s_{ee}} = 1.5\,\text{TeV}$ and 200 fb^{-1}, one can observe up to 700-GeV Higgs bosons [Jikia 1994b, Cheung 1994].

B.2.4 Strongly Interacting Electroweak Sector

The study of the electroweak symmetry breaking sector is difficult if it is strongly interacting as it is for any other colliders, *e.g., pp* or e^+e^-. The main reasons for the difficulty are that there is no light degrees of freedom in the sector and the only signature is the tail of strong interaction among longitudinal W-boson or top quark above the TeV scale. Even though a $\gamma\gamma$ collider has a huge cross section to produce a W-pairs, they are predominantly transversely polarized and are not sensitive to the strong interactions.

There are discussions to study $WWWW$ or $WWZZ$ final states at $\gamma\gamma$ colliders [Jikia 1994b, Cheung 1994]. However, their study typically requires center-of-mass energy *higher* than their e^+e^- cousin, and probably

not of a main target of the first stage $\gamma\gamma$ collider. For instance, it was discussed that at $\sqrt{s_{ee}} = 2\,\text{TeV}$, one needs a luminosity more than 200 fb^{-1} to observe strong interaction among the W_Ls [Jikia 1994b].

Another possible signature of the strongly-interacting electroweak sector is the energy dependence of the $t\bar{t}$ production cross section. Suppose the top quark mass is generated by an effective four-fermion interaction, $\mathcal{L} \sim \frac{1}{\Lambda^2} t\bar{t}Q\bar{Q}$, where Λ is the scale of extended technicolor or its analog, and Q is a techniquark which condenses to break electroweak symmetry. Due to a loop diagram of techniquarks Q, the $t\bar{t}$ production cross section can be significantly reduced [Asaka 1995]. A possible techni-eta meson may be observed at the $\gamma\gamma$ mode as well [Tandean 1995]. Clearly, more discussions and studies are necessary for the case of the strongly interacting electroweak sector.

B.2.5 Supersymmetry

If supersymmetry exists, charged superparticles can be produced at a $\gamma\gamma$ collider with reasonable cross sections. For many of them, W-pair is the main background. For instance, a pair of sleptons $\tilde{l}^+\tilde{l}^-$ can be produced which decays into $l^+\tilde{\chi}_1^0 l^-\tilde{\chi}_1^0$. While W-pairs can lead to the same signature, one can obtain a relatively clean sample of signals after suitable cuts [Kon 1993]. Mass measurement of sleptons and neutralinos can be done at a 5% level with 20 fb^{-1} [Murayama 1994]. Charginos suffer more from the W-pair background, and more studies are necessary.

The backscattered laser beam allows us to use the $e\gamma$ mode to extend the discovery reach of selectron $\tilde{e}$, and in the process $e\gamma \to \tilde{e}\tilde{\chi}_1^0$. A selectron can be produced even if the e^+e^- center-of-mass energy is below the threshold of its pair production, up to $m_{\tilde{e}} < \sqrt{s_{e\gamma}} - m_{\tilde{\chi}_1^0}$ [Kon 1992a].

B.2.6 Compositeness

If some of the particles in the Standard Model are a composite of more fundamental objects, they exhibit either (1) excited states decaying into the ground state by γ, Z, or g radiation, or (2) anomalous interactions at the low-energy limit of their form factors.

If the electron is a composite, one can look for its excited state e^* in the process $e\gamma \to e^* \to e\gamma$ [Kon 1992b].

If a W-boson is a composite, it may have an anomalous magnetic moment or electric quadrupole moment (assuming CP invariance). The huge W-pair production cross section from $\gamma\gamma$ allows us a precise measurement of such anomalous moments [Yehudai 1991, Gounaris 1995]. Another process $e^-\gamma \to \nu_e W^-$ can be also used [Yehudai 1990, Raidal 1995]. One can obtain constraints complementary to that from an e^+e^- mode.

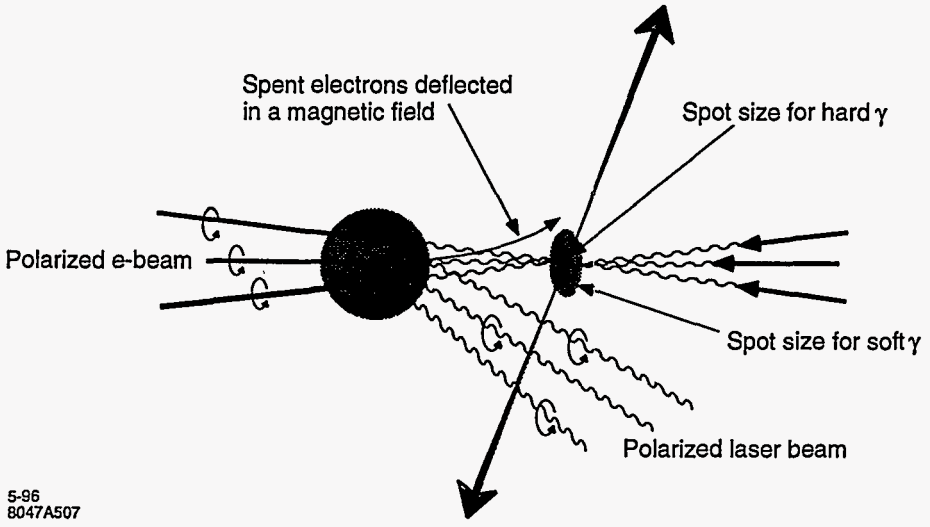


Figure B-2. *General scheme of $\gamma\gamma$ collision.*

B.3 Major Parameters

B.3.1 Basic Scheme

The basic scheme of the IR2 for $\gamma\gamma$ collisions is shown in Figure B-2. Two electron beams from their respective final-focus system (FFS) are heading toward the IP. At a location a short distance (5 mm for this design) upstream from the IP, referred to hereafter as the conversion point (CP), a laser beam is focused and Compton-backscattered by the electrons, resulting in a high-energy beam of photons. The photon beam follows the original electron motion (with a small angular spread of order $1/\gamma$) and arrives at the IP in a tight focus. It collides at the IP with an opposing high energy photon beam similarly produced by a second electron beam.

The spent electron beam, following its interaction at the CP, together with the photon beam will cause background γe^- and e^-e^- events as well as producing copious beamstrahlung photons, which will further increase the backgrounds, in interaction with the other electron beam. Extensive detector simulation is required to determine whether these background poses significant problem for a given experiment. One way to reduce the background is to place a bending magnet [Ginzburg 1983] between the CP and the IP, so that the spent beams miss each other. Implementing the sweeping magnet in the tight space in the IR2 is a challenging problem.

B.3.2 Laser Parameters

The laser beam must be chosen to optimize the generation of the gamma-rays via Compton scattering at the CP. The relevant laser parameters at the CP are summarized in Table B-1. The wavelength and the peak intensity of the required laser are similar to that available in the E-144 experiment at SLAC [Heinrich 1991].

Wavelength	$\lambda = 1.053\,\mu$m
Micropulse energy	$A = 1$ J
Repetition rate	The same as the electron beam pulse rate (90 micropulses separated by 1.4 ns repeating at 180 Hz)
Rayleigh length	$Z_R = 0.1$ mm
Rms spotsize at waist	$\sigma_{Lx} = \sigma_{Ly} = 2.90\,\mu$m
Rms angular divergence	$\sigma'_{Lx} = \sigma'_{Ly} = 28.9$ mr
Rms micropulse length	$\sigma_{Lz} = 0.23$ mm
Peak intensity	$\approx 1 \times 10^{18}$ W/cm^2
Peak power	0.5 TW
Average power	16.2 kW
Transverse coherence	Near diffraction limited
Polarization	Fully polarized with helicity switching capability

Table B-1. *Laser parameters for a reference design at $E_{cm} = 500$ GeV.*

However, the average power is two orders of magnitude larger than what is currently available. In computing the laser spotsize, it is useful to remember the following correspondence between the electron beam and the light beam:

$$\epsilon_x, \epsilon_y \leftrightarrow \frac{\lambda}{4\pi} \tag{B.2}$$

$$\beta_x^*, \beta_y^* \leftrightarrow Z_R \tag{B.3}$$

The quantity Z_R is known as the Rayleigh length in optics literature. Thus, the rms spotsize $\sigma_{Lx} = \sigma_{Ly}$ and the cross section Σ at the focus are respectively given by

$$\sigma_{Lx} = \sqrt{\frac{\lambda}{4\pi} Z_R}, \tag{B.4}$$

$$\Sigma \equiv 2\pi\sigma_{Lx}^2 = \frac{1}{2}\lambda Z_R \tag{B.5}$$

The considerations leading to the parameters in Table B-1 are given in B.4.

B.3.3 Electron Beam Parameters

The electron beam parameters for the reference design at 500-GeV CM energy are summarized in Table B-2.

The $\gamma\gamma$ or γe^- luminosity is approximately proportional to the e^-e^- geometric luminosity. Although we can in principle reexamine the design of the damping ring–linac complex from the point of maximizing the geometric e^-e^- luminosity, we have chosen for this initial design of the IR2 to use the same electron beam parameters before the IR1 for the e^+e^- collision. However, the beta-functions at the IP are chosen differently from the ones for the e^+e^- collision—with a relaxed β_y^* and a tighter β_x^*. This is due to the fact that the CP is separated from the IP (by 5 mm) to suppress the low-energy part of the $\gamma\gamma$ luminosity spectrum. The separation will introduce an increase in the spotsize of the gamma-ray photons at the IP due to their angular spread $\approx 1/\gamma$ relative to the electron beam. The vertical β_y^* could be larger than that in the case of the e^+e^- collision. It is necessary to reduce β_x^* to compensate the reduction in the luminosity when β_y^* is increased. With $\beta_x^* = \beta_y^* = 1$ mm, the geometric luminosity would be the same as in the case of the e^+e^- collision.

Luminosity goal	$\sim 10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for 10 % BW
	$\sim 5\times 10^{33}\,\mathrm{cm}^{-2}s^{-1}$ for broad band
Beam parameters before FFS	The same as e^+e^- design:
Electron energy	250 GeV
Rep. rate	90 bunches separated by 1.4 ns, 180 Hz
Particles per bunch	$N_e = 0.65\times 10^{10}$
Normalized rms emittance	$\gamma\epsilon_x = 5\times 10^{-6}$ mr, $\gamma\epsilon_y = 8\times 10^{-8}$ mr
Beta function at the IP	$\beta_x^* = \beta_y^* = 0.5$ mm
Rms spotsize at the IP	σ_x^*/σ_y^* =71.5/9.04 nm
Rms spotsize at the CP	σ_x^c/σ_y^c =718./90.9 nm
Rms angular divergence	σ_x'/σ_y' = 143./18.1 μrad
Rms bunch length	$\sigma_z = 0.1$ mm
Polarization	Fully polarized with helicity switching capability
Collision scheme	Vertical offset or sweeping magnet
CP–IP distance	b=5 mm
Crossing angle	$\phi_c \leq 30$ mr

Table B-2. *Electron beam parameters for a reference design at $E_{CM} = 500$ GeV*

The design goal for the FFS for the $\gamma\gamma$ collision is $\beta_x^* = \beta_y^* = 0.5$ mm. Note that, contrary to the e^+e^- case, there are no constraints on the beam profile at the IP arising from beamstrahlung effects.

The large crossing angle, 30 mr, is necessary in the collision scheme without a sweeping magnet, due to the large disruption of the low-energy electrons (coming from the high-order multiple scattering in the CP). The disruption is smaller when a sweeping magnet is employed.

A more detailed considerations leading to the parameters in Table B-2 are given in Section B.5.

B.4 CP Issues

B.4.1 Optimization of the Laser Parameters

Compton scattering of laser beam by relativistic electron beams is an efficient way to generate gamma-ray photons [Arutyunian 1963]. A review of the relevant kinematics can be found in [Telnov 1990].

The energy of the Compton-scattered photon is maximum when the scattered photon is in the direction of the incoming electron, *i.e.*, in the backscattering direction. The maximum energy is given by

$$\omega_{max} = \frac{x}{x+1}E_0, \tag{B.6}$$

where

$$x = \frac{4E_0\hbar\omega_0}{m^2c^4} \simeq 15.3\left[\frac{E_0}{TeV}\right]\left[\frac{\hbar\omega_0}{eV}\right] . \tag{B.7}$$

Here ω_0 is the laser frequency and E_0 is the initial energy of electrons. Hence, the energy of the backscattered photon increases with increasing value of the parameter x, but if x is larger than 4.8, high-energy photons

can be lost due to e^+e^- pair creation in collision with unscattered laser photons (Breit-Wheeler process). Thus, the optimum value is $x = 4.8$, corresponding to the maximum photon energy $\omega_{max} = 0.81E_0$. For $E_0 = 0.25\,\mathrm{TeV}$, this leads to a laser wavelength of about 1 μm. Thus it is convenient to choose the wavelength of Nd:Glass laser, 1.05 μm.

Neglecting multiple scattering, and assuming that the laser profile seen by each electron is the same, the "conversion" probability of generating high-energy gamma photons per individual electron can be written as

$$n_\gamma = 1 - \exp(-q) \quad . \tag{B.8}$$

If the laser intensity along the axis is uniform

$$q = \sigma_c \frac{N_L}{\Sigma} = \frac{\sigma_c A}{\hbar\omega_o \Sigma} = \frac{\sigma_c I \tau_L}{\hbar\omega_0} \quad . \tag{B.9}$$

Here σ_c is the Compton cross section, which for $x = 4.8$ is $1.75 \times 10^{-25}/cm^2$, N_L is the number of the laser photons, Σ is the transverse area of the laser spot, A is the the laser pulse energy, I is the laser intensity (power per unit area), and τ_L is the laser pulse length. The pulse energy corresponding to $q = 1$, *i.e.*, a conversion probability of 65%, is given by $A_0 = \hbar\omega_0\Sigma/\sigma_c \approx \hbar\pi c Z_R/\sigma_c$. The increase in the conversion probability with a laser pulse energy larger than A_0 is relatively small, and, furthermore, is expensive due to higher laser power requirements. Therefore we should choose $A \sim A_0$, *i.e.*, $q \sim 1$. Thus the pulse energy is minimized when the laser spot is focused tightly to match the electron pulse shape. However, the focusing may not be made arbitrarily strong: the laser intensity I could become so large that nonlinear QED effects may spoil the conversion process.

The nonlinear effect is characterized by the quantity

$$\eta = \frac{eE}{\omega_0 mc},\ \eta^2 = 0.4 \left[\frac{I}{10^{18}\mathrm{W/cm}^2}\right] \left[\frac{\lambda}{1.054\mu\mathrm{m}}\right]^2 \tag{B.10}$$

When $\eta^2 \gtrsim 1$, two or more laser photons can be scattered at the same time. The the maximum energy of the backscattered high energy photon in the non-linear Compton scattering involving n laser photons is given by

$$\omega_{max}^n = \frac{nx}{1 + \eta^2 + nx} E_0 \quad . \tag{B.11}$$

For $n = 1$, which corresponds to the single photon process, the maximum photon energy is smaller than that given by the linear approximation, Eq. B.6. This is not desirable since the $n = 1$ photons are usually the most useful ones. Another effect, which is essentially quantum mechanical, is that the pair production can now proceed via multiphoton scattering with a gamma-ray photon, leading to a depletion of the gamma-ray flux. In this design the laser intensity I is kept below 1×10^{18} W/cm^2.

A formula for the conversion efficiency n_γ, neglecting the nonlinear effect and multiple scattering, but taking into account the fact that different electrons see different laser profiles during the interaction, is derived in Section B.4.3. Given the laser pulse energy A and the intensity I, Eq. B.21 can be used to find the optimum value of Z_R and σ_{Lz} corresponding to the maximum conversion efficiency. The case for $A = 1$ J, $I = 1 \times 10^{18}$ W/cm^2, and the rms electron pulse length $\sigma_z = 0.1$ mm is shown in Figure B-3. It is seen that a maximum conversion efficiency $n_\gamma = 0.68$ can be achieved with $Z_R = 0.1$ mm. The corresponding laser pulse length (rms) is $\sigma_{Lz} = 0.23$ mm(which is larger than σ_z for the electrons).

The peak power corresponding to $A = 1$ J and pulse length $\tau \sim 2\sigma_{Lz}/c \simeq 1.8$ ps is about 0.5 TW. With the NLC pulse format of 90 micropulses repeated at 180 Hz, the required average laser power is 16.2 kW, which

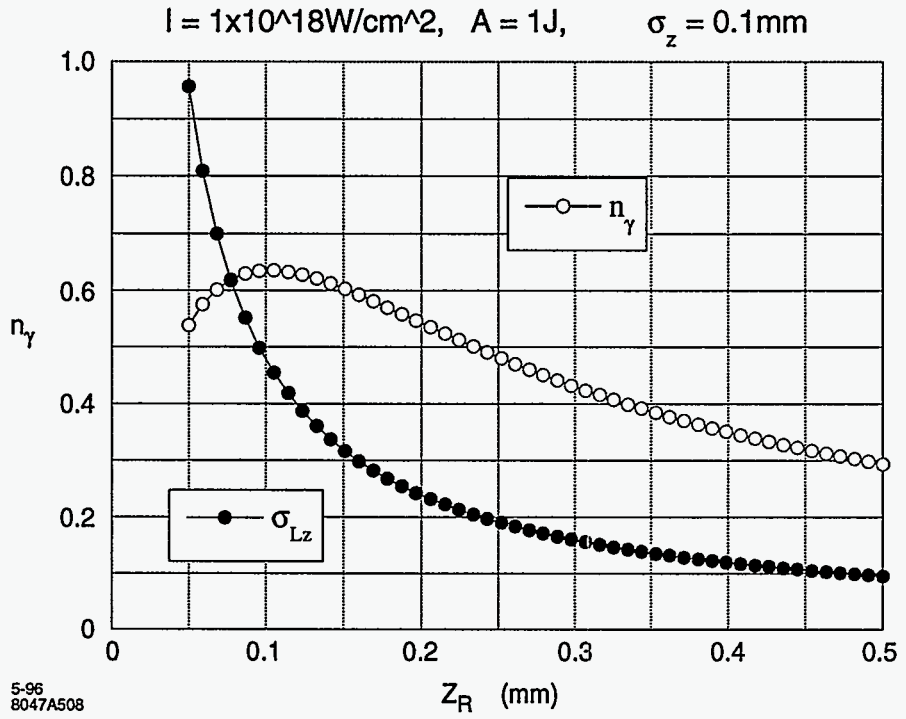


Figure B-3. *Compton Conversion efficiency as a function of z_r at fixed A and I.*

is rather high. The power can be reduced significantly if laser pulses can be reused by multipass optics or by storing the laser energy in an optical cavity.

With the laser intensity $I = 1\times10^{18}W/\,\mathrm{cm}^2$, the nonlinear effect is not negligible as we will see in Section B.6. To study the dependence of the conversion efficiency on A and I, we show in Figure B-4 the maximum conversion efficiency as a function of the laser pulse energy and the laser intensity. Figure B-5 and B-6 give respectively the corresponding Rayleigh length and the rms laser pulse length. Figure B-7 gives the contour line corresponding to the conversion efficiency $\eta_\gamma = 0.65$, showing that as the intensity is reduced the pulse energy must be increased to maintain the same conversion efficiency. As an example, with $A = 2\mathrm{J}$ and $I = 5\times10^{17}W/\,\mathrm{cm}^2$, one can obtain the maximum conversion efficiency $n_\gamma = 0.648$ with $Z_R = 0.201\,\mathrm{mm}$ and $\sigma_{LZ} = 0.46\,\mathrm{mm}$. Another example is $A = 3J$ and $I = 3.3\times10^{17}W/\,\mathrm{cm}^2$ for which $n_\gamma = 0.65$, $Z_R = 0.347\,\mathrm{mm}$ and $\sigma_{LZ} = 0.76\,\mathrm{mm}$. The non-linear effects in this case are smaller than the $A = 1J$, $I = 1\times10^{18}W/\,\mathrm{cm}^2$ example, but the pulse energy is higher. Since the laser power is expensive, we will adopt in this report $A = 1\mathrm{J}$, $I = 1\times^{18}W/\,\mathrm{cm}^2$ as the reference case.

Transverse coherence of the laser beam is important in obtaining a diffraction-limited focal spot. Versatile polarization control is also important; the helicity of the laser light should be opposite to that of the electron beam to obtain a higher conversion rate and the γ photon spectrum peaked around ω_m. Controlling the γ photon polarization by controlling the laser photon polarization is an important technique for many $\gamma\gamma$ or γe^- experiments [Barklow 1990]. Switching of helicity is proposed to characterize all helicity components of the luminosities [Telnov 1995].

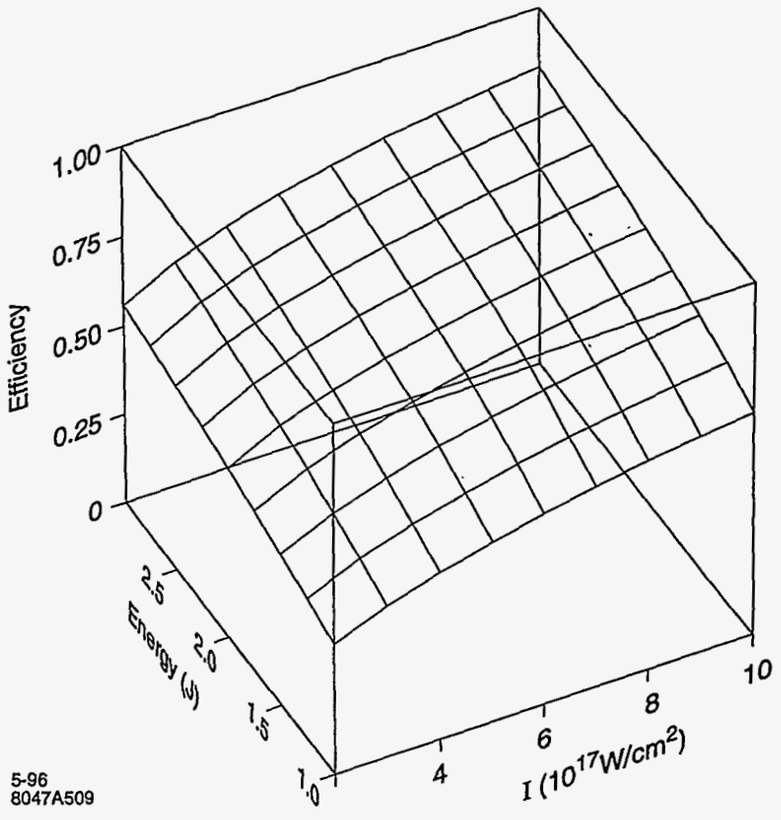


Figure B-4. *Optimized conversion efficiency as a function of pulse energy A and pulse intensity I.*

B.4.2 Low-Energy Electrons Due to High-Order Multiple Scattering

With the intense laser pulse required for an efficient conversion, the probability of multiple scattering is high, giving rise to soft electrons and photons. The multiple scattering process is roughly described by the Poisson distribution. A k-fold scattering has the probability

$$P_k \sim \frac{q^k e^{-q}}{k!} \quad , \tag{B.12}$$

and gives rise to a minimum electron energy E^k where [Telnov 1990]

$$E^k \sim \frac{E_0(1+\eta^2)}{1+\eta^2+kx} \quad . \tag{B.13}$$

For ten-fold multiple scattering, $k = 10$, the electron energy is about 2% of the incoming energy. There are about 1000 such particles, which could cause significant background signals if they are allowed to hit the quadrupole faces.

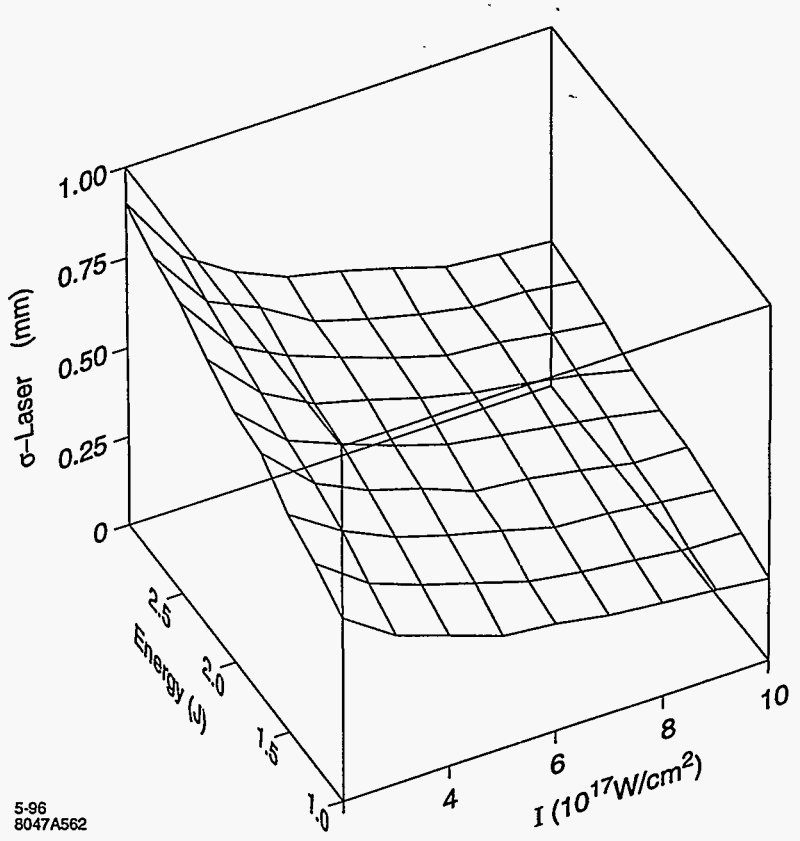


Figure B-5. *Optical pulse length corresponding to the optimized conversion efficiency.*

B.4.3 Compton Conversion Efficiency

Let an electron beam collide with a laser beam of density $n_L(\mathbf{x},t)$. The density of the electron beam will be written as $n_e(\mathbf{x}-\mathbf{v}t,t)$ to indicate that the beam moves with velocity $\mathbf{v}$. The number of scatterings in the space-time element $dxdt = dx'dt$, where $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$ is given by [Landau 1987]

$$d\nu = \sigma_c v_{rel} n_e(\mathbf{x}',t) n_L(\mathbf{x},t) dx'dt \quad . \tag{B.14}$$

Here σ_c is the Compton cross section, $v_{rel} = \sqrt{(\mathbf{v}-\mathbf{v}_L)^2 - (\mathbf{v}\times\mathbf{v}_L)^2}$ is the relative velocity, $\mathbf{v}_L = c\mathbf{n}$, and $\mathbf{n}$ is the direction of the laser propagation. Since we are interested in the case where the loss of the laser photons can be neglected, we may assume that the scattering does not change $n_L(\mathbf{x},t)$. On the other hand, the probability of Compton scattering per electron is large, and the electron after producing a γ photon after scattering may be regarded as lost. The rate of the loss is given by Eq. B.14 and can be written as follows:

$$-dn_e(\mathbf{x}',t)dx' = v_{rel}\sigma_c n_e(\mathbf{x}',t) n_L(\mathbf{x}'+\mathbf{v}t,t) dx'dt \quad . \tag{B.15}$$

From this, we derive

$$\frac{\partial}{\partial t} n_e(\mathbf{x}',t) = v_{rel}\sigma_c n_e(\mathbf{x}',t) n_L(\mathbf{x}'+\mathbf{v}t,t) \quad . \tag{B.16}$$

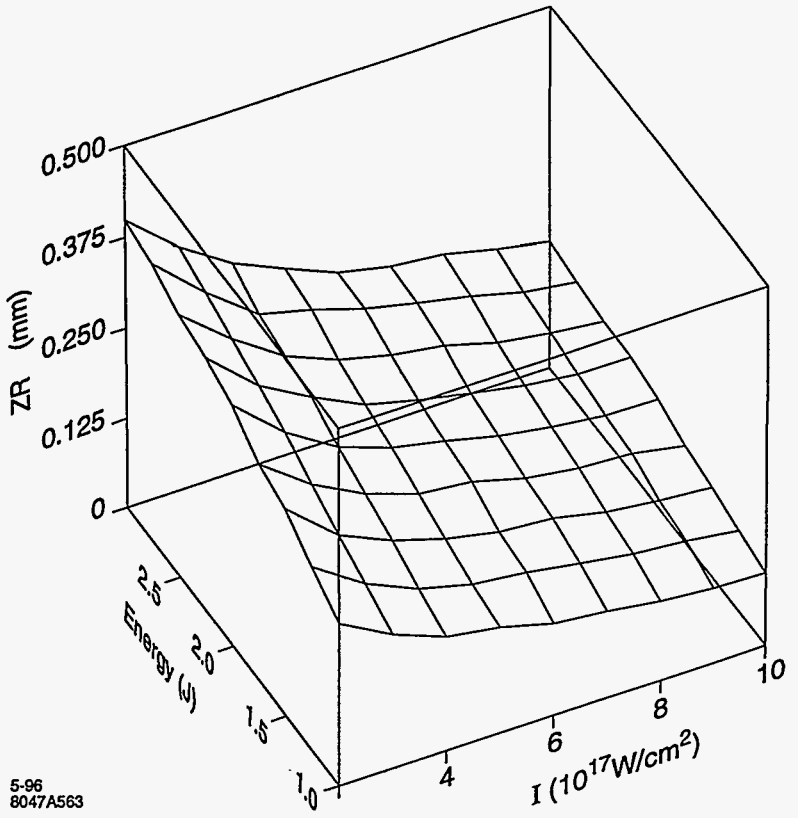


Figure B-6. *Rayleigh length corresponding to the optimized conversion efficiency.*

The solution of Eq. B.16 is

$$n_e(\mathbf{x}',t) = n_e^i(\mathbf{x}') \exp\left[-\int_{-\infty}^{t} v_{rel}\sigma_c n_L(\mathbf{x}'+\mathbf{v}t',t')dt'\right] \quad . \tag{B.17}$$

Here $n_e^i(\mathbf{x}')$ is the initial electron distribution. The total number of the γ photons generated is obtained by integrating the above equation:

$$N_\gamma = \int d\mathbf{x}' n_e^i(\mathbf{x}')\left(1-\exp\left[-\int_{-\infty}^{\infty} v_{rel}\sigma_c n_L(\mathbf{x}'+\mathbf{v}t',t')dt'\right]\right) \quad . \tag{B.18}$$

The transverse dimension of the electron beam is usually much smaller than that of the laser beam. For an electron beam traveling at an angle θ with respect to the z-axis, we have $x' = x + ct\sin\theta$, $y' = y$, and $z' = z - ct\cos\theta$. The initial electron density can be written as

$$n_e^i(\mathbf{x}') = \frac{N_e}{\sqrt{2\pi}\sigma_z}\delta(x')\delta(y')\exp(-z'^2/2\sigma_z^2) \quad . \tag{B.19}$$

Here N_e is the total number of electrons in the bunch, and σ_z is the rms bunch length. The laser pulse propagating along the negative z-direction can be written as

$$n_L(x,y,z,t) = \frac{N_L}{(2\pi)^{3/2}\sigma_{Lz}} \frac{\exp\left(-\frac{1}{2}\left[\frac{x^2+y^2}{\sigma_{Lx}(z)^2}+\frac{(z+ct)^2}{\sigma_{Lz}^2}\right]\right)}{\sigma_{Lx}(z)^2} \quad . \tag{B.20}$$

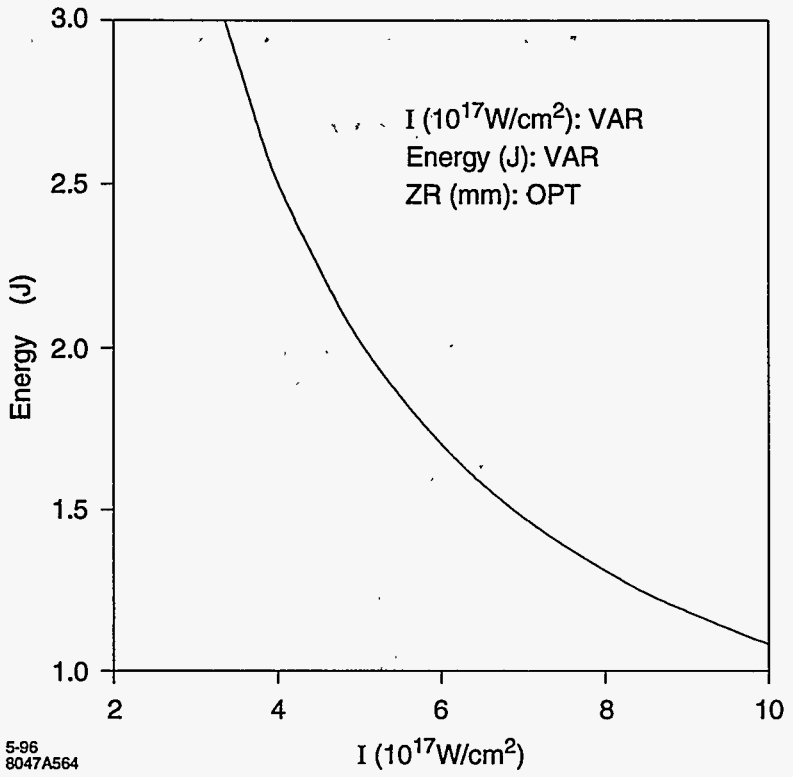


Figure B-7. *Optical pulse energy versus intensity at conversion efficiency = 0.65.*

Here N_L is the number of the laser photons in the pulse, σ_{Lz} is the rms length, $\sigma_{Lx}(z)^2 = \sigma_{Lx}^{*2}(1+(z/Z_R)^2)$, σ_{Lx}^* is the rms spotsize, and Z_R is the Rayleigh length. The last two quantities are related by $\sigma_{Lx}^{*2} = Z_R\lambda/4\pi$, where λ is the laser wavelength.

In the special case of head-on collision, $\theta = 0$, and one derives from the above formulas the result

$$n_\gamma \equiv \frac{N_\gamma}{N_e} = 1 - \frac{1}{\sqrt{2\pi}\sigma_z}\int dz \exp\left(-\frac{z^2}{2\sigma_z^2} - U(z)\right) \tag{B.21}$$

where

$$U(z) = \frac{4\sigma_c N_L}{\sqrt{2\pi}\lambda Z_R \sigma_{Lz}}\int ds \frac{\exp\left[-\frac{2(s-z/2)^2}{\sigma_{Lz}^2}\right]}{1+s^2/Z_R^2} \tag{B.22}$$

The spectrum of the photons is given by

$$\frac{dn_\gamma}{dy} = n_\gamma \frac{1}{\sigma_c}\frac{d\sigma_c}{dy} \tag{B.23}$$

where $y = \hbar\omega/E_o$, and $d\sigma_c/dy$ is the differential scattering cross section of the Compton process, given, for example, by Eq. 3 of Ref. [Telnov 1995].

Although we have neglected the multiple scatterings in the derivation here, Eq. B.23 is expected to be valid near $y \sim y_{max} = x/(1+x)$ even for a "thick" target. On the other hand, the effect of multiple scattering will be large in the soft photon region, $y \sim 0$.

B.5 IP Issues

B.5.1 Optimization of Spectral Luminosity

The total $\gamma\gamma$ luminosity is approximately given by $n_\gamma^2 \simeq 0.4$ times the geometric e^-e^- luminosity (with zero offset). However, the spectral luminosity of the $\gamma\gamma$ collision depends strongly on the distance b between the CP and the IP. This is because the energy of the gamma photon depends sensitively on the scattering angle, being maximum in the original electron direction and decreasing rapidly away from this direction. Introducing the parameter $\rho = b/\gamma\sigma_y^*$, where σ_y^* is the vertical rms electron beam size (assumed to be smaller than the horizontal size) in the absence of the CP, the spectral luminosity is broadly distributed as a function of the c.m. energy of the two-photon system when $\rho \ll 1$. As ρ is increased, the low-energy part of the luminosity spectrum becomes suppressed due to the larger spotsize occupied by low-energy photons. Thus, the luminosity spectrum develops a well defined peak at the high-energy end with a bandwidth of about 20% when $\rho \geq 1$. This region is also characterized by a high degree of polarization. For most applications, one would choose $\rho \simeq 1$ to obtain a narrow spectrum without suffering a large luminosity reduction. In our case, this corresponds to $b \sim 5\,\mathrm{mm}$. The spectral peak at the high-energy end of the invariant mass distribution accounts for about 20% of the total $\gamma\gamma$ luminosity, or about 10% of the geometrical e^-e^- luminosity.

For e^+e^- collisions, the beam spot at the IP is normally designed to be flat to minimize the beamstrahlung effect. In $\gamma\gamma$ collisions, the vertical beam size, which is determined by the condition $\rho \simeq 1$, is larger than that in the e^+e^- collisions for a reasonable value of the CP-IP distance b. The horizontal spotsize should be reduced in proportion to achieve a comparable luminosity. Thus the FFS for $\gamma\gamma$ collision must provide a value of β_x^* which is smaller and β_y^* which is larger than the corresponding values for the e^+e^- design. With $\beta_x^* = \beta_y^* = 1\,\mathrm{mm}$, the geometric luminosity would be the same as in the case of the e^+e^- collision. We therefore aim for $\beta_x^* = \beta_y^* = 0.5\,\mathrm{mm}$ for the $\gamma\gamma$ collision. In doing so, a proper account should be made of the Oide effect as well as the constraint that β_x and β_y be larger than the bunch length. A design of the FFS similar to that of the e^+e^- case, but with final quadrupoles reversed in x and y is presented in Section B.9.

B.5.2 Polarization

By varying the polarization of the electron and the laser beams, the polarization of the high-energy photon beams can be tailored to fit the needs of the individual experiment. Controlling the polarization is also important in sharpening the spectral peak in the $\gamma\gamma$ luminosity. Due to the polarization dependence of the Compton scattering, the spectral peak present in the case of $\rho \geq 1$ is significantly enhanced by choosing the helicity of the laser photons to be of the opposite sign to the helicity of the electrons.

B.5.3 Collision of the Spent Electron Beam

The background due to γe^- and e^-e^- collisions as well as the collision of the beamstrahlung photons is large if the spent electron beams are allowed to collide at the IP. The collisions of these particles would also produce positrons and minijets. These unwanted collisions give rise to the background events. Whether these backgrounds pose a significant problem will depend on the nature of the particular experiment, and can only be evaluated after detailed detector simulation.

Among the background events, the γe^- collisions appear to be the most significant, with a luminosity roughly equal to the $\gamma\gamma$ luminosity. The ee luminosity due to the collision of the "spent" electrons (*i.e.*, after Compton conversion) is suppressed significantly (by a factor of 5) due to the fact that the spectrum of the spent electron is broad, leading to a large disruption at the IP.

A way to avoid the collision of the electron beams would be to sweep the spent electrons away from the IP by an external magnetic field. The magnetic field should extend longitudinally to about 1 cm with a strength of about T. Such a magnet could in principle be designed either with a superconducting [Telnov 1990] or with a pulsed conductor [Silvestrov un]. Installing the sweeping magnet to the tight space in the interaction region with a minimum obstruction to the detector is a major challenge.

A plasma lens to overfocus the spent electron beam has also been proposed [Rajagopalan 1994]. This scheme must inject gases to produce plasma and also remove them from the interaction region. Another proposal is to arrange the electron beams to repel and miss each other entirely in "a heads-up" collision [Balakin 1994]. For this scheme to work, the electron beam intensity needs to be much higher than that contemplated in most linear collider proposals. We have not studied these options in detail in this study.

B.5.4 Disruption of Low Energy Electrons

A characteristic angle for the full-energy primary, disrupted electrons is [Hollebeek 1981]:

$$\theta_d \equiv \frac{2Nr_e}{\gamma\sigma_x} \quad . \tag{B.24}$$

Thus the main fraction of the electrons after the IP will be deflected into an angular cone given by θ_d in which γ is replaced by an average value. However, Eq. B.24 is valid only when the deflection angle is smaller than $4\sigma_x/\sigma_z$ (we assume $\sigma_x \ll \sigma_z$). For very low energy electrons for which $\theta_d\gamma e^-\sigma_x/\sigma_z$, the deflection angle is given by [Telnov 1990]

$$\theta \sim \sqrt{\frac{4\pi Nr_e}{\gamma\sigma_z}} \quad . \tag{B.25}$$

Low-energy electrons are generated by high-order multiple scattering at the CP, as discussed in Section B.4.2. For our parameters, a ten-fold or higher multiple scattering generates about one thousand electrons with energy as low as 2–3% of their initial energy . These particles will be deflected up to an angle of about 10 mr due to collision with the opposing electron beam. Since there are about 1000 such particles, which will contribute to the background signals if they are allowed to hit the quadrupole faces, the crossing angle should be larger than 10 mr plus an additional angle to clear the quadrupole faces closest to the IP. In the case of e^+e^- collisions, the quadrupole clearance is taken to be 20 mr. We therefore take the crossing angle for the $\gamma\gamma$ collision to be 30 mr. Crabbing the electron beam is essential for a $\gamma\gamma$ collider. The effect of the solenoidal field on beam collisions with a large crossing angle also needs to be studied.

B.6 Luminosity Calculations

B.6.1 Simulation Code Development

The physical processes occurring in the CP and the IP are complex and diverse, including linear and nonlinear, single and multiple Compton scattering at the CP, beamstrahlung, coherent and incoherent pair production, Bethe-Heitler and Landau Lifshitz processes at the IP. A reliable prediction of the $\gamma\gamma$ and γe^- luminosities and the backgrounds can only be done with a numerical code simulating the entire complex of CP and IP physics. It is desirable that several independent codes are available so that simulation results can be cross-checked. At the same time, simple analytical estimates for the relative importance of these processes are also highly desirable.

Ideally code for a full simulation of $\gamma\gamma$ or γe^- collisions must incorporate the following features [Chen 1995a]:

- The CP physics:
  - Linear and nonlinear Compton scattering: $e^{\pm} + laser \rightarrow e^{\pm} + \gamma$.
  - Linear and nonlinear Breit-Wheeler scattering: $\gamma + laser \rightarrow e^+e^-$.
- Beam propagation from the CP to IP including the effects of external magnetic fields (solenoidal field, sweeping magnet) and plasma lens (if any).
- The IP Physics:
  - Disruption effects; interaction of $e^{\pm}$ with the field of the opposing beam.
  - Beamstrahlung and coherent pair production via interaction of γ and $e^{\pm}$ with the collective field the opposing beam.
  - Incoherent processes (Bremsstrahlung: $ee \rightarrow ee\gamma$; Breit-Wheeler: $\gamma\gamma \rightarrow e^+e^-$; Bethe-Heitler: $e\gamma \rightarrow ee^+e^-$; Landau-Lifshitz: $ee \rightarrow eee^+e^-$).
- Beam propagation from the IP to the exit line.

The ABEL code (Analysis of Beam-beam Effects in Linear colliders) simulates the beam-beam interaction including disruption and beamstrahlung effects [Yokoya, 1986]. The code has been subsequently modified to include incoherent pair creation in the equivalent photon approximation (ABELMOD) [Tauchi 1993]. These codes were originally written for e^+e^- collisions but have been modified to simulate the e^-e^- collisions.

There are several Monte-Carlo codes for simulating the Compton conversion process: A code written by Horton-Smith [Horton-Smith phd] in connection with the E-144 experiment at SLAC, which takes into account the non-linear effect fully but is only applicable for unpolarized electrons; a code written by Ohgaki and Yokoya [Ohgaki 1995], which is based on the Compton scattering in linear approximation but is applicable for arbitrary electron and photon polarizations; a code by Telnov [Telnov 1995], which is similar to the previous one with the further approximation that all electrons see the same laser profile ("same-profile" approximation); a code recently written by Yokoya [Yokoya 1996], which takes into account the non-linearity of the Compton scattering, and can handle circularly polarized electrons.

The codes for the IP and the CP are being combined to a varying degree of sophistication to calculate the $\gamma\gamma$ and γe^- luminosities. In our preliminary calculation, we have used Telnov's code extensively, which

includes the multiple Compton scattering effects in linear Compton approximation and the same-profile approximation at the CP, deflection by external magnetic field and synchrotron radiation in the region between the CP and the IP, the beamstrahlung and the coherent pair production at the IP. A similar code has been assembled by Takahashi [Takahashi 1996] based on Ohgaki's Compton conversion package and ABEL. A more refined code incorporating Yokoya's non-linear Compton conversion and the ABEL-MOD is being assembled as a collaborative effort between Hiroshima University, KEK, SLAC and LBL. This code is referred to as CAIN 1.1. Recently, Yokoya has written a new code, named CAIN 2.0, which does not share any subroutines with ABEL [Yokoya 1996].

A simpler code (named BERT) aimed at a careful study of the transport and disruption of the two opposing electron beams with arbitrary initial energy distributions is being developed by W. Fawley. Such a code will be useful, for example, in evaluating the heads-up collision scheme [Balakin 1994] to suppress e^-e^- collisions at the the IP. The results from these codes have been cross-checked where applicable, and are found to be in reasonable agreement with each other.

B.6.2 Simulation of the CP

Figure B-8 gives the γ-photon spectrum after the CP, using the electron beam parameters in Section B.3, and the laser parameters in Section B.4. They are obtained using the code written by Yokoya [Yokoya 1996]. The top, middle and the bottom graphs correspond respectively to the case $A = 3J$, $I = 3.3 \times 10^{17}\,\mathrm{W/cm^2}$, the case $A = 2J$, $I = 5 \times 10^{17}\,\mathrm{W/cm^2}$, and the case $A = 1J$, $I = 10^{18}\,\mathrm{W/cm^2}$, For the top graph, the non-linearity is small, and the spectrum near the maximum photon energy agrees well with the theoretical formula, Eq. B.23. The non-linearity is visible but not pronounced for the middle graph. For the bottom graph however, the smearing of the spectral peak at the high-energy end of the photon is clearly seen.

B.6.3 Telnov's Simulation Results for $\gamma\gamma$, γe^- and e^-e^- Luminosities

This section summarizes the results of simulation calculations using Telnov's code. As discussed in the above, the code is valid under two assumptions: First, the laser profiles seen by all electrons are the same, and second, the non-linear effect can be neglected. The error introduced by the first assumption appears to be not significant in our parameter regime. The validity of the second assumption depends on the laser intensity as discussed in Section B.6.2: The non-linearity is negligible for the case $A = 3\,\mathrm{J}$, $I = 3.3 \times 10^{17} W/cm^2$, and significant for the case $A = 1J, I = 1 \times 10^{18} W/\,\mathrm{cm^2}$. Although the second case is adopted as the baseline design in this report, we use Telnov's simulation code in this section because it is the code currently fully debugged. The result of Telnov's simulation appears to be in general agreement with that calculated by Takahashi using the linear version of CAIN-1.

Table B-3 summarizes the main results. Here z = invariant mass of the colliding system/$2E_0$, where E_0 is the energy of the incoming electrons. The case (a) to (f) are for the collisions at various vertical offset Δy without the sweeping magnet. The case (d) is the same as (c) but suppressing the beamstrahlung at the IP. In the case (g), there is a 1-T sweeping magnet. The electron and the laser parameters are those given in Section B.3 and B.4, respectively. However, the distance between the CP and the IP is taken to be 7.8 mm for case (g), while it is 5 mm for all other cases.

Even without a sweeping magnet, the e^-e^- luminosity is significantly reduced, a factor of five already at a small offset, $\Delta y = 0.2\sigma_y$, and the reduction increases slowly as a function of the offset. The $\gamma\gamma$ luminosity

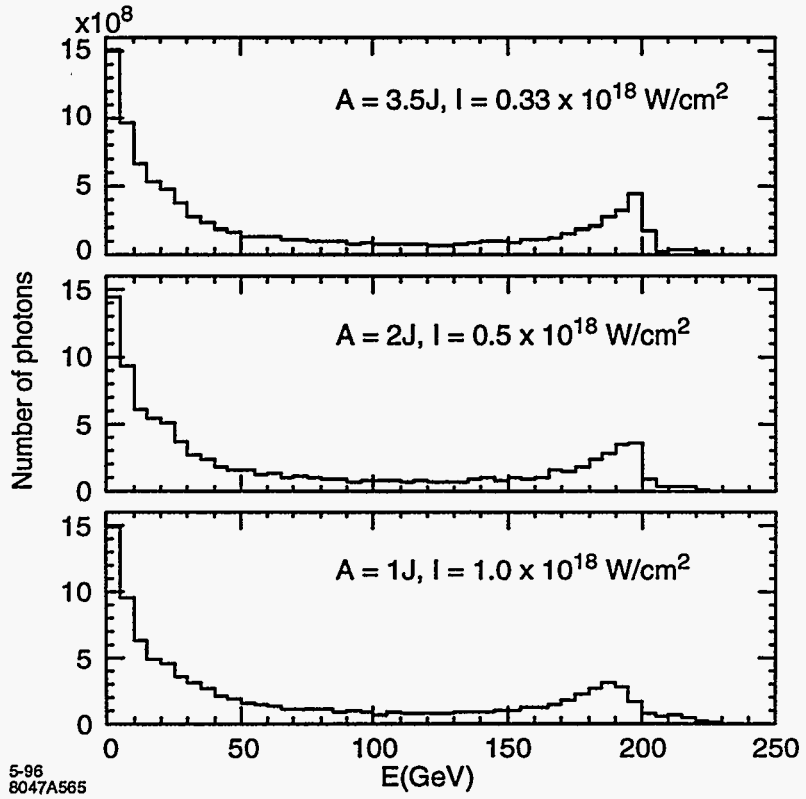


Figure B-8. *Spectrum of the γ-photons after the CP at different pulse energies and intensities.*

and γe^- luminosity are roughly equal to the geometric luminosity up to $\Delta y \approx 1\sigma_y$. The $\gamma\gamma$ luminosity at high energy end, $z > 0.65$, is about 10% of the geometric luminosity. A significant fraction of the total $\gamma\gamma$ luminosity is therefore in the low-energy region, and arises from the collisions of the beamstrahlung photons generated at the IP by the interaction of the spent electron beams. The low-energy $\gamma\gamma$ luminosity, as well as the γe^- luminosity, would pose a significant background problem in the collision scheme without the sweeping magnet.

From the table, it is apparent that the luminosity distributions are not a very sensitive function of the offset Δ_y. The $\gamma\gamma$ luminosity at high energy end ($z \geq 0.65$) is practically constant. The ee luminosity becomes smaller by a factor of 2 from $\Delta y = 0.25\sigma_y$ to $\Delta y = 1\sigma_y$. However the *ee* luminosity is already smaller than the γe^- luminosity by about a factor of five. Therefore the tolerance on Δy is rather relaxed; Δy up to about $1\sigma_y$ does not degrade the collision performance.

The column (g) in the table gives the result when a sweeping magnet is employed. Note that the high-energy $\gamma\gamma$ luminosity is about the same as before, but the background from the low-energy $\gamma\gamma$ or γe^- luminosities are significantly reduced.

The table also shows that the disruption angle for the low-energy particle ($E_{\min} = 3\,\text{GeV}$) is $\pm 8\,\text{mr}$ for collisions without a sweeping magnet. The low-energy particles are generated through multiple scattering at the CP. Analytical estimate shows that a ten-fold multiple scattering will generate of the order of one thousand particles with an energy of 2–3% of the initial electron energy (corresponding to about 5 GeV in

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
$L_{ee}/L_{\rm geom}$	0.2	0.19	0.15	0.16	0.12	0.096	0
$L_{ee}(z > 0.65)/L_{\rm geom}$	0.12	0.114	0.086	0.091	0.064	0.046	0
$L_{\gamma e^-}/L_{\rm geom}$	1.12	1.04	0.93	0.52	0.79	0.706	0.1
$L_{\gamma e^-}(z > 0.65)/L_{\rm geom}$	0.26	0.24	0.2	0.22	0.18	0.143	0.017
$L_{\gamma\gamma}/L_{\rm geom}$	1.23	1.22	1.16	0.38	1.08	1.05	0.37
$L_{\gamma\gamma}(z > 0.65)/L_{\rm geom}$	0.116	0.112	0.105	0.104	0.103	0.098	0.09
$L_{\gamma\gamma}(z > 0.75)/L_{\rm geom}$	0.057	0.0545	0.0514	0.051	0.05	0.046	0.051
$\theta_{y,\rm max}$ (mr)	8	8	8	8	8	8	2.5
$E_{\rm min}$	3	3	3	3	3	3	3

[a]No magnet deflection, $\Delta_y = 0.25\sigma_y$
[b] " " " , $\Delta_y = 0.5\sigma_y$
[c] " " " , $\Delta_y = 0.75\sigma_y$
[d] " " " , $\Delta_y = 0.75\sigma_y$ without beamstrahlung
[e] " " " , $\Delta_y = 1\sigma_y$
[f] " " " , $\Delta_y = 1.25\sigma_y$
[g]With magnet deflection, $b = 0.78\,\text{cm}$, $B = 10\,\text{kGauss}$

Table B-3. *Luminosities in $\gamma\gamma$ collision (V. Telnov).*

the present case). These are deflected to an angular cone of about 10 mr. This implies that the crossing angle should be larger than 20 mr. The exit beam pipe must be designed to accept these particles to avoid the background events.

For the case (g) with a sweeping magnet, the disruption of the 3-GeV particle is much smaller, about 3 mr.

Figures B-9–B-11 give a more detailed picture for the collision with the vertical offset $\Delta y = 0.75\sigma_y$. Figure B-9 gives the luminosity distributions for $\gamma\gamma$, γe^-, and ee collisions. Figure B-10 gives the $\gamma\gamma$ luminosities, where the solid curve is the same as in Figure B-9 and the dashed curve is for the case where the beamstrahlung contribution is suppressed. Figure B-11 gives the energy distribution of the final electrons.

Figure B-12 shows that the $\gamma\gamma$ luminosities for different offsets between $0.25\sigma_y \leq \Delta y \leq 1.25\sigma_y$ are more or less the same with each other. Figure B-13 shows the distribution in the vertical angle of the disrupted electron beams at various separations. The shape of the angular disruption may be used to calibrate the offset distance experimentally.

Finally, Figure B-14 shows the luminosity distributions for the case where a sweeping magnet is used (the case (g) in Table B-3). The γe^- luminosity is significantly reduced, and the ee luminosity is entirely suppressed.

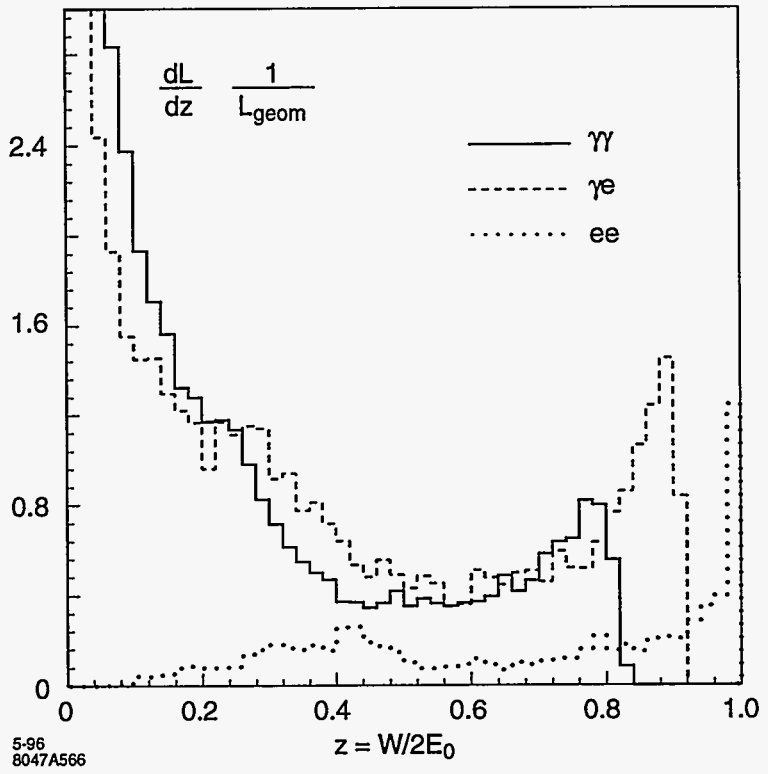


Figure B-9. *Luminosity distributions with $\Delta y = 0.75\sigma_y$ and no sweeping magnet.*

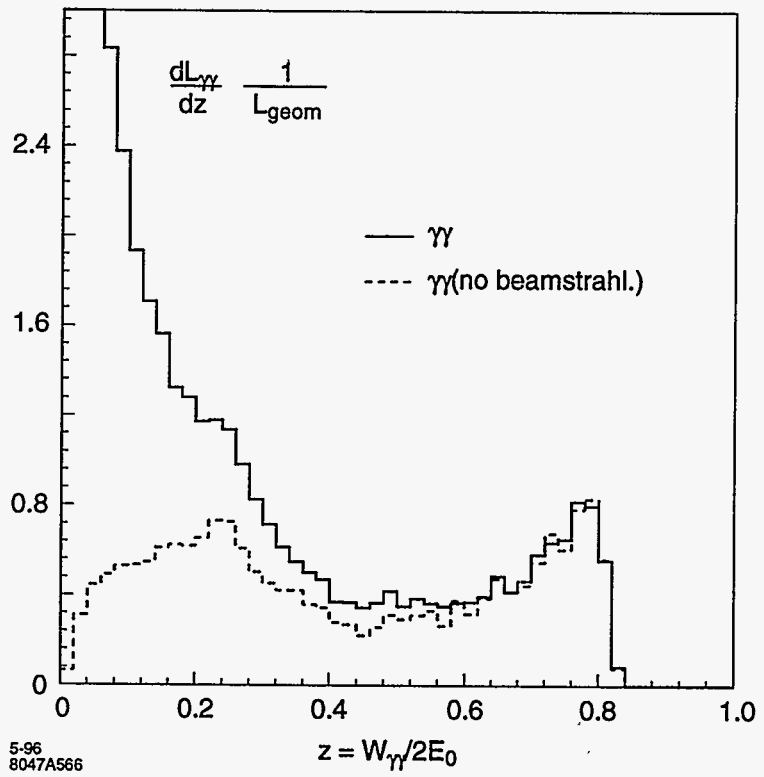


Figure B-10. *$\gamma\gamma$ luminosity for $\Delta y = 0.75\sigma_y$ and no sweeping magnet.*

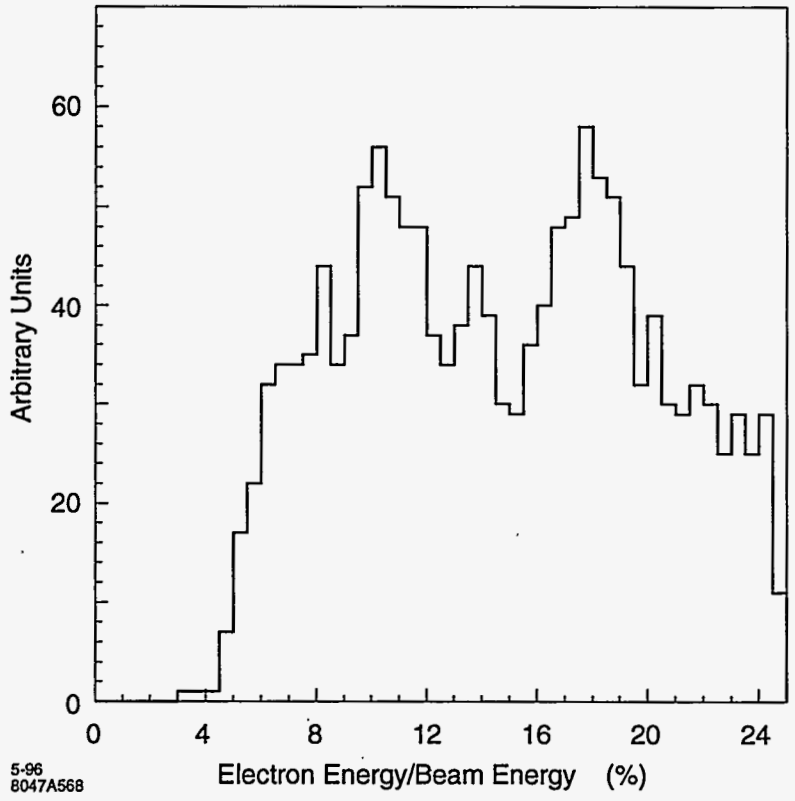


Figure B-11. *Energy distribution of final electrons for $\Delta y = 0.75\sigma_y$, no sweeping magnet.*

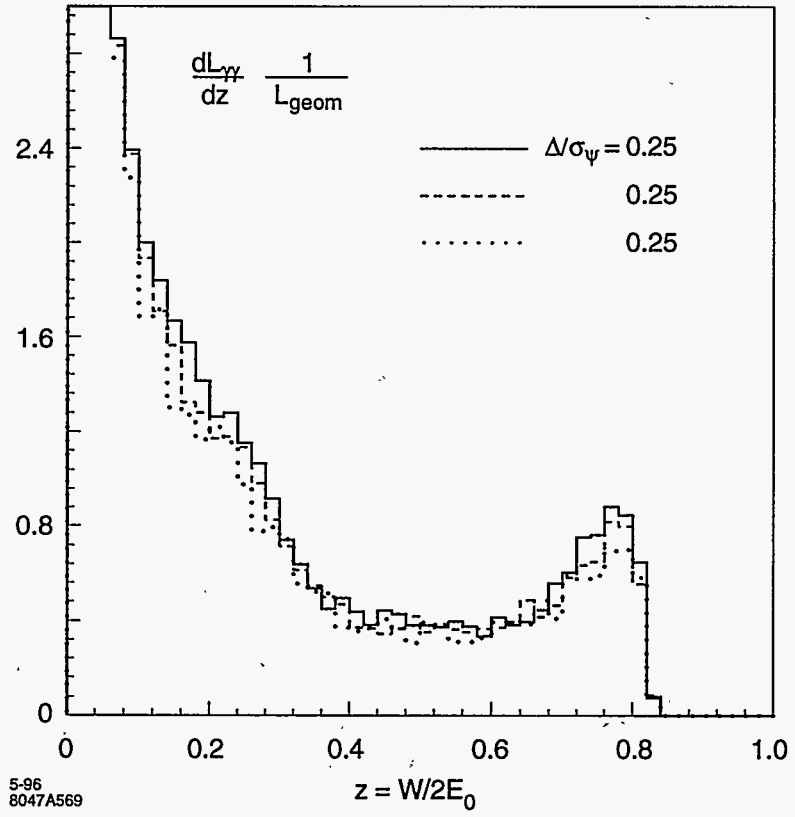


Figure B-12. *$\gamma\gamma$ luminosity distribution for various vertical separations, and no sweeping magnet.*

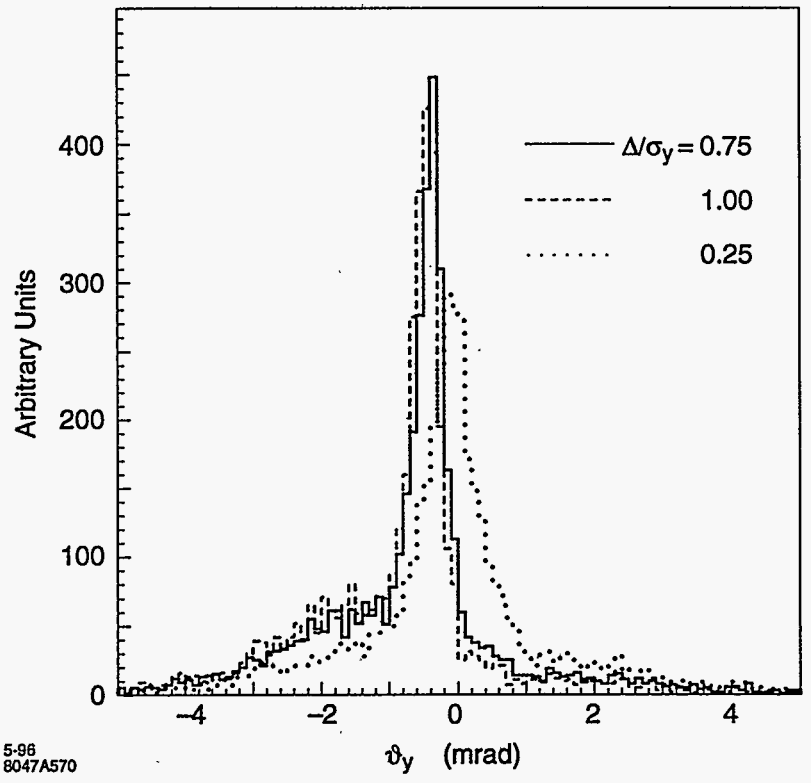


Figure B-13. *Distributions as a function of vertical angle for various vertical offset, and no sweeping magnet.*

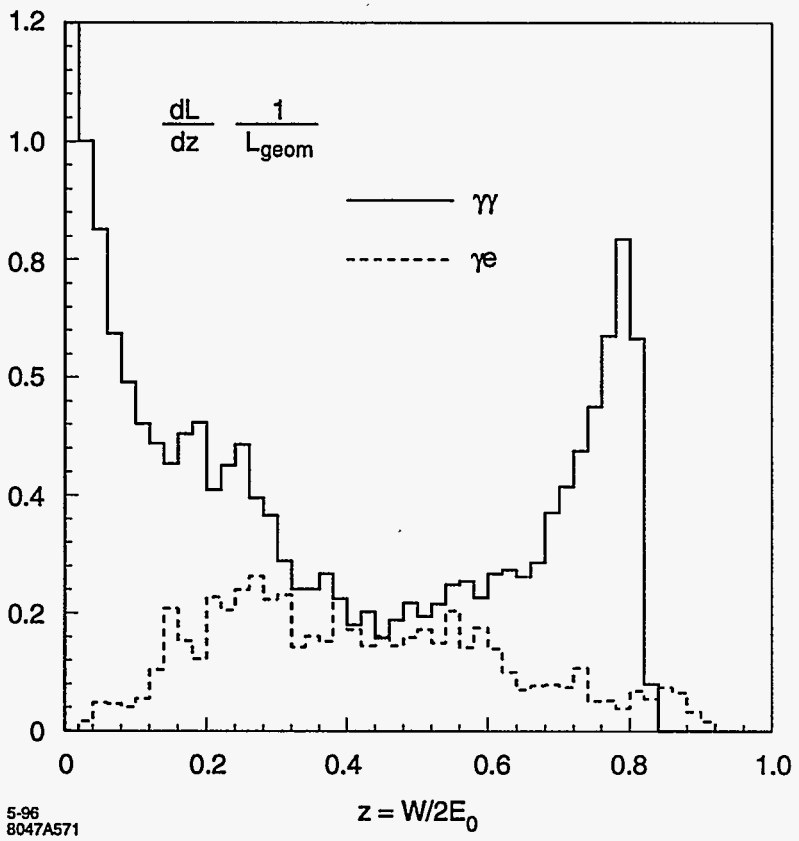


Figure B-14. *Luminosity distributions with magnetic deflection ($b = 0.78\,cm$, $B = 1\,T$).*

B.7 Backgrounds and Other Detector Considerations

B.7.1 Introduction

For the planned $\gamma\gamma$ physics program in IR2, the incoming polarized electron beams will be converted by an intense laser beam into hard photon beams. As a matter of routine operation in monitoring the incoming electron beam parameters, one or both of the beams will be left unconverted thus allowing high-luminosity studies of γe^- and e^-e^- collisions with the same detector. Given that the electron-photon CP will be located only a few millimeters from the IP, the resulting spent-electron beam background presents a significant challenge to the detector design for these interactions.

It is expected that the physics design criteria for the detector will be basically the same as those for the e^+e^- detector in IR1. Comprehensive, full solid-angle detectors are required to fully exploit the broad physics programs envisioned for both IRs, and the need for precision vertexing and excellent central tracking are identical for similar, specific physics reactions, such as Higgs production and detection in $\gamma\gamma \rightarrow H^0 \rightarrow b\bar{b}$ and $e^+e^- \rightarrow \nu\bar{\nu}H^0 \rightarrow \nu\bar{\nu}b\bar{b}$ through W fusion. The main differences expected in the IR2 design arise from the spent beam background and other detailed background differences, from the geometry of the laser optics and the luminosity monitoring systems, and from the unique physics opportunities offered by the different collision processes available.

In this section we detail the expected backgrounds for the $\gamma\gamma$ IR2 region and begin a discussion of detector design considerations. We provide a partial list of the unique physics requirements and a complete list of the machine-related backgrounds that are presently being considered. We review a set of detector geometry and performance parameters that are in use for the current level of simulation. Results of initial debugging runs of a detailed GEANT simulation and background analysis are also presented.

B.7.2 Physics Requirements

The Higgs $\gamma\gamma$ partial width measurement will be the central focus of the $\gamma\gamma$ physics program at a future linear collider. For a light mass Higgs, precision vertexing of b-quark jets will be essential in isolating a $H^0 \rightarrow b\bar{b}$ signal from charm and light quark backgrounds. For a larger mass Higgs, excellent electron and muon identification, and good jet energy measurements will be critical in identifying final state W bosons. In this case reasonable hermiticity would allow missing energy determination of neutrinos in the final state.

For detailed measurements of the photon structure function in γe^- collisions, good forward electron acceptance and energy measurements are required for accurate determination of q^2 and x. While in searches for singley produced supersymmetric (SUSY) particles, good electron and muon identification over the largest possible solid angle are important in testing various decay hypotheses.

In e^-e^- collisions, standard e^+e^- detector parameters are required in searches for new particles such as for a Z′. A comprehensive detector would be needed to explore fully any evidence for new physics, such as exotic doubley charged leptons or Higgses.

B.7.3 Backgrounds

Many of the background sources will be common to both the e^+e^- and $\gamma\gamma$ interaction regions. At this preliminary design stage, we rely heavily on the detailed background studies undertaken for the e^+e^- region as discussed in Chapter 12 of the main report.

The backgrounds from upstream sources such as the muon halo and quadrupole synchrotron radiation (QSR) backgrounds will be similar, but differing due to differences in beam parameters and collimation for the e^- beams needed for $\gamma\gamma$ collisions versus the very flat e^+e^- beams. Backgrounds from nearby sources, such as beam-gas scattering, should also be the same.

In the $\gamma\gamma$ interaction region, backgrounds due to the spent electron beams from the e^--γ conversion need particular attention. After the conversion, these charged beams are naturally deflected at the final focus by the long-range Coulomb interaction. This deflection will increase the angular divergence of the beams and also create beamstrahlung photons which will lead to additional backgrounds. As discussed in Section B.6, detailed simulations of the conversion process and interaction physics are being developed. A simple parameterization of this background is discussed below.

High-energy electron-positron pairs will be created due to nonlinear effects in the interaction of laser photons with high-energy photons at the conversion points. At the interaction point, beamstrahlung photons from the disrupted electron beams will interact with themselves, with the Weizsäcker-Williams virtual photons of the opposing e^- beam and with the oncoming high-energy photons to produce additional low-energy pairs. A display and simulation of the effect of these pairs are presented in a following section.

Physics backgrounds arise from the suppressed-luminosity collisions of the spent electron beams with each other and with high-energy photons, and from hadronic backgrounds from low-energy $\gamma\gamma$ interactions. Both backgrounds will add to the inherent backgrounds associated with any particular physics study. However, the backgrounds for $\gamma\gamma$ physics due to e^-e^- and γe^- collisions will provide in themselves additional physics opportunities, while the low-energy $\gamma\gamma$ interactions provide parasitic physics similar to what has been available at lower energy machines. More detailed simulation of the hadronic backgrounds from $\gamma\gamma$ interactions and studies of the background e^-e^- collisions are being planned.

B.7.4 Detector Considerations

For a comprehensive study of $\gamma\gamma$, γe^- and e^-e^- physics, the detector chosen for IR2 will be expected to provide precision vertexing for b-quark separation, and accurate momentum and energy measurements of electrons, muons, and jets up to the beam energy over its full solid-angle coverage. Particle identification of electrons and muons will be accomplished by the calorimeter and muon tracking systems, while a central tracker with dE/dx and other detection techniques would extend the electron-hadron separation over a larger solid angle. Excellent pattern recognition and fine segmentation is required to minimize the sensitivity to machine backgrounds.

Over a large rapidity (η) range, special consideration has to be given to monitoring both the total luminosity and the differential spin-dependent luminosity, $dL_J/d\sqrt{s}/d\eta$, for each of the processes, e^-e^- , γe^- , and $\gamma\gamma$. Detector issues arise from possible interference with the laser optics in the small angle region, and from the lower rates for particular interactions at larger angles which may allow only offline luminosity measurements. Small-angle Møller scattering would provide an excellent monitor of the colliding beams in the e^-e^- mode of operation. Placement of the luminosity detectors behind the laser optics should not limit online monitoring,

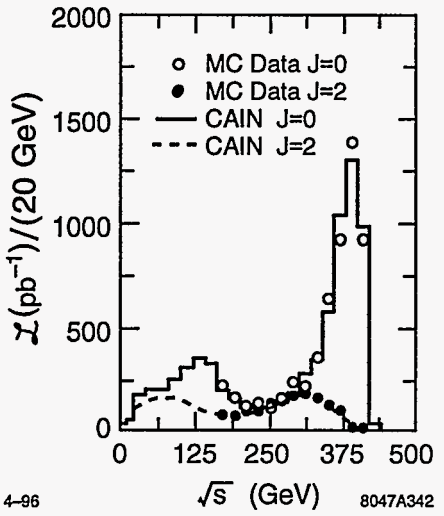


Figure B-15. *Comparison of measured and generated luminosity. The circles and asterisks are measured luminosity for $J = 0$ and $J = 2$ component while solid and dashed lines are generated luminosity for $J = 0$ and $J = 2$ respectively. The normalization corresponds to total integrated luminosity of $10 fb^{-1}$.*

or precision offline measurements at the few per cent level. The two-photon processes, $\gamma\gamma \to e^+e^-$ and $\gamma\gamma \to \mu^+\mu^-$, allow online monitoring at small angles and precision measurements at intermediate angles for both the $\gamma\gamma$ and γe^- processes. The process of W pair production in two-photon reactions, $\gamma\gamma \to W^+W^-$, also provides an excellent measure of the $\gamma\gamma$ luminosity.

For a typical $\gamma\gamma$ collider luminosity distribution, the total weighted cross section for W pair production of about 50 pb and a selection efficiency of 15% for 4-jet reconstruction of the Ws yields 75K events per 10 fb^{-1} of integrated luminosity [Takahashi 1995]. For the spin-2 process $\gamma\gamma \to e^+e^-$ the corresponding total luminosity-weighted cross section with $\sqrt{s_{\gamma\gamma}} > 200\,\text{GeV}$ of 10 pb, and a detection efficiency of 70% yields 70 K events per 10 fb^{-1} [Takahashi 1995]. The spin-0 luminosity would be measured by flipping the electron and laser beam polarization simultaneously [Telnov 1993]. Figure B-15 compares the simulated $J = 0$ and $J = 2$ luminosity measurements with the generated luminosity spectra. The expected statistical errors in the W and lepton pair luminosity measurements are shown in Figure B-16.

In detailed GEANT Monte Carlo simulations of a generic detector, we have shared the basic geometry definitions for studies of both the e^+e^- and $\gamma\gamma$ interaction regions. The detector is taken to be 2 m in radius and ± 2.5 m along the beam with a 2-T solenoidal magnet field. At this stage, it consists of simple models of beam line elements and scoring planes only in the vertex and central tracking detector regions. The eventual calorimetry and muon tracking systems are not included, and the resulting detector self-shielding is not taken into account.

For the IR2 design studies, we increase the crossing angle from 20 to 30 mr as presently chosen for the laser optics design, and increase the acceptance of the outgoing quadrupoles from about ~3 mr to ~10 mr to transport the spent electron beam outside of the interaction region.

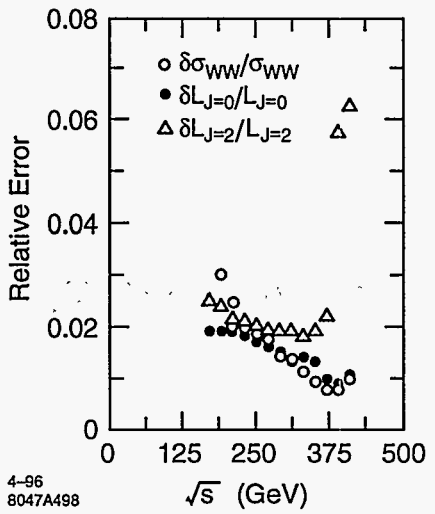


Figure B-16. *The statistical error in W pair production cross-section and luminosity measurement with integrated luminosity of* $10 fb^{-1}$. *The circles are for W measurement while the asterisks and stars are for luminosity measurement for* $J = 0$ *and* $J = 2$ *component.*

B.7.5 Initial Simulations

A GEANT drawing of the detector with an expanded vertical scale is shown in Figure B-17. At the center of the drawing are the scoring planes used in the vertex detector region from 2–10 cm in radius and ±20 cm in length. On either side are the final-focus (FF) quadrupole magnets and beam line masking chosen to minimize backscattering into the detector. For display purposes only, we plot 100 low-energy electron and positron tracks from a simulation of the beam-beam interaction in e^+e^- collisions. (We expect that the electron-positron pairs produced in the conversion and interaction points of the $\gamma\gamma$ region will be similar in energy and angular distributions.) The pairs were generated only in the forward direction of one of the beams. These tracks radiate photons as they spiral along the strong magnetic field lines to the face of the inboard FF quadrupole magnets. Secondary backgrounds are produced from electrons and photons that backscatter into the detector. Figure B-18 displays the hits in the central tracking chamber due to this background source for the $\simeq 10^4$ pair tracks expected to be produced at each bunch crossing. Earlier simulations of a conventional e^+e^- detector with an approximate crossing angle geometry indicated that this background should be tolerable [Ronan 1993].

To obtain an initial estimate of the spent electron beam backgrounds, we generate in our simulation electrons originating from the interaction point with a flat energy spectrum from 15 to 85% of the incoming beam energy. The generated angular spread of these electrons is parameterized by two Gaussian fits to independent detailed simulations of the conversion process. The fitted distributions in both transverse dimensions have a central component and a broader tail with angular spreads of ~1 mr and ~3–4 mr, respectively. On a highly expanded vertical scale, Figure B-19 displays 100 such spent electron tracks exiting through the downstream quadrupole magnets. More detailed simulations of the conversion process and tracking of the spent electrons through the interaction region are in progress. Also, realistic modeling of the fields within the quadrupole magnets are needed in the GEANT simulations to take account of the spent beam as it is transported out of the interaction region.

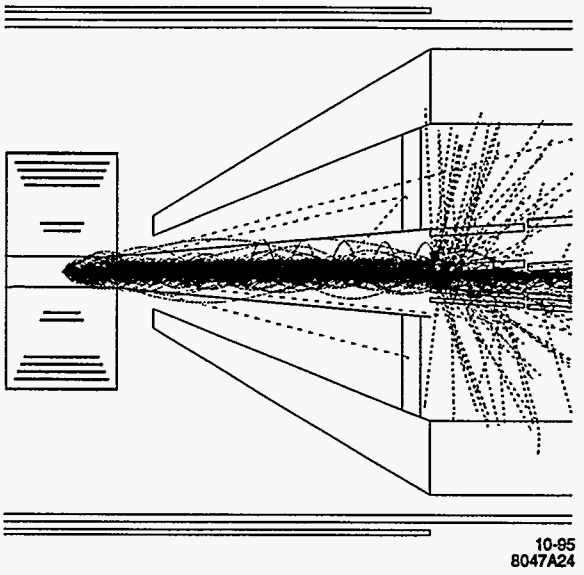


Figure B-17. *Low-energy electron-positron pair background tracks hitting the the face of downstream quadrupole magnets. Only 100 such tracks are shown.*

B.7.6 Conclusions

We are just beginning to understand the relevant design issues and parameters for a detector which could perform the physics anticipated at a future electron linear collider. We welcome volunteer help and encourage international collaboration in the development of simulation tools and eventually in the conceptual design of a detector for $\gamma\gamma$, γe^- and e^-e^- physics.

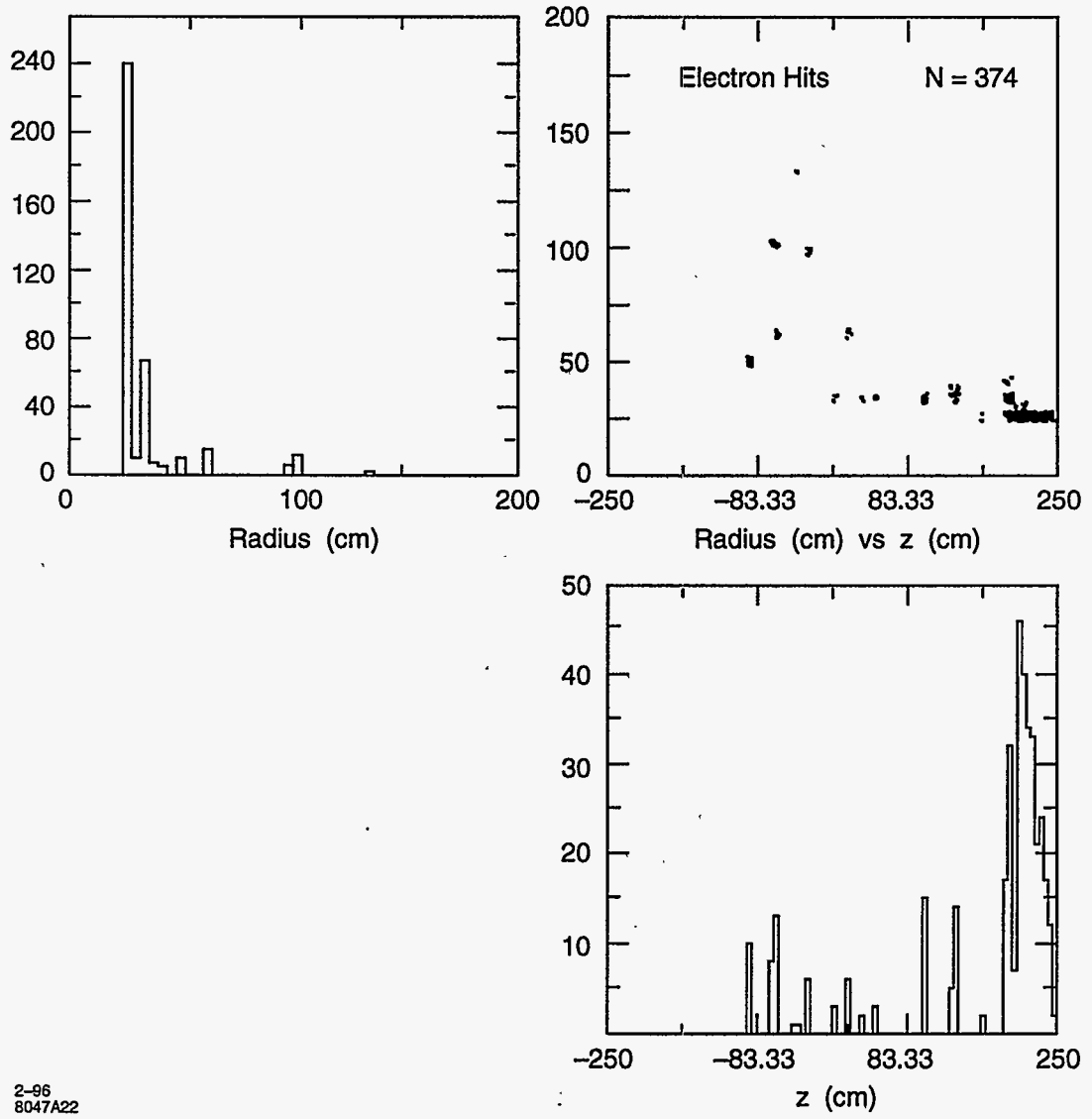


Figure B-18. *Scatter plot (2b), and radial and axial projections (2a,2c) of secondary background hits in central tracker from pair background.*

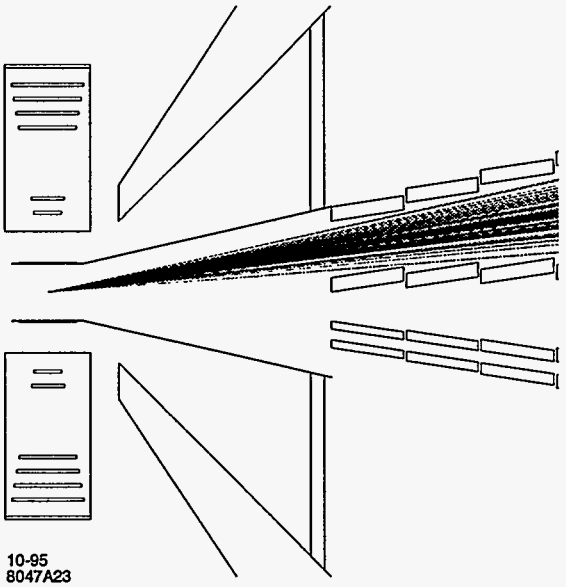


Figure B-19. *Spent electron beam exiting through downstream beam-line magnets. Only 100 rays are shown.*

B.8 Laser Optical Path in IR

B.8.1 Single-Pass Scheme

This section considers some of the practical issues in actually bringing the laser beams into (and out of) an interaction point centered in a HEP detector of the sort that is described in the previous section. A solution is shown in Figure B-20. The case considered here is the simplest one. Each electron beam has an opposing laser beam which is backscattered from it immediately prior to the interaction point. The laser beam, which is essentially unaltered in the process, is then disposed of. The beams are brought in from opposite sides and follow very similar (but not identical) paths along opposite directions. Before discussing the scheme in more detail, it is worthwhile to explain some of the constraints that are encountered in this problem.

- Because of the extremely high intensities involved, transmissive optics are, for the most part, not feasible. This occurs because in the final analysis, the index of refraction of all optical materials is a function of the intensity of the laser light in it. These nonlinearities give rise to a filamentation instability if the total length of material in the system is sufficiently large. The limitation from this so called "B-integral" problem [Siegman 1986] means that such devices as lenses, Pockels cells, and polarizers may not be used after the gratings in the compressor. In fact, the limitations on the design of the optical path produced by this problem are probably such as to just allow a pair of quartz windows for the beam to enter and leave the vacuum system of the accelerator. Beyond these two windows, all optical elements will presumably be reflective; *i.e.*, dielectric mirrors. This limitation also implies that the laser beam will be transported in vacuum.

- Focusing of the beam must be optimized to produce the desired peak intensity (limited by nonlinear effects in the backscattering process as discussed above) and a sufficiently long length of such intensity that most of the electron bunch is converted into photons. This optimization was discussed above. The practical consequence of this is that the "$f_{\#}$" of the laser focusing will be fixed by this optimization. The $f_{\#}$ is defined roughly as the ratio of the focal length to the diameter of the focusing mirror, as illustrated in Figure B-21, assuming that the mirror is uniformly illuminated. For Gaussian beams, the diameter of the focusing mirror is somewhat arbitrary, but we will take it as the intensity $1/e^2$ diameter. With this definition, one can show that $f_{\#} = 1/4\sigma_{x'}$, where $\sigma_{x'}$ is the rms angular divergence of the focussed laser beam. In the current design, we have $\sigma_{x'} \approx 27.9\,\mathrm{mr}$ and $f_{\#} = 8.65$.

- Effective use of the laser requires that the optical axis of the laser beam be parallel with the direction of the electron beam to within an angle small compared to the aspect ratio of the laser bunch. The degradation in the general case is given by [Xie 1995] (We do not consider the possibility of "crabbing" the laser beam here.)

$$\frac{L(\theta)}{L(0)} = \frac{1}{\sqrt{1+(\theta/\theta_0)^2}}, \quad \text{where} \quad \theta_0 = \sqrt{\frac{\sigma_{x1}^2+\sigma_{x2}^2}{\sigma_{z1}^2+\sigma_{z2}^2}}\ , \tag{B.26}$$

  The σ_xs are the transverse sizes of the two crossing beams and the σ_zs are the longitudinal sizes. In our case, the expression for θ_0 is dominated by the size of the laser beam and reduces to $\theta_0 = \sigma_x/\sigma_z$. From the discussion above, one can infer that the aspect ratio of the laser is approximately $4f_{\#}$. Since the angle between the optical axis and the electron beam is $1/2f_{\#}$, the minimum degradation if the disrupted beam does not pass through the mirror is $1/\sqrt{5}$. Because of this degradation, a through-the-lens design has been implemented.

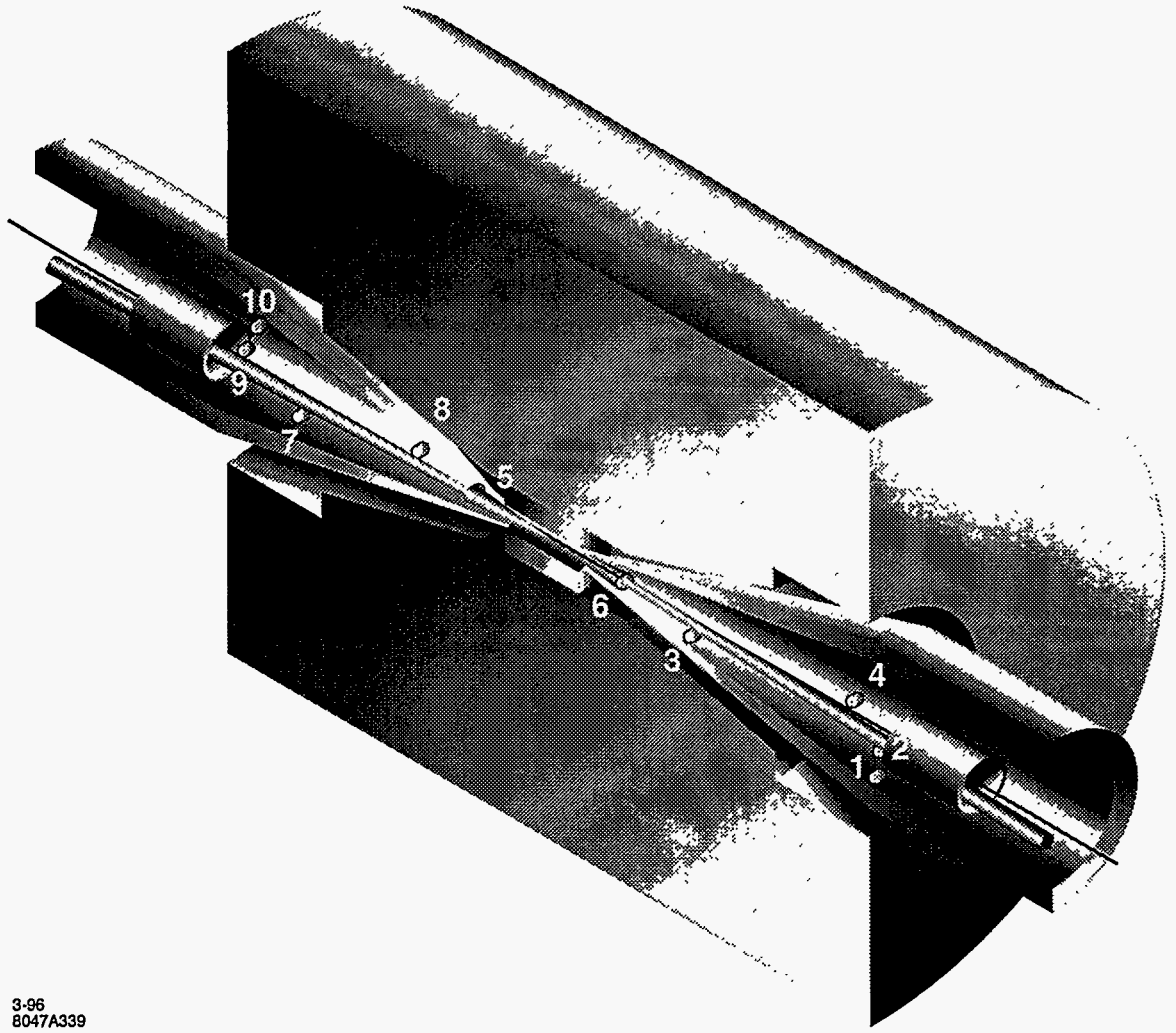


Figure B-20. *The layout of laser optical path in the IR2. The paths followed by the various beams are described in the text. The mirrors have an elliptical shape due to the overlap of the two circular beams. The shapes shown for the mirrors represent the outer edge of the flattop beams. While no provisions are shown for the edges of the mirrors or for mirror supports, space appears to be available.*

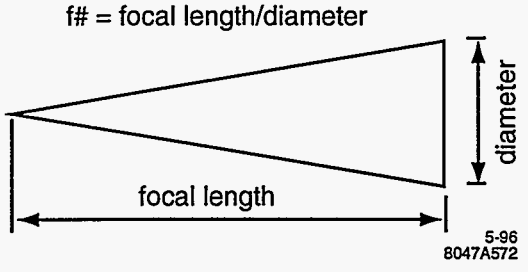


Figure B-21. *The definition of the $f_{\#}$ of an optical system.*

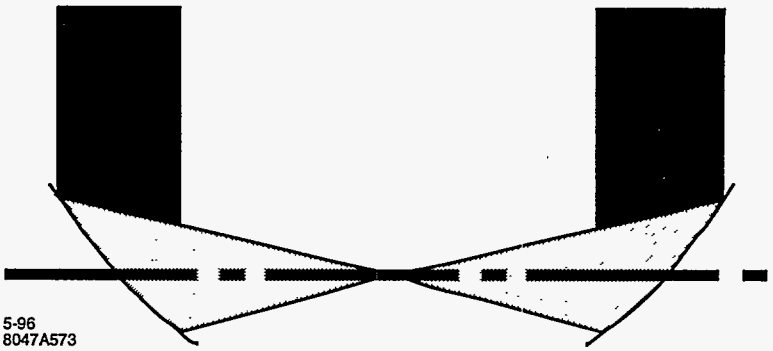


Figure B-22. *The off-axis paraboloid is both difficult to align and sends the beam path through space better occupied by the detector.*

- Given both the high peak and average power involved, it is almost certainly necessary to transport the "spent" laser beam out of the IR to an external dump. This also allows external monitoring of the optical quality of the beam after it has gone through focus.
- The two electron beams cross at only a small angle. This angle is limited by the creation of synchrotron radiation in the B-field of the detector. It must be greater than zero in order that the disrupted beam from the opposite side has a separate path by which to exit the detector. Since this angle, which in the current design is 30 mr, is small compared to the half angle associated with the focusing of the laser beam ($1/2f_{\#} \approx 57.8$ mr), it follows that the "used" beam from one side will land on the focusing optic for the other side.
- The two conversion points are separated (1 cm) in space so that softer photons (which are produced at larger angles) will diverge before the interaction point and not interact. This will introduce a "walk" into the paths of the two laser beams so that at a large distance from the IP the beams eventually separate.
- If the two beams share the same mirrors at some point (and this seems inevitable), then the mirrors probably should not be located at the points in space where the two pulses (one from each direction) overlap in time as this will significantly increase the peak intensity on the mirror.
- At first glance, there might seem to be advantages to bringing in the laser beam using an off-axis paraboloid as shown in Figure B-22. Such schemes suffer, however, both from difficulties in alignment and from a need to direct the beam through a region which is naturally occupied by the detector.
- Damage to the dielectric coatings on both the flat and curved mirrors used to transport the laser is an issue. Ultimately, this sets a limit on how close the closest optic can come to the interaction point.

The design for the laser-optical path is shown in Figure B-20. The central cylinder represents a vertex chamber. The more-or-less cylindrical object outside of that represents the rest of a generic cylindrical detector. A previous proposal [Miller 1995] in which the laser beam focuses twice has been implemented. This makes it possible to maintain near-normal incidence on all the optics in the system, and keep the laser beams in the vicinity of the conversion points inside a cylinder which is roughly defined by the outer radius of the final focusing optic used. The incident electron beams are shown as thin lines on the far side of the detector. The disrupted electron beams are shown as cones slightly offset from the incident beams. The axis of the detector is located between the incoming and outgoing electron beams.

As was mentioned above, the two laser beams enter from opposite sides of the detector and trace similar, but not identical, paths in opposite directions. These paths are distinct because the incoming electron beams make a small angle with each other, and because the conversion points are offset from each other by 1 cm.

The path followed by the laser beam incident from the right in Figure B-20 is as follows. A round and collimated (*i.e.*, parallel) beam incident upon the right (far) side of mirror 1 is reflected vertically onto mirror 2. It is, in turn, reflected in a more-or-less horizontal direction onto the right side of mirror 3. This mirror reflects the beam onto the right side of mirror 4. At this point the beam is still parallel. Mirror 4 focuses the beam to a point over the center of mirror 6. This is done to minimize the clear region which must be provided as the beam is brought onto mirror 5 (the final focusing mirror for this side). After passing through this focus, the beam diverges and lands on the left side of mirror 5. Mirror 5 focuses the beam to a diffraction limited spot 0.5 cm past the interaction point where it backscatters from the electron beam which entered the detector from the same side.

At this point, the only task left is to extract the beam from the inside of the detector. Mirrors 6–10 are used to bring in the second laser beam to convert the second electron beam. Because of the very small angle between the two electron beams, and because the final focusing optic has its axis coincident with the electron beam, the "spent" beam will follow a path out of the IR which is nearly coincident with that of the other incoming beam. These paths are not exactly coincident because of the 20 mR angular offset of the two electron beams. This offset of laser beams is what produces the requirement for oval (as opposed to circular) mirrors.

After passing through the conversion point, the original laser beam diverges and strikes mirror 6. Mirror 6 will refocus the beam to a point just under mirror 5, from which it will diverge and land on mirror 7. Mirror 7 will recollimate the beam and send it onto mirror 8. After leaving mirror 7 the beam is once again parallel. It is then reflected from mirrors 9 and 10 in a manner similar to that by which is entered on the opposite side.

It must be noted that because mirrors 6–10 will be optimized to produce a diffraction-limited spot for the beam entering on that side, and because the beam exits at a slightly different angle, the optical path on exit will introduce aberrations into the beam. Whether or not these aberrations can be corrected subsequently so that the beam may be diagnosed after exiting the detector is an open question.

The mirrors in this scheme have been located where one finds the luminosity monitor in a conventional e^+e^- detector. Since there does not appear to be an analog of Bhabha scattering in $\gamma\gamma$ collisions, this is presumably not a major difficulty. If this region needs to be instrumented (to ensure, for instance, maximum hermiticity of the detector), then the material associated with the mirror is probably tolerable. A typical mirror will have a thickness of $\frac{1}{6}$ of its diameter. This will produce 10% of a radiation length of fused quartz over some fraction of the azimuth.

B.8.2 Optical Beam Focusing

Previous work [Meyerhofer 1995a] has sometimes assumed that Gaussian laser beams would be used at a $\gamma\gamma$ collider. This is not optimum for a couple of reasons. High-power lasers such as will be needed for this project are typically build with flattop beams to achieve the maximum fill factor. The idea is roughly the following. The cost of the laser depends on both the aperture and the peak intensity within that aperture. Having paid for both of these, the maximum energy is extracted by uniformly filling the aperture; *i.e.*, maximum fill factor. Since, as is explained below, it is probable that the disrupted electron beam will pass through the middle of the final laser focusing optic, a beam profile which peaks in the center is not optimum. Figure B-21

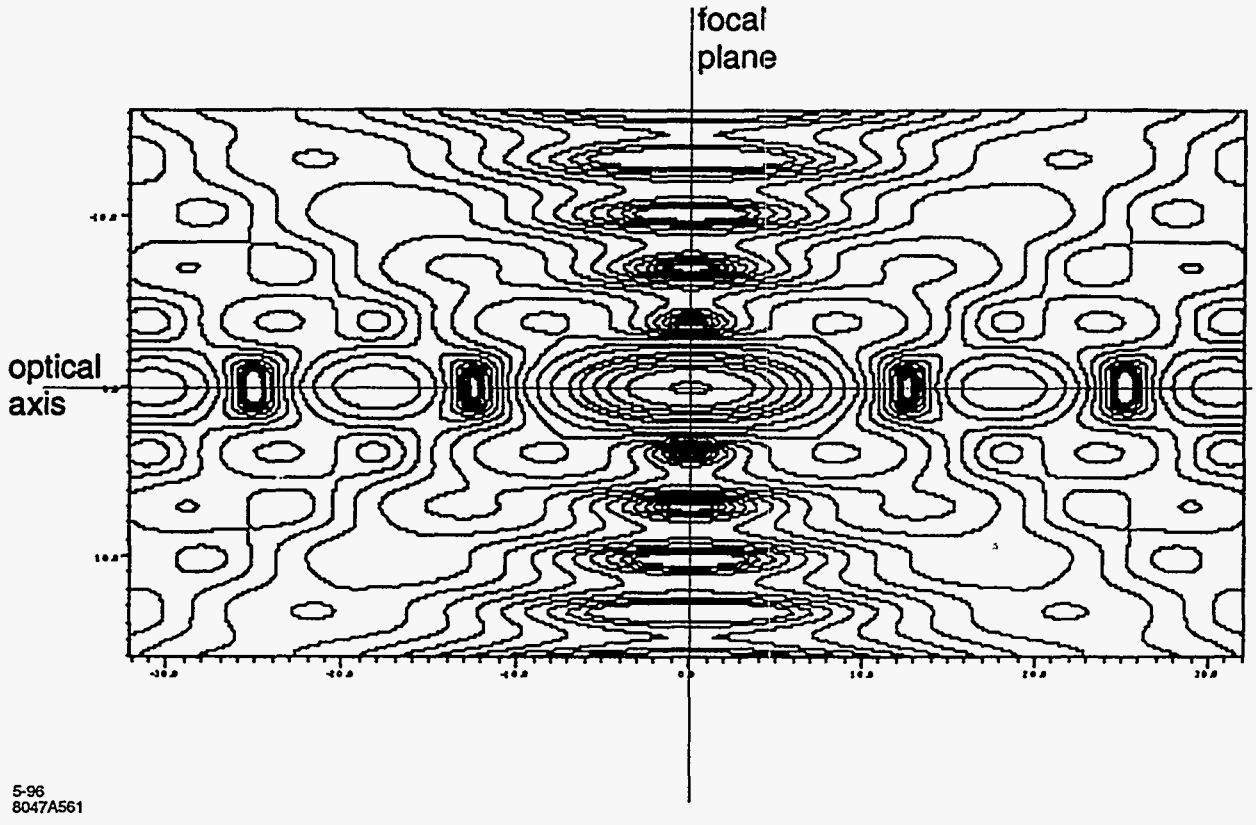


Figure B-23. *The behavior near best focus from a flattop beam.*

shows the definition of the $f_{\#}$ of a focused beam. The resulting properties of the beam at the best focus are given in Table B-4. It is seen that while the Gaussian profile provides simple analytical expressions near best focus, the flattop beam produces a much more complicated pattern as shown in Figure B-23.

Because the optimization has been done for Gaussian beams, a comparison has been made between the focal spots for the two cases. This is shown in Figure B-24. A flattop beam requires a slightly larger beam (smaller $f_{\#}$) to produce a similar spot. The exit hole required to let the disrupted beam pass through produces only a 2% energy loss in this case. Table B-5 gives a summary of the laser and electron beam parameters. Table B-6 is a summary of the parameters relevant to the laser optics design.

B.8.3 Laser Damage of Optics

The particular optics used in this design are not expensive. They are neither extremely large, nor are the surfaces expected to be particularly complicated. On the other hand, their reliability will remain a critical issue if a $\gamma\gamma$ collider is to operate successfully. Definitive statements on this subject can not be made at this time because no data exists on the damage threshold for multilayer surfaces for the particular time structure of the laser pulse which would be required in this case. Data exists for the case of single picosecond scale

	Flattop	Gaussian
$I(r, z=0)$	$I_0 \left[\frac{2J_1(r/s)}{r/s}\right]^2$	$I_0 \exp\left(-2\frac{r^2}{w^2(z)}\right)$
	$s = \frac{f_\# \lambda}{\pi}$	$w(z) = w_0 \left[1 + \left(\frac{z}{z_R}\right)^2\right]^{1/2}$
$I(r=0, z)$	$I_0 \left[\frac{sin(z/z_0)}{z/z_0}\right]^2$	$I_0 \left[1 + \left(\frac{z}{z_R}\right)^2\right]^{-1}$
	$z_0 = \frac{8}{\pi} f_\#^2 \lambda$	$z_R = \frac{4}{\pi} f_\#^2 \lambda$
Effective area	$\frac{f_\#^2 \lambda^2}{\pi}$	$\frac{\pi w_0^2}{2} = \frac{2 f^2 \lambda^2}{\pi}$

Table B-4. *The beam parameters at focus for Gaussian and flattop beams.*

Electron beams:	
μ-pulse / macro-pulse	90
repetition rate	180 Hz
angular offset	±15 mR
spent beam divergence	±10 mR
Laser beams:	
wavelength, λ	1.05 μm
beam profile	5-cm diam. flattop
energy	1 J / pulse
length	1.8 psec
μ -bunches	90 at 1.4 ns
repetition rate	180 Hz
power (per beam)	16 kW

Table B-5. *Beam parameters used in this study.*

Optics:	
$f_\#$ (flattop)	7 (±71 mr)
Distance to first mirror	35 cm
Area of first mirror	20 cm^2
Fractional area of first hole	2%
Fluence	0.05 J/ cm^2/*pulse*
Vertex chamber:	
Inner radius/solid angle	4 cm/ 0.986
Outer radius/solid angle	12 cm/ 0.894
Total length	48 cm
Distance to first quad.:	200 cm
Masking:	100–150 mr

Table B-6. *Parameters of the design proposed in this study.*

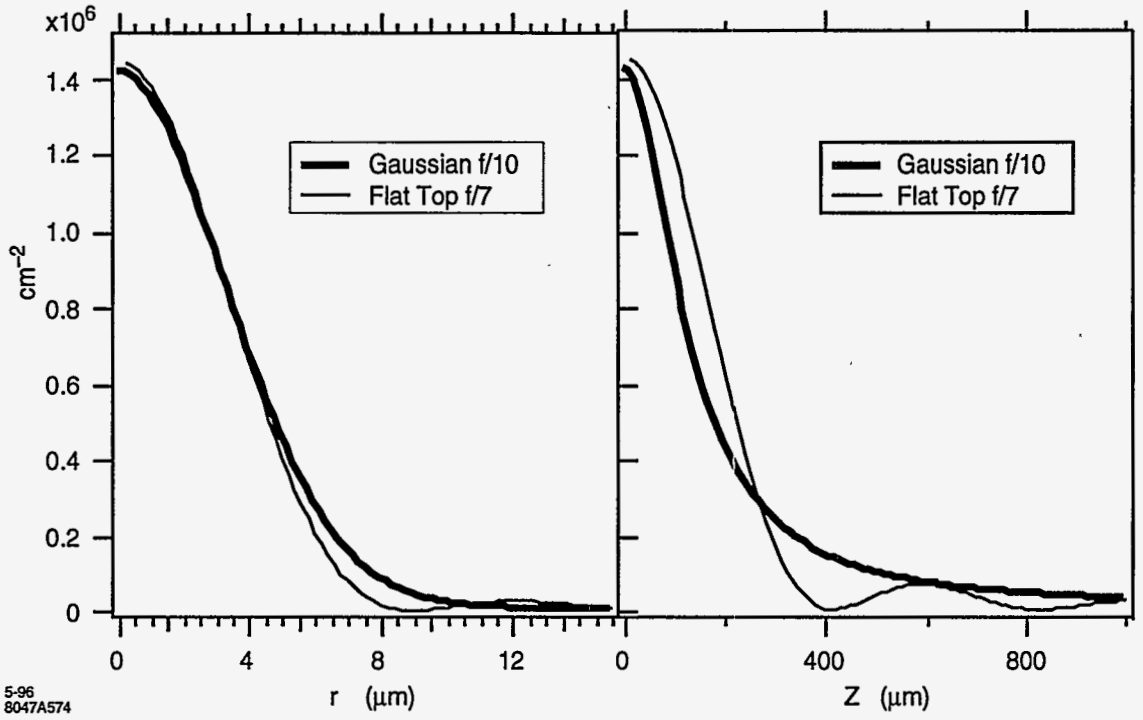


Figure B-24. *A comparison of the radial and axial dependence of the intensity for Gaussian and flattop beams. Similar sizes are obtained when the f/# of the flattop beam is about 0.7 times that of the Gaussian beam.*

pulses. Extensive data exists for much longer length pulses (nanosecond time scales), and for collections of nanosecond scale pulses which produce very high average powers. The three limits are considered below.

- On the time scale of a single pulse (1.8 ps), measurements have recently been made at LLNL on commercial multilayer surfaces [Stuart 1995a]. Damage thresholds in the range of 0.7 to 2 J/cm^2 have been observed. This is more than a factor of four above the anticipated fluence of 0.05 J/cm^2.
- On nanosecond time scales, the situation is less clear. The anticipated fluence is 4.5 J/cm^2 spread over 126 ns. The damage threshold for continuous pulses of this length is 100 to 200 J/cm^2. While this is well above the anticipated operating point, the validity of this averaging has not been demonstrated. No data exists for collections of picosecond pulses separated by a few nanoseconds and which accumulate to fluences of this magnitude.
- The average flux on the final optic is 0.83 kW/cm^2. Since this is below the levels of 3 to 5 kW/cm^2 which are routinely used in the laser isotope separation program (AVLIS) at LLNL, it does not appear to be a problem.

It is worth noting that the overall scale of the optics, masking, and vertex chamber is set by the inner radius of the vertex chamber and by the damage threshold of the first optical element. If fluences higher than the design fluence are possible (and this certainly cannot be ruled out at this point), then all dimensions can be scaled down if it is desired to place the vertex chamber at a smaller radius.

At this point there is a clear need for a detailed optical design of the elements involved here. This would determine the particular surfaces required on the focusing elements to verify that a defraction-limited spot can be produced. It would also make it possible to determine the magnitude of the aberration introduced on the exit path and to determine if it is possible to correct these aberrations externally. Such corrections would be needed to use the transmitted beam to verify the quality of the focal spot.

There are also questions about the damage thresholds of the optics as well as questions about the optics resistance to damage by radiation.

The optical design here has been based on Table B-1, with A =1 J and $I = 1 \times 10^{18}$ W/cm^2. As discussed elsewhere, the non-linear effect in Compton conversion is not negligible. the non-linearity can be avoided by increasing A to 3 J and decreasing I by a factor of three. This will increase the requirement on the laser power, as well as power on the mirror by a factor of 3. The focusing optics also changes from f/7 to f/10, with the result that the mirrors are located a little further out from the IP.

B.8.4 Ideas on Relaxing Average Laser Power Requirement Via Multipass Optics

In a gamma-gamma collider, a high-energy electron bunch is converted into a burst of gamma rays by Compton scattering with an intense laser pulse. Assuming the laser source has the same pulse structure and repetition rate as that of the electron beam and assuming that each laser pulse is used only once, the required average laser power would be around 20 kW, three orders of magnitude higher than what has been achieved with any laser having TW peak power. However, the required average power can be significantly reduced if the laser pulses can be reused, because the laser pulse suffers little loss in energy after each scattering. In this section, we explore the possibilities of reusing the laser pulses with specially designed optics.

There are two approaches to this problem. The first one is a multipass approach in which a laser pulse is made to pass through the conversion point a finite number of times before being thrown away. In this way, the average power is reduced by reducing the number of pulses needed. The second approach may be called pulse stacking in which a train of weaker laser pulses are stacked up in an optical cavity to make a stronger pulse for intracavity conversion. In designing optics for both approaches, using transmissive optical elements should be avoided if possible to minimize the nonlinear transverse and longitudinal pulse distortion at high power.

To illustrate the idea, an example of multipass optics based on all reflective elements is schematically shown in Figure B-25, in which a laser pulse is made to pass through a CP eight times, four in each direction. This is done by using two mirror banks each with eight individual mirrors. Each individual mirror is placed in the numbered order along the optical pass according to the sequence the laser pulse is kicked. Such a kicking pattern guarantees the same pass length for each round trip the laser pulse makes passing through a CP. With some modification, the scheme in Figure B-25 can also be made to accommodate two conversion points separated by a few mm while requiring laser pulses to pass through the two conversion points in opposite direction.

The reduction factor in the required laser average power for the multipass approach is limited because the number of mirrors that can be utilized is limited due to the tight space near the interaction region, and due to the build-up of aberration. To make better use of the limited space, it is desirable to have laser pulses bounced back and forth along the same pass, thus forming a cavity mode. By stacking up weaker externally injected pulses inside a cavity, the factor of reduction in average power could be up to the cavity Q.

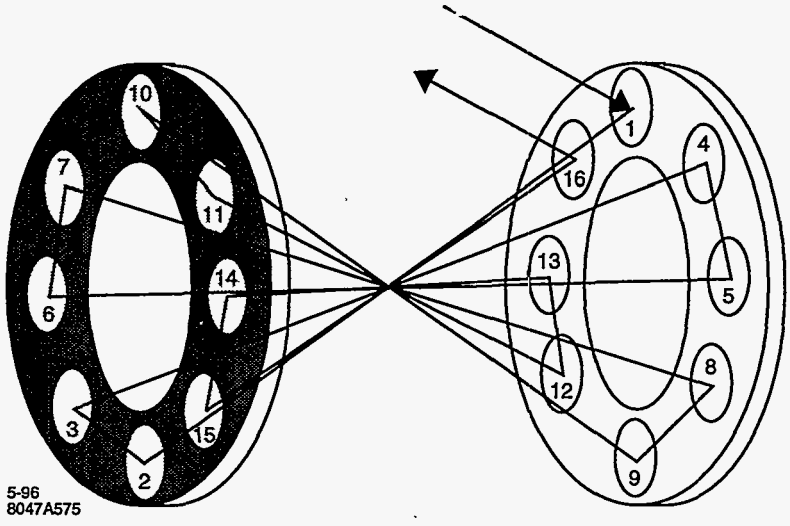


Figure B-25. *Multipass optics with reflecting mirrors.*

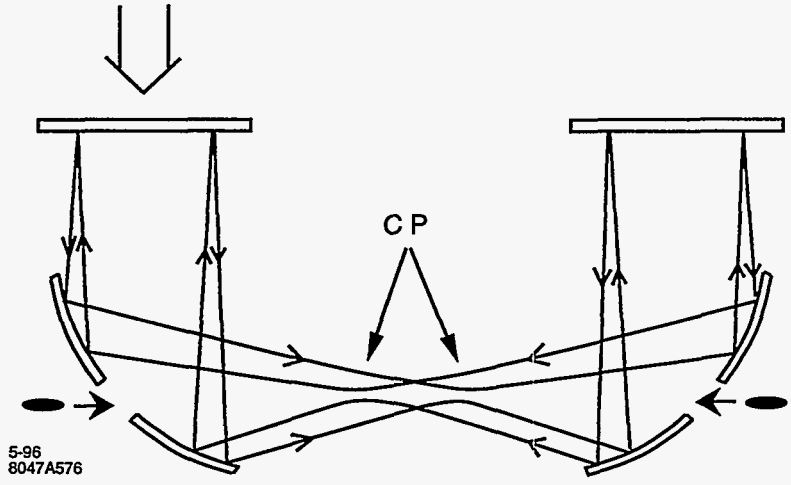


Figure B-26. *A double confocal resonator.*

For gamma-gamma colliders the desirable cavity mode should have certain characteristics. The double confocal resonator seems to be an ideal choice for this purpose. A double confocal resonator is effectively made of two usual confocal resonators. The usual confocal resonator is formed by two curved mirrors separated by a distance equal to the radius of curvature of the mirrors. Putting two confocal resonators together, one gets a ring resonator with four identical mirrors separated by equal distance. Folding such a ring resonator with two flat mirrors one gets the double confocal resonator shown in Figure B-26.

The double confocal cavity has several advantages:

- A dominant cavity mode can be made to have a central peak at each focal point and an annular shape at each mirror location.
- The focal spotsize or Rayleigh range is determined by the limiting intracavity apertures and are therefore easily adjustable.
- The annular mode distribution on the cavity mirrors allows large holes on the mirrors for electron beam and debris passage without sacrificing cavity Q.
- The annular shape on the mirrors also gives larger mode sizes thus reducing the power loading.
- With proper external mode for injection, two focal points can be formed each opposing to an incoming electron beam.
- In addition to these advantages, only one unavoidable transmissive window is used in Figure B-26 to couple laser pulse into the cavity in the vacuum system. Because a weaker pulse is needed for stacking, the power going through the window can be quite small compared to other approaches requiring full peak power injection.

The excitation of the desired mode in the cavity is largely dependent on the external mode preparation. Roughly speaking, the injection mode should have an annular amplitude distribution and a nearly flat phase front at the location of the transmissive window. If the external mode exactly matches the desired cavity mode, the cavity mode can be established right away, otherwise the resulting mode deterioration and slower intracavity power build-up could affect collider performance. There are ways to convert the usual Gaussian mode from a laser into an annular shape, for example by using an axicon or a profiled transmissive element. Fortunately, all these can be done outside the crowded interaction region and high-vacuum environment.

B.9 Gamma-Gamma Final Focus System

The goal of the FFS for the IR2 is to produce $\beta_x^* = \beta_y^* = 0.5\,\text{mm}$, as explained in Section B.3.3. However, the current version of the $\gamma\gamma$ final focus system has $\beta_x^* = 0.9\,\text{mm}$ and $\beta_y^* = 0.7\,\text{mm}$. This is not accidental because with these beta functions at the IP the chromaticity of the final focus doublet is the same as in the e^+e^- case. Efforts to find smaller β^* values resulted in higher x and y chromaticities, implying a greater sensitivity to the quadrupole placement tolerance and also a greater complexity and length of the FFS. The current solution of the FFS for the IR2 has a similar complexity as that for the IR1, being essentially an adaptation of the already existing solution for e^+e^- case to the new regime of the $\gamma\gamma$ collision.

B.9.1 Beam Parameters

Table B-7 lists the electron beam parameters necessary for a discussion of a final focus system of a $\gamma\gamma$ interaction region. Note that beta-functions at the IP are optimized differently from the ones for the e^+e^- collision—with a relaxed β_y^* and a tighter β_x^*. The following considerations are behind this choice: High-energy γ quanta appearing in a conversion of photons on electrons are emitted within the angular spread $1/\gamma$ towards a direction of the parent electron.

Thus, a spotsize of the high-energy γ beam at the IP reflects a spotsize of the electron beam plus additional contribution due to the angular spread of the γ quanta at the CP. This contribution, $\delta\sigma_{x,y}^*$, depends from a

E (gev)	250
ϵ_x (m×rad)	1×10^{-11}
ϵ_y (m×rad)	1×10^{-13}
β_x^* (mm)	0.9
β_y^* (mm)	0.7
Horizontal beam size at the IP (nm)	96
Vertical beam size at the IP (nm)	8.5

Table B-7. *Beam parameters*

distance b from the CP to the IP

$$\delta\sigma_{x,y}^* = b/\gamma \quad . \tag{B.27}$$

We would like to have

$$\delta\sigma_{x,y}^* \lesssim \sigma_{x,y}^* \quad . \tag{B.28}$$

Therefore,

$$b \lesssim \sigma_{x,y}^* \gamma \quad . \tag{B.29}$$

Thus, it is more convenient to have larger β_y^* and, correspondingly, larger σ_y^* because it allows the CP to be placed further away from the IP. At the same time, it is necessary to reduce β_x^* to compensate the reduction in the luminosity when β_y^* is increased. Note that, contrary to the e^+e^- case, there are no constraints on the beam profile at the IP arising from the beamstrahlung effects.

B.9.2 Final Focus Doublet

For linear colliders, the chromaticity is defined as the change in the IP waist position when the particle has an energy unequal to the design energy. A change in the horizontal or vertical waist position is characterized by the presence of a term x'^2 or y'^2 in the beam line Hamiltonian, which is precisely the terms in the Hamiltonian of a drift. The horizontal or vertical chromaticity is thus characterized by the presence of terms $\delta x'^2$ or $\delta y'^2$, where δ is the fractional energy error. Since the change in the horizontal or vertical IP position due to the these terms is given by

$$\Delta x = \frac{\partial H}{\partial x'} \qquad \text{or} \qquad \Delta y = \frac{\partial H}{\partial y'} \tag{B.30}$$

and the derivative of a quadratic function introduces a factor of 2, it is usual to define a horizontal and vertical chromatic length L_x^c and L_y^c as the coefficient of $\delta x'^2$ or $\delta y'^2$ respectively.

The spread in spotsize from spread in incoming angle can be written

$$\frac{\Delta x}{\sigma_x^*} = \delta\xi_x \frac{x'}{\sigma_{x'}^*} \quad , \quad \text{where} \quad \xi_x \equiv \frac{L_x^c}{\beta_x^*} \quad . \tag{B.31}$$

The quantity ξ_x is called the chromaticity. It is a dimensionless number. A value of one would indicate that the chromatic aberration gives a contribution to the beam size equal to the linear term. the contribution to this term coming from the final doublet can be calculated by the formula

$$\int ds\, k_x(s)\beta_x(s) \quad . \tag{B.32}$$

	Length, (m)	Gradient, (kg/mm)	Bore radius, (mm)	Beam-stay-clear
F-quad	1.43	4.50	3	$11\sigma_x$, $50\sigma_y$
D-quad	1.16	−3.38	4	$19\sigma_x$, $41\sigma_y$

Table B-8. *Doublet parameters*

since the phase advance from the IP to elements of the doublet is very close to $\pi/2$.

It is well known that the chromaticity of a lattice with only quadrupoles will be non zero, and to compensate the chromaticity one must add sextupoles to the beam line. To minimize sextupole strength, one first seeks the quadrupole configuration that minimizes the chromaticity.

The minimum chromaticity from quadrupoles occurs by placing strong quadrupoles as close to the IP as possible, without interfering with the function of the detector. Doing this in both planes of course requires a doublet. The rays on the IP side of the doublet will be focused to the IP. The divergence of the rays on the upstream end of the doublet will have a divergence that is characterized by the length of typical beam line modules of phase advance $\pi/2$ or π. Thus the function of the doublet is to focus parallel rays to the IP.

Thus the doublet design can be separated from the beam line design. One first seeks the double giving the minimum chromaticity parameters, taking into account detector requirements, constraints of materials available for quadrupole fabrication, tolerances on quadrupole position and field strength, and synthrotron radiation within the doublet.

As a first attempt to design the final focus system for $\gamma\gamma$ collisions, we decided to keep chromaticity of the final focus doublet close to the chromaticity of the e^+e^- final focus. Thus, with $L^* = 2\,\mathrm{m}$, the minimum beam-stay-clear requirement of $10\sigma_{x,y}$, and the maximum pole-tip field in the permanent magnet quadrupoles of 1.35 T, we arrived at the doublet parameters described in the Table B-8. Note, that the quadrupole nearest to the IP is of the F-type and the length of the drift space between the F and D quadrupoles is 0.3 m. This doublet has an x-plane chromaticity $\xi_x = 3100$ and y-plane chromaticity $\xi_y = 24500$. These values are to be compared with $\xi_x = 1100$ and $\xi_y = 23000$ in the e^+e^- case.

The size of these chromaticities indicate the precision with which the chromatic correction must be made. This has, of course, direct implications for the system tolerances. It also places constraints on the synchrotron radiation allowed in the system, because the change of particle energy within the system implies a failure of the chromatic balance built into the optics. However, for the present design of the doublet, the increase of the beam spotsize at the IP due to the Oide effect is negligible.

B.9.3 Chromaticity Compensation

We follow a standard approach to the chromaticity compensation of the final focus doublet. Similar to the e^+e^- final focus system, the $\gamma\gamma$ final focus system consists of the five modules. These are, in order of their location beginning from the IP: final transformer (FT), vertical chromaticity correction section (CCY), beta-exchanger (BX), horizontal chromaticity correction section (CCX) and beta-matching section (BMS). All modules have exactly the same functions as in the e^+e^- case. Figure B-27 shows the beta and the dispersion functions along the final focus system from the entrance of the BMS to the IP. The total distance from the entrance of the BMS to the IP is about 1600 m.

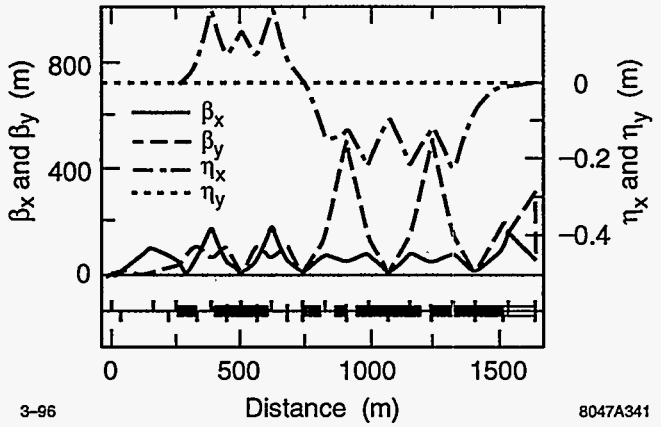


Figure B-27. *Horizontal and vertical beta functions from BMS to IP for 1600 m-long $\gamma\gamma$ final focus system at 500 GeV.*

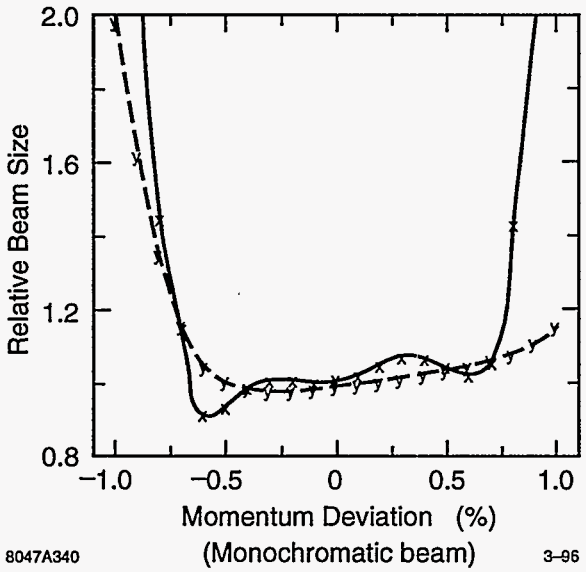


Figure B-28. *Relative increase of spotsizes as a function of momentum offset Δ of a monoenergetic beam.*

Figure B-28 shows the vertical and horizontal beam size as a function of the momentum offset Δ of a monoenergetic beam. The spotsizes are given in units of the values for zero momentum offset. The momentum bandwidth for a 10% blow-up of either spotsize is $\pm 0.6\%$.

It was recently shown (Dick Helm) that by increasing the overall length of the Final Focus section to 1750 m one can have $\beta_x^* = \beta_y^* = 0.5\,\text{mm}$ with the energy bandwidth about $\pm 0.5\%$.

B.10 Extraction and Diagnostic Line

The extraction line has the important functions of optimizing luminosity, characterizing beam properties at the IP and transporting beams from the IP to a dump with minimal background in the particle detector from this or any secondary function along the way. Beam characterization includes measurements of current, position, profile, energy, polarization, and low-order correlations on a bunch-to-bunch basis for feedback and stabilization. A prototype optical and diagnostic layout will be developed that provides such functions. In many respects it will be similar to the corresponding e^+e^- dump line that was presented in Section 11.8.

One presumed difference with that line is the difficulty of disposing of the high-power photon beam. We note that the outgoing beamstrahlung power approaches 10% of the incident, primary beam power for the e^+e^- IR. This was dealt with by implementing a common dump for both leptons and photons. The main question to be answered here is what distance is then available and whether it is reasonable from the standpoint of detector backgrounds and the various secondary functions that one would like to implement.

To accomplish all of the various tasks, we need to know the detailed composition and characteristics of the outgoing beam(s) under different, possible circumstances. Clearly, these characteristics depend on those assumed for the incoming beam. Based on some assumed set of incoming beam parameters, the transport optics, and the CP and IP conditions, we can then make predictions for the outgoing beams that are used to guide the design of the outgoing, beam line optics. Typically, the outgoing beams have a significant number of pairs as well as the primary, degraded electrons from the Compton conversion and more photons than leptons from multiple scattering within the strong laser conversion field.

Thus, the primary difference for this IP is the degraded electron beam that is highly disrupted but also necessarily includes a significant fraction of electrons with their full incident energy.

Clearly, the detector fields begin to have a serious influence on the incoming and outgoing beam characteristics with such large crossing angles. Thus, if background simulations for the detector imply an unacceptable situation, this procedure has to be iterated until a consistent solution is achieved. This is just now being done for the e^+e^- channel as described in Chapter 12.

While our overall goal is to optimize the luminosity while disposing the various beam components into their respective dumps, it is also important to provide any monitoring and feedback that can optimize the usable collision rate at the IP. Thus, beyond simply dumping the beam, there are other functions that run from the absolutely necessary to the desirable that will be considered as part of the dump line in roughly descending order of importance:

- Beam control and stabilization.
- Diagnostics and monitoring—including luminosity and polarization.
- e, μ, n, γ secondary beams and parasitic experiments.
- Polarized sources for γ, μ and e^+ beams.
- Energy recovery.

Separate discussions on these possibilities can be found in Section 11.8 of the ZDR for the e^+e^- channel. While the outgoing line is more difficult here, it can be accomplished in a similar way to that discussed in Section 11.8 because that design was made to be compatible with such a possibility there.

Energy per micropulse	1 J
Pulse duration	1.8 ps
Focusing f-number	7
Wavelength	0.8-1 μm
Pulse structure	90 micropulses(1.4-ns separation)
	180-Hz macropulse rate
Average power	16 kW

Table B-9. *Laser requirements for NLC γ-γ option*

B.11 Laser Technology I: Solid State Lasers

As we have seen in Section B.3.2, the optical beam for $\gamma\gamma$ or γe^- colliders consists of a sequence of TW micropulses, each a few ps long, with an average power of tens of kW. The requirements are summarized in Table B-9. Unless multipass optics can be used in the interaction region, these are the requirements for the lasers. Solid state lasers meeting the requirements would probably be available for the colliders to 500 $\times$ 500 GeV, but FEL would be used for higher energy colliders.

The solid state lasers for the $\gamma\gamma$ colliders have been discussed in two recent articles [Meyerhofer 1995a, Clayton 1994, Clayton 1995]. There will be two of these laser systems, one each for each colliding electron beam. While the energy, pulse duration, and focusing can be met with currently operating lasers, [Perry 1994a], these lasers have not yet met the average power requirements. The average power of high peak power systems has, however, been increasing rapidly recently, driven by activities such as the Isotope Separation program at LLNL and facilitated by the development of high power laser diode pump sources. The system requirements could also be reduced by using a multipass conversion point. It is expected that the system requirements will be met with a series of 1-kW, diode-pumped, solid-state, chirped pulse amplification laser systems. These unit cells will be fed by a single, phase-locked oscillator to insure timing stability.

Many of the components of the required laser system can be achieved with technology which is currently being developed for applications other than the $\gamma\gamma$ collider. As a result, a single unit cell prototype laser module could be developed over the next few years.

B.11.1 Laser Materials

Although both dye and excimer laser systems can easily meet the short-pulse requirement of the NLC, achieving the energy and beam quality requirements with lasers based on these materials would be difficult and expensive. The difficulty is associated with the low saturation fluence (energy storage) of these materials. The saturation fluence and upper state lifetime limit the amount of energy which can be stored and extracted per unit area (volume) from a laser material. Pulse energies on the order of one J would require laser apertures of approximately 500 cm^2 for dye- and excimer-based systems. Solid-state lasers offer an increase in saturation fluence between two and three orders of magnitude greater than dye or excimer lasers. This makes possible the development of extremely compact, high-energy lasers based on solid-state lasing media. Unfortunately, the high saturation fluence of solid-state materials cannot be accessed directly with short-pulses due to limitations on peak power imposed by the nonlinear refractive index.

In principle, production of a 1-J, 1-ps laser pulse could be accomplished by producing a low-energy 1-ps, 1.053-μm laser pulse and directly amplifying it in a Nd:Glass laser chain. Because of the high saturation fluence, the final amplifier cross section could be less than $1\,\mathrm{cm}^2$. The extracted power density would be in excess of $1\,\mathrm{TW/cm}^2$, close to the damage threshold of most materials [Stuart 1995a]. A second, more severe limitation on the amplification chain is provided by the nonlinear index of refraction in the material [Siegman 1986]. Self-focusing and filamentation of the laser pulse can occur when the accumulated nonlinear phase (B-integral) exceeds π,

$$B = \frac{2\pi}{\lambda}\int_0^L n_2 I(t)dz = \frac{2\pi}{\lambda} n_2 I_{\mathrm{peak}} L \quad , \tag{B.33}$$

where n_2 is the nonlinear index of refraction and L is the length of the medium. The power density in a medium is thus limited by the B-integral. For a given length amplifier, the peak intensity must satisfy $I_{\mathrm{peak}} < \lambda/2n_2L$. For a 10-cm-long Nd:Glass laser amplifier, the peak intensity must be less than $2\,\mathrm{GW/cm}^2$. For a pulse length of 1 ps and an energy of 1 J, the cross-sectional area must be greater than $A_L > E/I_{\mathrm{peak}}\tau = 500\,\mathrm{cm}^2$, which is to be compared with the $> 1\,\mathrm{cm}^2$ estimated from the saturation fluence. One solution to keeping the B-integral small is to use a larger diameter amplifier. Unfortunately, the use of a large diameter amplifier for short-pulse amplification eliminates the advantage of solid-state media.

B.11.2 Chirped-Pulse Amplification

This problem can be overcome by the use of Chirped-Pulse Amplification (CPA) [Strickland 1985]. This technique allows smaller amplifiers to be used, which reduces the cost and increases the repetition rate. Chirped-Pulse Amplification circumvents self-focusing by temporally stretching the pulse before amplification and recompressing it afterwards. This reduces the B-integral in the lasing medium by the compression ratio, the ratio of the stretched to compressed pulse durations. While optical fibers and prisms can be used to stretch or compress a pulse, the simplest technique involves a grating pair. A grating pair can be used to impart a positive [Martinez 1987] or negative [Treacy 1969] chirp to a short pulse. The sign of the chirp is defined by the time derivative of the frequency.

The CPA concept is shown in Figure B-29. A short, low-energy pulse is generated in an oscillator. The pulse is then stretched by a factor greater than 1000 in grating pair. To obtain a positive chirp a telescope is used between the grating pairs to invert the sign of the natural negative chirp associated with grating dispersion. To limit the size, complexity and cost of the telescope, a positive chirp is generated when the beam size and energy are low (before amplification). The pulse is now long enough for safe amplification to high energy. The pulse is recompressed in a second, grating pair (no telescope), generating a high-energy, ultrashort laser pulse.

The initial grating pair imparts a phase delay proportional to the frequency. This produces a pulse which has different frequencies spread out in time (chirped pulse). The second grating pair imparts a phase delay which is the inverse of the first grating pair, thereby removing the chirp and recovering the short pulse. Ideally, with the amplifiers placed between the two sets of gratings, the only change in the chirped pulse is its amplitude and the temporal characteristics of the pulse at the input and output are the same. Unfortunately the amplifier can modify both the amplitude [Perry 1990a, Chuang 1993] and phase [Chuang 1993] structure of the pulse. The amplitude changes include gain narrowing and pulling [Perry 1990a]. The primary phase change is self-phase modulation of the chirped pulse in the amplifier chain [Chuang 1993, Perry 1994b]. All of these effects must be minimized in order to obtain optimal pulses upon compression.

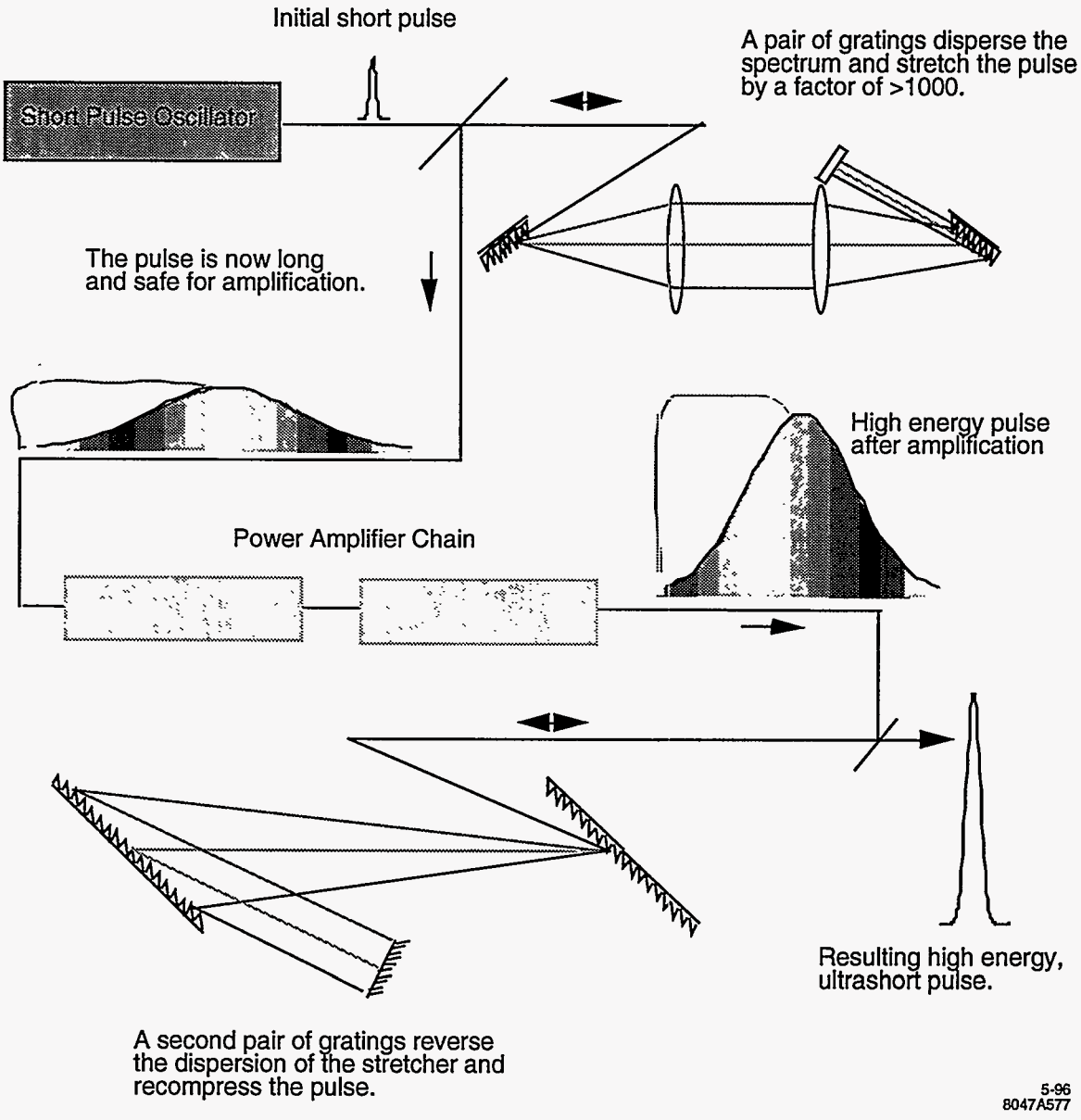


Figure B-29. *Chirped-pulse amplification concept.*

B.11.3 High-intensity, Short-pulse Laser Systems

There have been very dramatic advances in the technology and application of chirped-pulse amplification technology in the past five years [Perry 1994a]. Original CPA systems employed conventional mode-locked Nd:YAG or Nd:YLF oscillators producing transform-limited 50–100 ps pulses. The bandwidth necessary to achieve picosecond pulses was obtained by passing the 50–100-ps output of the oscillator through long (50–1000 m) lengths of single-mode fiber. Self-phase modulation in the fiber produced a chirped pulse with increased bandwidth. The pulse duration was increased from the original 50–100 ps by group-velocity dispersion in the fiber or the addition of a grating-pair pulse stretcher as mentioned previously. This method for obtaining stretched, chirped pulses has been rendered obsolete by the development of Kerr lens mode-locked oscillators which directly produce transform-limited 10–1000 fsec pulses. These pulses are then directly stretched in time by factors on the order of 10^4 by the grating stretcher. Elimination of the complex nonlinear interaction in the fiber is one of the many important advances which have led to the rapid and widespread acceptance of chirped-pulse amplification lasers.

The pulse exiting the stretcher has an energy of 0.10-1 nJ and a duration of ~1 ns. In the power amplifier section, the pulse must be amplified by a factor of 10^9 to the Joule level. This is typically achieved through the use of a multipass regenerative amplifier and a series of single or double pass amplifiers. In a regenerative amplifier, a pulse is switched into a cavity and makes multiple passes through an amplifier and is subsequently switched out. The regenerative amplifier brings the energy to the few milliJoule level (an increase of seven orders of magnitude) while the single and double pass amplifiers provide the remainder of the gain. Because of the losses inherent in the amplifiers, the total gain is of order 10^{10}–10^{11}.

After amplification, the pulse is compressed back to a short pulse by a pair of compression gratings. The damage threshold of these diffraction gratings is one of the most important limits on CPA lasers. The gratings must exhibit high damage threshold for both the long duration of the stretched pulse (ns) and the short duration of the compressed pulse (ps). Since the physical mechanism of optical damage changes from the nanosecond to picosecond regime [Stuart 1995a], producing high-damage threshold diffraction gratings is both a scientific and technical challenge. Commercially-available metallic gratings exhibit damage thresholds as high as 250 mJ/cm^2 for ps pulses. Recent advances in grating technology [Boyd 1995] have increased the damage threshold of metallic gratings to over 400 mJ/cm^2 for pulses in the range 0.1 to 50 ps. While these gratings do enable a factor of two increase in the peak power density achievable with most CPA systems, they will not be useful for the combined high average and high peak power requirement of the NLC. A new concept in diffraction gratings based on multilayer dielectric materials [Perry 1996] can, in principle, meet both the average power and peak power requirements of the NLC. These gratings have achieved a damage threshold of 600 mJ/cm^2 for 100 fsec pulses [Perry 1995]. They should exhibit a significantly increased threshold for average power damage relative to thin-film metallic gratings, however this is yet to be demonstrated.

One must also be concerned about power density of the compressed pulse in transporting it to the interaction point. If the intensity exceeds 10^{12} W/cm^2 it may cause plasma formation on solid surfaces [Stuart 1995a]. In addition, the nonlinear index of refraction of the pulse in air, in any windows, and in the focusing elements will destroy the beam quality of the pulse and make it impossible to focus. For example a 10^{12} W/cm^2 pulse passing through a 1-cm-thick glass window ($n_2 = 3 \times 10 - 16$ cm^2/W) the B-integral is 6π so that significant self-focusing would occur. The output diameter of the TW pulse must be large enough so that the cumulative B-integral in the transport line is limited to less than approximately π. This means that in practice, the intensity of the compressed pulse should be kept below 10^{11} W/cm^2 before focusing, implying that the beam cross section must exceed 10 cm^2.

B.11.4 Synchronization and Repetition Rate

Two additional topics are important in the design of the laser system for a $\gamma\gamma$ collider. The laser beam must be synchronized to the electron beam to a fraction of the laser pulse duration and the repetition rate of the laser system must be matched to the repetition rate of the linac. The laser can by synchronized to the electron beam by driving the laser oscillator at a subharmonic of the linac rf. The phase relationship between the rf and the laser pulses can be maintained by electronic feedback [Rodwell 1986, Rodwell 1989]. Subpicosecond timing jitter has been demonstrated using these techniques [Rodwell 1986, Rodwell 1989]. A typical mode-locked oscillator operates with a driving frequency of 35–60 MHz, producing a 70–120-MHz pulse train. The length of the cavity is matched to driving frequency to produce the mode-locked, short duration, pulse train. The pulse train is detected with a fast photo-diode. The photo-diode signal is mixed with the input rf and error signal is used to phase shift the rf driving the mode-locker [Rodwell 1989]. It is important to note that to maintain the synchronization, the oscillator must be actively mode-locked. In addition, the path length of the laser after the oscillator, through the laser system and through the transport must be stable to less than the pulse duration, 0.3 mm for a 1-ps laser pulse.

B.11.5 1-ps, 1-J Laser System for Nonlinear QED Experiments

Recently a 0.5-Hz repetition rate, 1-μm, 1-ps, 1-J, chirped-pulse amplification (CPA) laser system has been demonstrated using a flashlamp-pumped, Nd:glass, zig-zag slab amplifier [Bamber 1995]. The system has been installed at the Stanford Linear Accelerator Center as part of the E-144 experiment to study nonlinear QED [Bula 1992] in collisions with 50-GeV electrons. With the exception of repetition rate and pulse format, many of the performance parameters required for $\gamma\gamma$ colliders are being examined with this system. In addition to demonstrating the laser energy, and pulse width, the jitter of the laser pulse with respect to the linac rf is currently less than 2 ps [Bamber 1995]. Compton-scattered gammas and recoil electrons due to multiphoton Compton scattering have been observed [Bula 1995]. Unfortunately, the average power of this laser system is four orders of magnitude less than required for the NLC $\gamma\gamma$ collider.

B.11.6 Average Power

As mentioned previously, the single pulse laser requirements for converting the NLC e^+e^- collider into a $\gamma\gamma$ or γe^- collider can be met by solid-state lasing materials. A number of different solid state materials are used for short-pulse, high-intensity, laser systems [Kmetec 1991, Perry 1991, Ditmire 1993, Beaudoin 1992, Salin 1991, White 1992] Both Ti:Sapphire [Salin 1991, Stuart 1995b] and Nd:Glass [Strickland 1985] have been used to generate high-intensity, ultrafast, laser pulses with wavelengths in excess of 1 μm. The advantage of using Ti:Sapphire is that its larger gain bandwidth allows shorter pulses to be generated and amplified, whereas a pure Nd:Glass system is limited to pulse durations of order 1 ps, which is sufficient for this application. The saturation fluence of Nd:Glass is approximately an order of magnitude higher than Ti:Sapphire, making Nd:Glass an attractive candidate for the NLC. Unfortunately, the thermal conductivity and thermal shock limit are low for glasses. As a result, although conventional Nd:glasses can meet the peak power requirements of the NLC, they cannot meet the average power requirements. The relevant properties of Nd:Glass are listed in Table B-10.

New glass hosts currently under development offer a nearly factor of two increase in the thermal shock limit. These glasses could, in principle, make possible a diode-pumped Nd:Glass based NLC laser. No laser has

Central wavelength	$1.053\,\mu m$
Gain bandwidth	$\sim 200\,Å$
Saturation fluence	$5\,J/cm^2$
Nonlinear index, n_2	$6 \times 10^{16}\,cm^2/W$

Table B-10. *Selected properties of Nd:phosphate glass lasing materials.*

yet been constructed from these advanced glasses. However, two kilowatt-class (long pulse) lasers are under development at LLNL utilizing these new glasses. Performance data from these lasers will be invaluable in analyzing the suitability of glass-based systems for the NLC.

In addition to developments in laser glass, recent advances in crystal hosts are also encouraging for NLC options. New crystals which have been "engineered" for diode-pumping and high average power operation are now emerging. One attractive candidate is Yb:S-FAP. This is a fluroapatite crystal host for the Yb lasing ion. The material functions well as a laser near $1\,\mu m$ and has sufficient bandwidth to support pulses of 2–5 ps in duration. It has thermo-mechanical properties which are substantially better than even the advanced glasses. Furthermore, it has a long upper-state lifetime ($\approx$1 ms) and an absorption band at 900 nm which make it nearly ideal for diode pumping with efficient AlGaAs diodes. The Yb:S-FAP crystal can not yet be grown in sufficient sizes to meet NLC requirements. However, a large amount of effort is currently being devoted to further developing S-FAP and related crystal hosts. It is reasonable to expect that large scale crystals would be available within one to two years.

In short, there are several options for meeting the average power requirements of the NLC. These options include: 1) direct, diode-pumped Nd:Glass based lasers incorporating advanced athermal glass, 2) direct, diode-pumped, broad-bandwidth crystals (*e.g.,* Yb:S-FAP or others) and, 3) two-stage laser-pumped lasers such as a long pulse ($\approx$10 ns) neodymium based laser pumping a short-pulse Ti:Sapphire laser. We have not yet performed an optimization study for the NLC laser which would compare the performance and cost of these various options. Such a study would be part of the conceptual design of the NLC $\gamma\gamma$ collider.

In addition to requiring advances in high average power laser materials, advances in diode laser technology are also required to meet the NLC specifications. However, there are major efforts on advancing diode laser technology already underway as part of both military- and civilian-led projects. High average power diode laser arrays which would meet the requirements of the NLC are already under development at LLNL and elsewhere. Current high peak power diode arrays have generated 1.45-kW average power [Beach 1994]. The continued development of diode laser technology and the associated thermo-mechanical systems will be only moderately influenced by the approval of the NLC. Instead, the NLC will reap the benefit of substantial development effort which is expected to produce diode packages which can meet the NLC requirements well in advance of the NLC construction schedule.

B.11.7 NLC Laser Concept

The proposed laser system for the NLC $\gamma\gamma$ option consists of two $\sim$16-kW laser systems built out of 1-kW unit cells. A schematic of the unit cell is shown in Figure B-30. All of the cells are fed by a single, phase-stabilized oscillator, ensuring synchronization of all of the laser pulses with the electron beam. Each of the unit cells consists of a series of diode-pumped, solid-state, laser amplifiers. The pulses are subsequently compressed in a grating pair and stacked into a single pulse train.

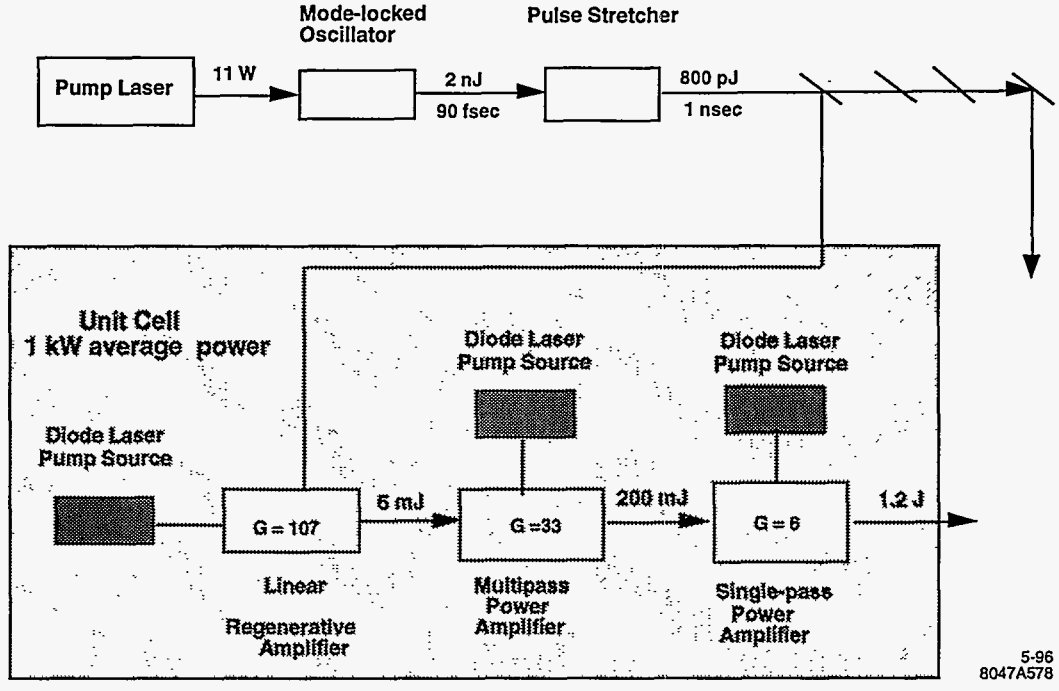


Figure B-30. *1-kW average power unit cell.*

Pulse stacking from individual unit cells into a single pulse train occurs via polarization switching as shown in Figure B-31. The output of a single unit cell is a small pulse train which is s-polarized (linearly polarized out of the paper in the figure). These pulses are reflected along the primary axis by a thin film polarizer. These polarizers are designed to reflect s-polarized light with greater than 99% efficiency while simultaneously transmitting p-polarized light with similar efficiency. Light from the first unit cell passes through a Pockels cell which has an applied voltage sufficient to provide a half-wave retardation. This rotates the polarization of the pulse 90° from s-polarization to p-polarization (linearly polarized in the plane of the paper). This p-polarized light now passes through the next thin film polarizer. The s-polarized pulses from the next unit cell are reflected from the thin-film polarizer. Combined with the pulses from the first unit cell, we have two sets of orthogonally polarized pulses incident on the next Pockels cell. The Pockels cell is initially held at ground while the p-polarized pulses from the first unit cell pass. After these pulses pass, the Pockels cell voltage is switched to half-wave voltage. The s-polarized pulses from the second unit cell are rotated to p-polarization upon transmission through the Pockels cell. The result is now a combined train of p-polarized pulses along the same optical axis. The procedure is repeated for each subsequent cell. In this scheme, the pulse-compression gratings probably should be placed after the pulse stacking so that all of the pulses will be compressed in the same grating pair and to keep the B-integral down.

B.11.8 A Ring Configuration for Multiplexing and Polarization Control

The baseline design would be to use each pulse once, with a single pass through a Pockels cell for polarization control. However, we will explore during the design the possibilities of re-using the pulse, thus significantly reducing the average laser power requirements, and hence cost. This was discussed in Section B.8.4. Here we

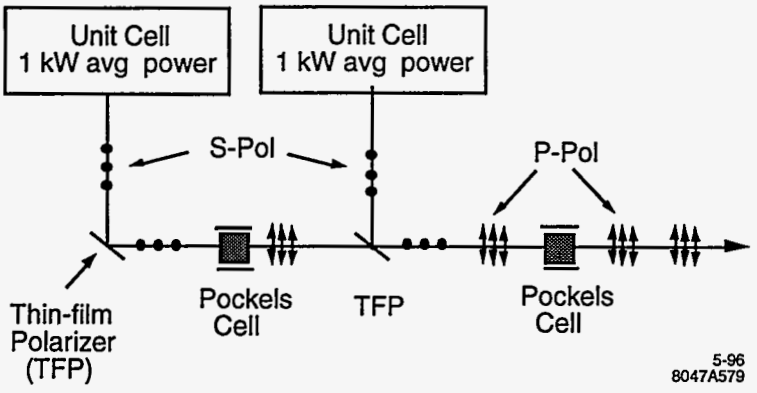


Figure B-31. *Pulse stacking/combination from individual unit cells onto a single optical axis is accomplished with electro-optic polarization switching.*

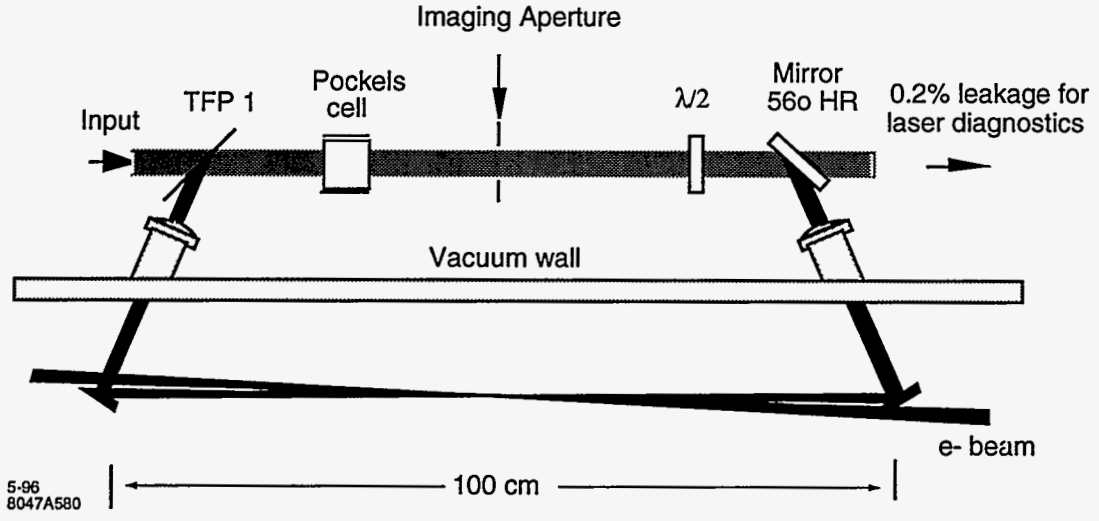


Figure B-32. *Relay imaged ring regenerative cavity.*

discuss in detail an example of such schemes based on a ring regenerative device first demonstrated at LLNL in 1990 [Perry 1990b]. The regenerative ring (Figure B-32) is a photon recirculator which is based on the original ring regenerative amplifier concept. The pulse is initially p-polarized (from the pulse stacker) and enters the cavity through a thin-film polarizer, TFP 1. The Pockels cell is initially at ground providing no phase retardation. The half-wave plate ($\lambda/2$) rotates the polarization 90° (now s-polarized, out of the plane of the paper). The s-polarized pulse reflects off of the high reflector and enters the vacuum chamber through a window. The laser beam is directed to the IR in the manner described in Section B.8.1 and directed out of the IR through a second window, before striking the original polarizer (TFP 1). The s-polarized pulse now reflects off of the polarizer and passes on its original path through the Pockels cell. However, the Pockels cell is now switched to half-wave voltage providing a 180° phase retardation which rotates the polarization back to the plane of the paper (p-polarization). The pulse is now trapped in the cavity and retraces its original path. Beam quality (focusabilty) is maintained by constructing the cavity either as a relay-imaged ring or as a TEM_{00} resonator.

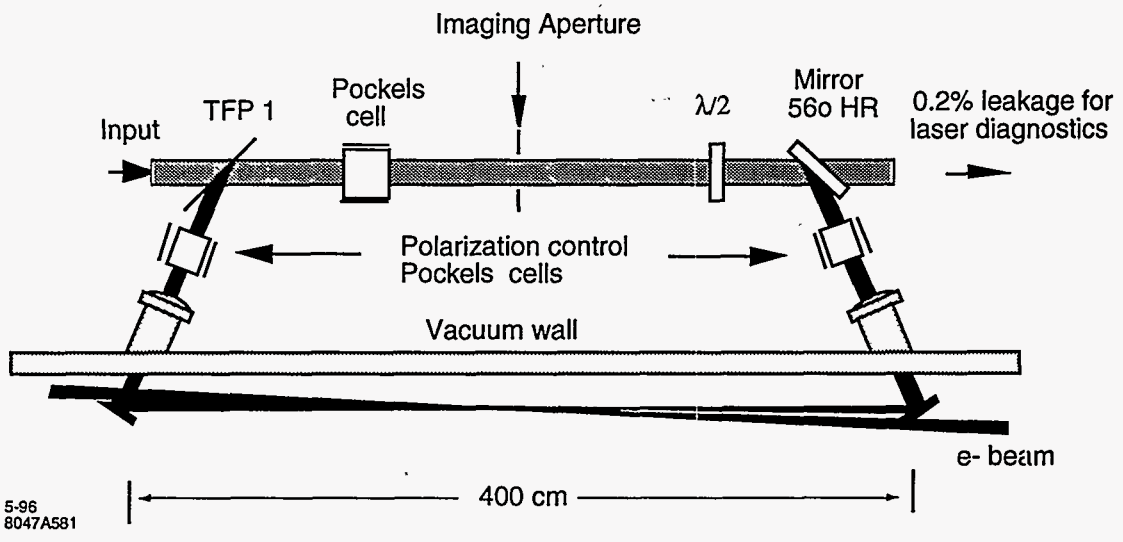


Figure B-33. *Laser photon recirculator and polarization control at the interaction point.*

The ring is constructed such that its cavity length exactly matches a multiple of the separation between microbunches of the electron beam. In this case, each time the laser pulse traverses the ring, it strikes a new electron microbunch at the IP. This effectively multiplies the repetition rate of the laser by the number of passes around the ring. A typical cavity round-trip time is 10 ns corresponding to a 3-m cavity. However, this can easily be adjusted to match the optimum electron bunch spacing as dictated by the rf accelerator. We have constructed rings for a round-trip time as short as 3 ns to over 30 ns.

The pulse will slowly decay in energy with each pass of the ring (ring down). The rate of decay is determined by the optical quality and reflectivity of the cavity optics. We have achieved a net cavity loss as low as 4% in a ring cavity of a design similar to Figure B-32. With extremely high quality optics as are commonly used in the Atomic Vapor Laser Isotope Separation (AVLIS) program, we should be able to achieve a cavity loss lower than 2%. With a 2% loss, 80 round trips drops the pulse energy to 20% of the original input. Even with only moderate quality optics, we routinely achieve 70 round trips in existing systems.

For the picosecond pulses envisioned for the $\gamma\gamma$ collider, self-focusing and self-phase modulation will limit the number of round-trips achievable for the standard design of Figure B-32. These problems can be overcome by incorporating a passive pulse stretcher/compressor into the ring and/or using all reflective focusing of the beam to the interaction point. Optimization of these designs and the effect on system performance is an important development task.

B.11.9 Polarization Control at the Interaction Point

Helicity control is a unique requirement of the laser system for the $\gamma\gamma$ collider. The desire to change the polarization from linear to circular and back on either a macrobunch or several macrobunch time scale in order to investigate the helicity dependence of various $\gamma\gamma$ reactions is a complicating factor to the laser design. The strong polarization dependence of the laser amplifier section, pulse compressor and pulse stacker demand that all helicity manipulation be performed after the pulse has left these elements. The use of the

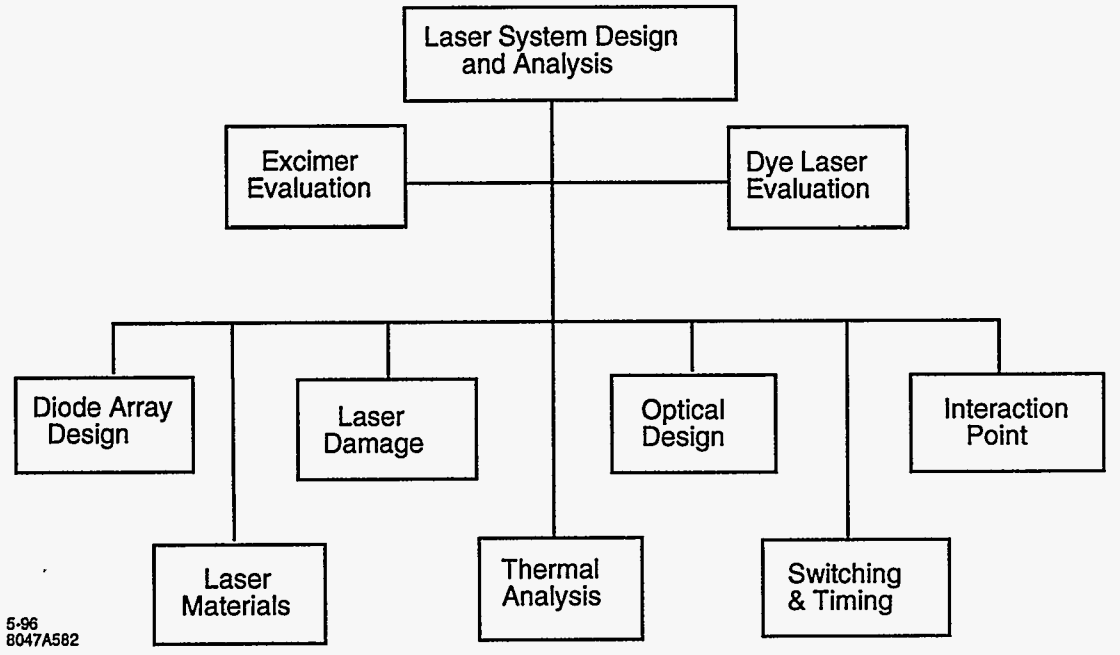


Figure B-34. *Conceptual design organization for NLC $\gamma\gamma$ collider laser system.*

photon recirculator of Figure B-32 is amenable to polarization control by addition of two additional Pockels cells as shown in Figure B-33. Before entering the vacuum chamber, the pulse passes through a Pockels cell which is either held at ground to produce no phase retardation and leaves the polarization unchanged (linearly polarized) or is switched to quarter-wave voltage to produce circularly polarized light. This Pockels cell can be easily switched at a 5-ms repetition rate (180 Hz) to enable switching between linear and circular polarization at the interaction region on alternating macropulses.

The ring scheme for reusing the laser beam and polarization control requires an extensive R&D. It should be demonstrated that there is the necessary space around the detector and that the B-integral in the Pockels cells, or the large cross-sectional areas of the cells, does not pose a problem.

These laser concepts should be considered preliminary at this time. Substantial conceptual design and optimization is yet to be done. A schematic of the conceptual design organization and issues to be addressed is shown in Figure B-34.

B.11.10 Conclusion

In summary, many of the technological advances required for the NLC $\gamma\gamma$ option have recently been achieved. The $\gamma\gamma$ collider portion of the NLC benefits substantially from the large national efforts which are devoted to the development of high peak power and high average power laser systems. While substantial design, optimization and development still needs to be done, our preliminary study suggests that within the next few years, many of the required laser components will be demonstrated and a prototype NLC laser module could be developed.

B.12 Free-Electron Lasers

Free-electron lasers (FEL) are another option for photon colliders, and they are especially interesting for higher energy colliders, where the required wavelength of the laser is longer than $\sim 1\,\mu$m, for which solid-state lasers do not presently exist. Several schemes have been proposed based on different combinations of FEL oscillators, amplifiers, and optical switching techniques. A scheme based on the chirped pulse amplification and compression, similar to the technique used in solid state lasers but replacing the amplifier with an FEL driven by an induction linac, is another attractive option.

B.12.1 An FEL Scheme Using Induction Linac and Chirped Pulse Amplification Technique

With the usual high-gain FEL amplifiers, it is difficult to produce the laser pulses of the characteristics outlined in this section. This can be readily understood if we note that the saturation power in high-gain FELs is given approximately by $P_{\rm sat} \sim \rho P_{\rm beam}$ where ρ is the FEL scaling parameter [Bonifacio 1984] and $P_{\rm beam}$ is the electron beam power; $P_{\rm beam} = EI$, where E and I are the beam energy and current respectively. Suppose we require the pulse energy $A = P_{\rm sat}\tau$ to be about 1 J. Assuming $\rho \sim 10^{-2}$, $I \sim 1\,$kA, and $\tau = 2$ ps, we find that E needs to be about 50 GeV. On the other hand, producing $\lambda = 1\,\mu$mFEL with such a high-energy electron beam requires a strong and long (100-m) wiggler magnet.

The discussion above also points to the solution of the problem. Namely if the pulse length were much larger, about 1 ns, then the required electron beam energy becomes $E = 100\,$MeV, which is quite reasonable for a 1-μm FEL. Thus the solution is to amplify the 1-ns pulse and later compress it to a few ps. A laser pulse can be compressed if it is chirped. Thus we are led to the idea of employing the chirped pulse amplification technique to FEL [Telnov 1991] extensively developed for solid-state lasers [Perry 1994a].

The scheme is schematically illustrated in Figure B-35. A solid-state laser produces a sequence of 1.8-ps, 0.14-mJ micropulses, with the same time structure as the collider beam, with an average power of 2.3 W. The micropulses are stretched (and hence chirped) to slightly less than 1.4 ns by means of a dispersive element schematically represented by a grating pair in the figure. The resulting optical beam which becomes amplified in a high-gain FEL driven by an induction linac producing $1.4 \times 90 = 130$-ns-long electron pulses at a 180-Hz repetition rate. The energy of each induction linac pulse may be modulated to match the chirped optical beam as indicated. Each amplified micropulse contains a few Joules of energy. The micropulses are compressed by another dispersive element to 1.8 ps. Of course, these optical components must designed to withstand high peak power (as is already true in SS laser compression systems) and high average power.

The electron beam and wiggler parameters required are shown in Table B-11.

A design of the induction linac producing the required electron beams is challenging but within the current state-of-the-art, and will be discussed in Section 9.2.3. The FEL consists of two sections, a 7.2-m-long uniform section in which the power gain length is about 1 m, and the input peak power of 100 kW is amplified to 140 MW. It is followed by a 7.2-m-long tapered section to further amplify the power to 1.6 GW. Thus, the energy contained in each micropulse is $1.6\,\text{GW} \times 1.4\,\text{ns} \simeq 2.24\,\text{J}$, which is larger than 1 J required for a conversion efficiency of 68%.

A scheme to amplify chirped pulses, similar to the one discussed here, but based on a regenerative FEL amplifier driven by an rf linac, was independently proposed recently by an LANL group [Chan 1995]. The

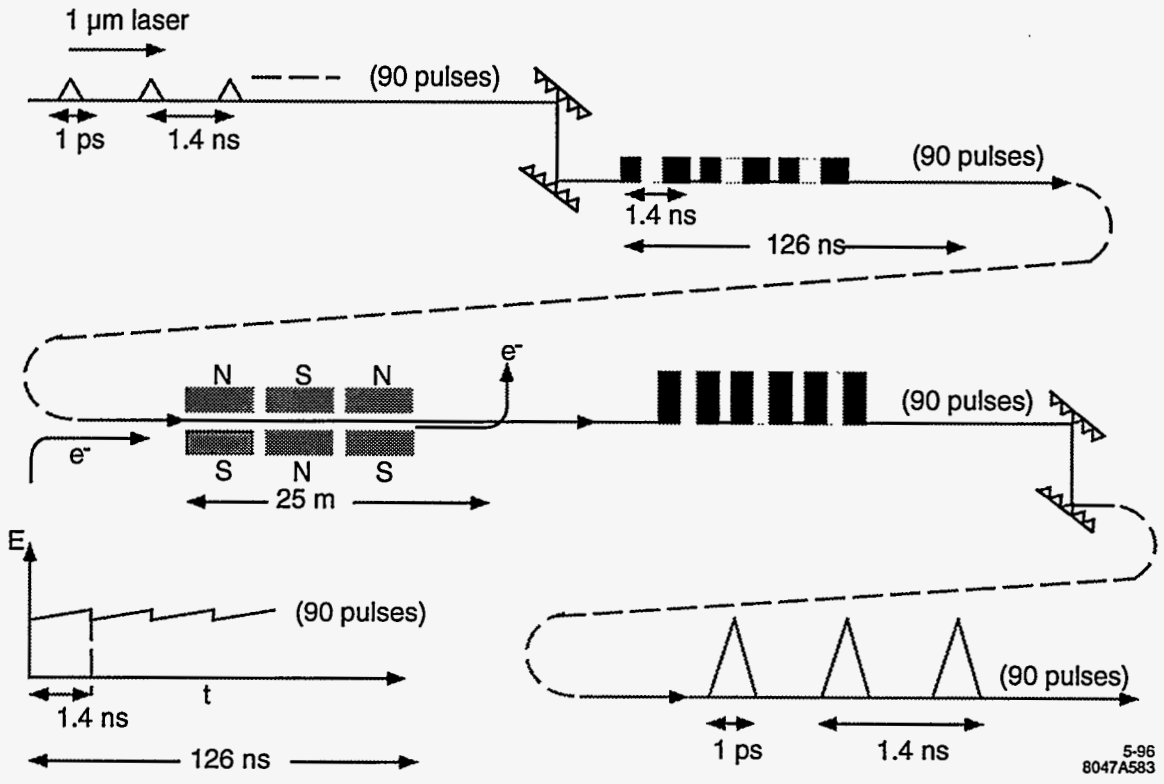


Figure B-35. *FEL pulse stretching/amplification/compressing scheme.*

Electron Beam Parameters:	
I	1 kA
E	100 MeV
$\Delta E/E$	10^{-3}
Rms normalized emittance (ϵ_N)	50×10^{-6} m-rad
Betatron wavelength (λ_β)	11.3 m
Undulator Parameters:	
λ_W	4.0 cm
K	1.4
Length of uniform section	7.2 m
Length of tapered section	10 m
Micropulse FEL Power:	
Input power	100 kW
Power after uniform section	140 MW(power gain length = 1 m)
Power after tapered section	1.6 GW
Energy per micropulse	1.6 GW × 1.4 ns = 2.24 J

Table B-11. *Parameters for an FEL Pulse Compression Scheme.*

scheme requires the use of intracavity optical switches operating in a high-power environment and focusing mirrors operating at a peak power density of $1\,\mathrm{GW/cm^2}$.

B.12.2 Chirping Requirement and Tolerance

The compression of the input pulse with electric field amplitude $E_i(t)$ to the output amplitude $E_f(t)$ can be represented by the linear transformation

$$\begin{aligned} E_f(t) &= \frac{1}{\sqrt{2\pi}} \int G(t-t')E_i(t')dt', \\ G(t) &= \frac{1}{\sqrt{\tilde{\mu}}} \exp\left(\frac{it^2\tilde{\mu}}{2} - i\omega_0 t\right) \end{aligned} \tag{B.34}$$

where $1/\tilde{\mu}$ is the time delay per unit frequency interval introduced by the dispersive element. An input chirped pulse may be represented by

$$E_i(t) = \exp\left(-i\left[\omega_0 t + \frac{\mu}{2}t^2\right] - \frac{t^2}{4\sigma_{\tau i}^2}\right) \quad , \tag{B.35}$$

where $\sigma_{\tau i}$ is the rms pulse length, and μ is the frequency change per unit time interval (chirping rate). Inserting Eq. B.35 into Eq. B.34, we find that the output pulse is also Gaussian, with the rms pulse length

$$\sigma_{\tau f} = \sqrt{\frac{1}{4\sigma_{\tau i}^2\tilde{\mu}^2} + \sigma_{\tau i}^2\left(\frac{\mu}{\tilde{\mu}} - 1\right)^2}. \tag{B.36}$$

Thus, if the grating is designed so that $\tilde{\mu} = \mu$, then $\sigma_{\tau f} = 1/2\mu\sigma_{\tau i}$. If $\sigma_{\tau i} = 1.4\,\mathrm{nsand}$ $\sigma_{\tau f} = 1.8\,\mathrm{ps}$, then we require $\tilde{\mu} = 2 \times 10^{20}/\mathrm{s}^2$. The total frequency chirping over the uncompressed pulse is $\Delta\omega \equiv= 2\mu\sigma_{\tau i} = 5.6 \times 10^{11}/$, or for $1\,\mu\mathrm{m}$ radiation $\Delta\omega/\omega = 2.8 \times 10^{-4}$. This is well within the gain bandwidth of the FEL. Therefore, it may not be necessary to modulate the electron beam energy.

There may be jitter introduced by the FEL amplification process. The jitter must be small so that $\sigma_{\tau i}^2(\mu/\tilde{\mu} - 1)^2 \ll 1/4\sigma_{\tau i}^2\tilde{\mu}^2$ or $\mu - \tilde{\mu} \ll 1/2\sigma_{\tau i}^2$. Thus the frequency error over the pulse must satisfy $\Delta\omega/\omega \ll 1/\omega\sigma_{\tau i} = 3.8 \times 10^{-7}$. To see the implication of the above requirement, we consider a high gain FEL in the exponential growth regime. The phase error due to a fluctuation $\Delta\rho$ in ρ is $\Delta\phi \sim 2\pi\Delta\rho L_W/\lambda_W$, where L_W is the length of the wiggler. Since ρ is proportional to $I^{1/3}/\gamma$, we have $\Delta\rho/\rho \sim (\Delta I/3I - \Delta\gamma/\gamma)$. Hence, $\Delta\omega/\omega = \Delta\phi/\omega\sigma_{\tau i} = (\lambda/c\sigma_{\tau i})(L_W/\lambda_W)\rho(\Delta I/3I - \Delta\gamma/\gamma)$. Taking $L_W = 20\mathrm{m}$, $\lambda_W = 4\,\mathrm{cm}$, $\lambda = 1\,\mu\mathrm{m}$, $\rho = 2 \times 10^{-3}$, we obtain $\Delta\omega/\omega \sim 10^{-5}(\delta I/3I - \Delta\gamma/\gamma)$. Therefore, if the fluctuations in $\Delta I/I$ and $\Delta\gamma/\gamma$ are a few percent, then the compression requirement is satisfied. Such a tolerance should be achievable.

B.12.3 Induction Linac Driver

The $\gamma\gamma$ induction linac driver beam parameters required for driving the FEL are listed in Table B-12 with beam parameters of several other induction linacs. The ATA was used to drive the Paladin FEL [Orzechowski 1990], the ETA II was used to drive the Intense Microwave Prototype (IMP) FEL [Allen 1992], and the TBNLC is a proposed induction accelerator driven relativistic klystron for powering the NLC [Caporaso 1995]. The $\gamma\gamma$ Induction Linac Driver will require significant improvement in the areas of beam

Parameter	$\gamma\gamma$ Induction Linac Driver	ATA (Paladin)	ETA II (IMP)	TBNLC Injector/main
E_b (MeV)	100	45	7	2.8/10[a]
I_b (kA)	1	0.8	2.4	1.2/0.6
ϵ_N(edge)	300	650(300 A)	900	400/800
(mm-mr)		1,116 (2.7 kA)		
$B_N(A/m^2 - r^2)$	2.2×10^9	1.4×10^8 (300 A) 4.6×10^6 (2.7 kA)	4×10^8 (1.5 kA) (2.5 MeV)	$1.5 \times 10^9/2 \times 10^8$
Flattop	150	>20	40	200
Pulse Length (ns)				
ΔE/E	$\pm$ 0.1%		$\pm$1%	$\pm$0.2%

[a] Final beam energy without rf power extraction would be 90 MeV.

Table B-12. *Comparison of the $\gamma\gamma$ Induction Linac Driver to other induction accelerators.*

brightness (B_N) and energy flatness (ΔE/E) with respect to measured values for ATA or ETA II. Note that the electron beam of the last induction linac built at LLNL (ETA II) had a measured brightness 1/5 of that required. ETA II, however, had much better beam quality than the earlier ATA. Thus, although induction linacs built to date have not produced the beam quality required for the $\gamma\gamma$ induction linac driver, it is reasonable to expect that improved computer modeling of induction injectors will lead to designs with higher beam brightness meeting the FEL requirements for $\gamma\gamma$ colliders. The use of photocathodes in induction injectors could also be explored to achieve the required beam quality.

ATA and ETA II represent about 10-year-old induction accelerator technology. A program to construct a prototype induction linac for the TBNLC is in progress at LBNL. Simulations have shown that the induction linac injector can produce a B_N of about $1.5 \times 10^9\ A/m^2 - r^2$ for a 1.2-kA, 2.8-MeV beam. A demonstration of this brightness will be significant to the $\gamma\gamma$ induction linac driver as its injector represents the largest source of brightness degradation (emittance growth). Energy regulation to achieve a ΔE/E of $\pm$0.2% is planned for the prototype, and will determine the feasibility of a ΔE/E of $\pm$0.1% for the $\gamma\gamma$ driver.

Beam energy is primarily a cost issue. Induction modules can be added to provide the desired beam energy. However, as the length of the induction linac increases, transverse instability of the beam motion can become a difficulty. The beam tube radius of the induction cells can be increased to lower the growth in the transverse instability as described below. The larger inner radius will lead to a larger induction core volume increasing the core cost and reducing efficiency. The transverse instability mechanism is well understood for induction linacs [Capporaso]. Assuming constant acceleration ($\gamma = \gamma_0 + \lambda z$) and a solenoid focusing field proportional to the beam energy, a figure of merit for transverse instability is given by the product of the betatron phase advance (φ_β) times the number of e-folds (f_e) of gain:

$$\varphi_\beta f_e = \frac{2}{\lambda^2}\frac{(\omega_0 Z_\perp)}{L_g}\frac{I}{I_0}\left(\sqrt{\gamma_f} - \sqrt{\gamma_0}\right)^2 \quad , \tag{B.37}$$

where I is the beam current, $I_0 = 17.03$ kA, L_g is the separation between induction module gaps, $(\omega_0 Z_\perp)$ is the transverse impedance of the gaps, and λ is the gap energy increase ($\Delta\gamma$) divided by L_g. The transverse impedance scales approximately as:

$$(\omega_0 Z_\perp) = \eta\frac{4w}{b^2} \quad , \tag{B.38}$$

where w is the gap width, b is the inner radius, and η is a design factor of order unity ($\eta = 1.3$ for ATA).

Three e-folds of gain ($f_e = 3$) in an accelerator length of about 16 betatron wavelengths ($\varphi_\beta = 100$) would be a reasonably conservative design for controlling transverse instability. Assuming a maximum field stress in the gap of 100 kV/ cm ($\Delta\gamma = 0.196$ and w = 1.0 cm), gap spacing of 30.0 cm, injector voltage of 1.5 MV, and $\eta = 0.7$ ("good" design), Eqs. B.37 and B.38 are used to find b $\geq$ 11.2 cm. The focusing field will increase from 183 G at 1.5 MeV to 1.3 kg at 100 MeV. The linac will have approximately 1,000 induction modules and be over 300 m in length.

Induction cells with 11-cm beam pipe radius will have about twice the core volume of the TBNLC cells (pipe radius = 5 cm). This volume increase will not only lead to added cost for core material, but also require an appreciable increase in the number of pulse power units than that of the TBNLC design. Although the design parameters have not been optimized, the $\gamma\gamma$ Induction Linac Driver is expected to cost significantly more per unit length than the TBNLC induction linac and have lower wall plug to beam power conversion efficiency.

References

[Allen 1992] S.L. Allen and E.T. Scharlemann, "The ETA II Linear Induction Accelerator and IM Wiggler: A High-Average-Power Millimeter-Wave Free-Electron-Laser for Plasma Heating", *Proc. 9th Intl. Conf. on High Power Part. Beams*, Washington, DC, (1992).

[Arutyunian 1963] F.R. Arutyunian and V.A. Tumanian, *Phys. Lett.* **4** 176 (1963); R. H. Milburn, *Phys. Rev. Lett.* **10** 75 (1963).

[Asaka 1995] T. Asaka, N. Maekawa, T. Moroi, Y. Shobuda, Y. Sumino, TU-483, May 1995, talk given at 5th Workshop on Japan Linear Collider (JLC), Tsukuba, Japan, 16-17 Feb 1995; e-Print Archive: hep-ph/9505371.

[Baillargeon 1995] M. Baillargeon, G. Belanger, and F.Boudjema, *Phys. Rev.* **D51**, 4712 (1995).

[Balakin 1994] V. Balakin, A. Sery, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355** 157 (1995).

[Bamber 1995] C. Bamber, T. Blalock, S. Boege, J. Kelly, T. Kotseroglou, A.C. Melissinos, D.D. Meyerhofer, W. Ragg, and M. Shoup III, *Opt. Lett.* submitted (1995).

[Barklow 1990] T. Barklow, SLAC–PUB–5364 (1990).

[Barletta 1993] W. Barletta *et al.*,, (Plasma Lens Collaboration), "Proposal for Plasma Lens Experiments at the Final Focus Test Beam", SLAC–Proposal–E–150 (April 1993).

[Beach 1994] R. Beach *et al.*, Conf. on Lasers and Electro-optics (CLEO), Anaheim, CA (1994).

[Beaudoin 1992] Y. Beaudoin, C.Y. Chien, J.S. Coe, J.L. Tapi, and G. Mourou, *Opt. Lett.* **17** 865–867 (1992).

[Ben-Menachem 1993] S. Ben-Menachem and P. Chen, SLAC–436, in *Proc. of the 5th Int. Workshop on Next-Generation Linear Colliders*, 426 (1993).

[Berestetskii 1982] V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, *Quantum Electrodynamics*, Section 101 (Pergamon Press 1982).

[Berkeley 1995] *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, 1994; *Nucl. Instr. Methods A* **355**, 1–194 (1995).

[Billhardt 1993] F. Billhardt, M. Kalashnikov, P.V. Nickles, and I. Will, *Opt. Commun.* **98**, 99–104 (1993).

[Bonifacio 1984] R. Bonifacio, C. Pellegrini and L.M. Narducci, *Opt. Commun.* **50**, 373 (1984).

[Borden 1993a] D.L. Borden, plenary talk presented at Workshop on Physics and Experiments with Linear e^+e^- Colliders, Waikoloa, Hawaii, 26–30 April 1993, ed. F.A. Harris, S.L. Olsen, S. Pakvasa, and X. Tata, (World Scientific, Singapore, 1993).

[Borden 1993b] D.L. Borden, D.A. Bauer, D.O. Caldwell, *Phys. Rev.* **D48**, 4018 (1993).

[Borden 1994] D.L. Borden, V.A. Khoze, W.J. Stirling, J. Ohnemus, *Phys. Rev.* **D50**, 4499 (1994).

[Boyd 1995] R.D. Boyd, J.A. Britten, D.E. Decker, B.W. Shore, B.C. Stuart, M.D. Perry, and L. Li, *Appl. Opt.* **34**, 1697–1706 (1995).

[Brodksy 1994] S. Brodsky, P. Zerwas, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 19 (1995).

[Bula 1992] C. Bula *et al.*, Study of QED at Critical Field Strength at SLAC, E-144 proposal (1992).

[Bula 1995] C. Bula *et al., Phys. Rev. Lett.* , submitted (1995).

[CAIN] CAIN (Conglomérat d'ABEL et d'Interactions Non-Linéaires) being developed by P. Chen, G. Horton-Smith, T. Ohgaki, A. Spitkovsky, A.W. Weidemann, and K. Yokoya.

[Caporaso 1995] "Relativistic-Klystron Two-Beam Accelerator Based Power Source for a 1-TeV Center-of-Mass Next Linear Collider: Preliminary Design Reports", LBID–2085 and UCRL–ID–119906, (Feb. 1995); G.J. Caporaso and A.G. Cole, "High Current Electron Transport", in *AIP Conference Proceedings* 249, The Physics of Particle Accelerators, eds. M. Month and M. Diens, AIP, NY, 1662–1672.

[Capporaso] G.J. Caporaso and A.G. Cole, "High Current Electron Transport", in AIP Conference Proceedings 249, The Physics of Particle Accelerators, eds. M. Month and M. Diens, 1662–1672, (AIP, NY).

[Chan 1995] K.C.D. Chan, J.C. Goldstein, D.C. Nguyen, and H. Tekeda, "A Chirped Pulse Regenerative - Amplifier FEL for the Gamma-Gamma Collider", LANL preprint, LA–UR–95–1502, presented at the 1995 Part. Acc. Conf., Dallas, TX (1995).

[Chanowitz 1994] M. Chanowitz, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355** 42 (1995).

[Chen 1989] P. Chen, V. Telnov, *Phys. Rev. Lett.* **63** 1796 (1989).

[Chen 1995a] P. Chen, G. Horton-Smith, T. Ohgaki, A.W. Weidemann, K. Yokoya, *Nucl. Instr. and Methods* **A35** 107 (1995).

[Chen 1995b] P. Chen, D. Bullock and D. Yu, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355** 130 (1995).

[Cheung 1994] Kingman Cheung, *Phys. Rev.* **D50**, 4290 (1994).

[Chien 1995] C.Y. Chien, G. Korn, J.S. Coe, J. Squier, G. Mourou, and R.S. Craxton, *Opt. Lett.* 20, 353-355 (1995).

[Chuang 1991] Y.-H. Chuang, D.D. Meyerhofer, S. Augst, H. Chen, J. Peatross, and S. Uchida, *J. Opt. Soc. Am.* B 8, 1226–1235 (1991).

[Chuang 1993] Y.-H. Chuang, L. Zheng, and D.D. Meyerhofer, *IEEE J. Quantum Electron.* **29** 270–280 (1993).

[Clayton 1994] C. Clayton, N. Kurnit, and D. Meyerhofer, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994).

[Clayton 1995] C.E. Clayton, N.A. Kurnit, and D.D. Meyerhofer, *Nucl. Instr. and Methods* **A355**, 121–129 (1995).

[Corkum 1988] P.B. Corkum, F. Brunel, N.K. Sherman, and T. Srinivasan-Rao, *Phys. Rev. Lett.* **61**, 2886 (1988).

[Ditmire 1993] T. Ditmire and M.D. Perry, *Opt. Lett.* **18** 426 (1993).

[Eboli 1993] O.J.P. Eboli, M.C. Gonzalez-Garcia, F. Halzen, D. Zeppenfeld, *Phys. Rev.* **D48**, 1430 (1993); O.J.P. Eboli, talk given at Workshop on Gamma-Gamma Colliders, Lawrence Berkeley Laboratory, Berkeley, CA, USA (March 28–31, 1994).

[Fawley 1980] W.M. Fawley and E.P. Lee, "Modelling of Beam Focusing and Kink Instability for Colliding Relativistic Electron and Positron Beams", LLNL report *UCID–18584*, (1980).

[Fawley 1987] W.M. Fawley and E.P. Lee, E.P. "Modelling of Beam Focusing and Kink Instability Particle in Cell Simulations of Disruption", in *Proc. New Developments in Par. Accel. Techniques*, ed. S. Turner, **CERN 870-11, ECFA 87/110**, 605–609 (1987).

[Ferray 1990] M. Ferray, L.A. Lompre, O. Gobert, A. L'Huillier, G. Mainfray, C. Manus, A. Sanchez, and A.S. Gomes, *Opt. Commun.* **75**, 278 (1990).

[Ginzburg 1981] I. Ginzburg, G. Kotkin, V. Serbo, V. Telnov, *Pizma ZhETF* **34** 514 (1981); *JETP Lett.* **34** 491 (1982); *Prep. INF* 81–50, Novosibirsk, in English (Feb. 1981).

[Ginzburg 1983] I. Ginzburg, G. Kotkin, V. Serbo, V. Telnov, *Nucl. Instr. and Methods* **205**, 47 (1983); *Prep. INP* 81–92, (Novosibirsk, Aug. 1981).

[Ginzburg 1984] I. Ginzburg, G. Kotkin, S. Panfil, V. Serbo, V. Telnov, *Nucl. Instr. and Methods* **219**, 5 (1984).

[Ginzburg 1994] I. Ginzburg, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 63 (1995).

[Gounaris 1995] G.J. Gounaris, J. Layssac, F.M. Renard, PM–95–11, May 1995. 23 pp.; e-Print Archive: hep-ph/9505430

[Grzadkowski 1992] B. Grzadkowski and J.F. Gunion, *Phys. Lett.* **B294**, 361 (1992).

[Gunion 1993] J.F. Gunion and H.E. Haber, *Phys. Rev.* **D48**, 5109 (1993).

[Gunion 1995] J.F. Gunion, J.G. Kelly, J. Ohnemus, *Phys. Rev.* **D51**, 2101 (1995).

[Heinrich 1991] J.G. Heinrich, C. Lu, K.T. McDonald, C. Bamber, A.C. Melissinos, D. Meyerhofer, Y. Semertzidis, Pisin Chen, J.E. Spencer, R.B. Palmer, "Proposal for a Study of QED at Critical Field Strengths in Intense Laser–High-Energy Electron Collisions", SLAC–Proposal–E–144 (October 1991).

[Heinz 1989] P. Heinz and A. Laubereau, *J. Opt. Soc. Am.* **B 6**, 1574 (1989).

[Hiramatsu 1995] S. Hiramatsu, S. Hashimoto and Y. Ishida, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 133 (1995).

[Hollebeek 1981] R. Hollebeek, "Disruption Limits for Linear Colliders", *Nucl. Instr. and Methods* **184**, 333 (1981).

[Horton-Smith phd] G. Horton-Smith, Ph.D. Thesis, SLAC.

[Injean 1994] H. Injean *et al.*, "Diode array, kilowatt laser development", Conference on Lasers and Electro-Optics (CLEO), pp. CThC1 (Anaheim, CA, 1994).

[Jikia 1993] G.V. Jikia, *Phys. Lett.* **B298**, 224 (1993); *Nucl. Phys.* **B405**, 24 (1993).

[Jikia 1994a] G. Jikia and A. Tkabladze, invited talk at Workshop on Gamma-Gamma Colliders, Berkeley, CA, 28-31 (Mar 1994).

[Jikia 1994b] G. Jikia, IFVE-94-77, 34 pp. (Jul 1994); e-Print Archive: hep-ph/9407393.

[Kamal 1995] B. Kamal, Z. Merebashvili, A.P. Contogouris, *Phys. Rev.* **D51**, 4808 (1995).

[Kim un] K.-J. Kim, A. Sessler, M. Xie, report al LC95, to be published.

[Kim 1994] Prepared by K.-J. Kim, P. Pierini, A. Sessler, V. Telnov at LC95, KEK, Japan, (March 1994).

[Kmetec 1991] J.D. Kmetec, J.J. Macklin, and J.F. Young, *Opt. Lett.* **16**, 1001–1003 (1991).

[Kon 1992a] T. Kon and A. Goto, *Phys. Lett.* **B295**, 324 (1992); F. Cuypers, G.J. van Oldenborgh, R. Rückl, *Nucl. Phys.* **B383**, 45 (1992).

[Kon 1992b] T. Kon, I. Ito, and Y. Chikashige, *Phys. Lett.* **B287**, 277 (1992).

[Kon 1993] T. Kon, *Phys. Lett.* **B316**, 181 (1993); F. Cuypers, G.J. van Oldenborgh, R. Rückl, *Nucl. Phys.* **B409**, 144 (1993).

[Kondratenko 1983] A. Kondratenko, E. Pakhtusova, E. Saldin, *Dokl. Akad. Nauk* **264**, 849 (1982).

[Kramer 1994] M. Kramer, J. Kuhn, M.L. Stong, and P.M. Zerwas, *Z. Phys.* **C64**, 21 (1994).

[Landau 1987] L.D. Landau and E.M. Lifshitz, "The Classical Theory of Field", Course of Theoretical Physics 2, (Pergamon Press, 1987).

[Martinez 1987] O.E. Martinez, *IEEE J. Quantum Electron.* **QE-23**, 59–64 (1987).

[Meyerhofer 1995a] D. Meyerhofer, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 113–120 (1995).

[Meyerhofer 1995b] D.D. Meyerhofer, *Nucl. Instr. and Methods* **A355** 113-120 (1995).

[Miller 1995] D.J. Miller, *Nucl. Instr. and Methods* **A355**, 101 (1995).

[Morton 1995] P. Morton and S. Chattopadhyay, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 138 (1995).

[Murayama 1994] H. Murayama, talk at Workshop on Gamma-Gamma Colliders, Berkeley, CA (28-31 Mar 1994).

[Norem 1995] J. Norem, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 166 (1995).

[Ohgaki 1995] T. Ohgaki and X. Yokoya, to be published.

[Orzechowski 1990] Reference on Paladin–T.J. Orzechowski at 1990 USPA School.

[Perry 1990a] M. Perry and F. Patterson, *Opt. Lett.* **15**, 381 (1990).

[Perry 1990b] M.D. Perry and F.G. Peterson, LLNL Internal Documents (1990); M.D. Perry, T. Ditmire, and D. Strickland, *Opt. Lett.* **17**, 601 (1992).

[Perry 1991] M.D. Perry and R. Olson, *Laser Focus World* bf 27, 69-74 (1991).

[Perry 1994a] For a review, see M. Perry and G. Mourou, M.D. Perry and G. Mourou, *Science* **264**, 917-924 (1994).

[Perry 1994b] M.D. Perry, T. Ditmire, and B.C. Stuart, *Opt. Lett.* **19**, 2149 (1994).

[Perry 1995] M.D. Perry, R.D. Boyd, J.A. Britten, D. Decker, B.W. Shore, C. Shannon, and E. Shults, *Opt. Lett.* **20**, 940–942 (1995).

[Perry 1996] M.D. Perry, J.A. Britten, R.D. Boyd, H. Nguyen, B.W. Shore, U.S.Patent Pending.

[Raidal 1995] M. Raidal, *Nucl. Phys.* **B441**, 49 (1995).

[Rajagopalan 1994] S. Rajagopalan, D. Cline and P. Chen, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 169 (1995).

[Rajagopalan 1996] S. Rajagopalan, D.B. Cline, P. Chen, "Application of a Plasma Lens to $e^-\gamma$ Colliders", these Proceedings.

[Richard 1995] F. Richard, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA (1994); *Nucl. Instr. Methods A* **355**, 92 (1995).

[Rodwell 1986] M.J.W. Rodwell, K.J. Weingarten, D.M. Bloom, T. Baer, and B.H. Kolner, *Opt. Lett.* **11**, 638-640 (1986).

[Rodwell 1989] M.J.W. Rodwell, D.M. Bloom, and K.J. Weingarten, *IEEE J. Quantum Electron.* **25**, 817–827 (1989).

[Ronan 1993] M. Ronan, *Proc. of Workshop on Phys. and Exper. with Linear Colliders*, Waikoloa, Hawaii, US (1993).

[Salin 1991] F. Salin, C. Rouyer, J. Squier, S. Coe, and G. Mourou, *Opt. Commun.* **84**, 67 (1991).

[Saldin 1995] E. Saldin *et al.*, Preprint JINR E-9-94-74, Dubna,1994, Submitted to Particle and Accelerators; *Nucl. Instr. Methods A* **355**, 171; DESY 94-243,1994 (1995).

[Siegman 1986] A.E. Siegman, *Lasers* (University Science, Mill Valley, Ca, 1986).

[Silvestrov un] G. Silvestrov, V. Telnov (unpublished).

[Squier 1991] J.A. Squier, F. Salin, J.S. Coe, P. Bado, and G.A. Mourou, *Opt. Lett.* **15**, 85 (1991).

[Strickland 1985] D. Strickland and G. Mourou, *Opt. Commun.* **56**, 219 (1985); P. Maine, D. Strickland, P. Bado, M. Pessot, and G. Mourou, *IEEE J. Quantum Electron.* **QE-24**, 398 (1988).

[Stuart 1995a] B.C. Stuart *et al.*, submitted to JOSAB, UCRL–JC–120, 225.

[Stuart 1995b] B. Stuart, S. Herman, and M.D. Perry, *IEEE J. Quantum Electron.* **31**, 528 (1995).

[Stuart 1995c] B. Stuart, M.D. Feit, M.D. Perry, A.M. Rubenchick, and B.W. Shore, *Phys. Rev. Lett.* **74**, 2248 (1995).

[Takahashi 1994] T. Takahashi, "Physics of e^+e^-, γe^- and $\gamma\gamma$ collisions at Linear Accelerators", *Proc. INS Workshop*, Tokyo, Japan, 93 (1994).

[Takahashi 1995] T. Takahashi, Hiroshima Univ. Preprint, HUPD–9526 (1995).

[Takahashi 1996] T. Takahashi, in preparation.

[Tandean 1995] J. Tandean, *Phys. Rev.* **D52**, 1398 (1995).

[Tauchi 1993] T. Tauchi, K. Yokoya, P. Chen, *Particle Acclerators* **41**, 29 (1993).

[Telnov 1990] V. Telnov, *Nucl. Instr. Methods A* **294**, 72 (1990).

[Telnov 1991] V. Telnov, private communication and *Proc. of Workshop on Phys. and Exper. with Linear Colliders*, Lapland, Finland, 551 (1991).

[Telnov 1993] V. Telnov, *Proc. of Workshop on Physics and Experiments with Linear Colliders*, Wokoloa, Hawaii (World Scientific, 1993).

[Telnov 1995] V. Telnov, *Proc. of Workshop on Gamma-Gamma Colliders*, Berkeley CA, USA, (1994); *Nucl. Instr. Methods A* **355**, 3 (1995).

[Telnov priv] V. Telnov, private communication.

[Treacy 1969] E.B. Treacy, *IEEE J. Quantum Electron.* **QE–5**, 454 (1969).

[Watanbe 1995] I. Watanabe, talk presented at Workshop on Physics and Experiments with Linear e^+e^- Colliders, Appi, Japan, (1995).

[White 1992] W.E. White, J. Hunter, L. Van Woerkum, T. Dimitre, and M.D. Perry, *Opt. Lett.* **17**, 1067 (1992).

[Xie 1995] M. Xie, K.-J. Kim, and A.M. Sessler, *Nucl. Instr. and Methods* **355**, 163 (1995).

[Yehudai 1990] E. Yehudai, *Phys. Rev.* **D41**, 33 (1990).

[Yehudai 1991] E. Yehudai, *Phys. Rev.* **D44**, 3434 (1991).

[Yokoya, 1986] K. Yokoya, KEK–Report 85–9, October 1985; also *Nucl. Instr. and Methods* **A251**, 1 (1986).

[Yokoya 1996] *User's Manual of CAIN 2.0*, (KEK publication, April, 1996).

Contributors

- Karl van Bibber
- Swapan Chattopadhyay
- Bill Fawley
- Dick Helm
- Tim Houck
- John Irwin
- Kwang-Je Kim
- Dan Klem
- David Meyerhofer
- Hitoshi Murayama
- Micheal Perry
- Mike Ronan
- Andy Sessler
- Jim Spencer
- Tohru Takahashi
- Valery Telnov
- Achim Weidemann
- Glen Westenskow
- Ming Xie
- Kaoru Yokoya
- Sasha Zholents

C

Ground Motion: Theory and Measurement

Contents

C.1 Introduction 1044
C.2 Theory 1044
C.2.1 Ground Motion Spectra and Related Functions 1045
C.2.2 The Lattice Response Function $G(k)$ 1052
C.2.3 Combining Spatial and Frequency Dependence 1056
C.3 Ground Motion Measurement Devices 1057
C.3.1 Ground Motion Levels 1057
C.3.2 Accelerometers 1057
C.3.3 Geophones 1057
C.3.4 Sensor Internal Noise Sources 1060
C.3.5 *Feedback Seismographs* 1061
C.3.6 Other New Sensors 1063
C.4 SLAC Ground Motion Measurements and Analysis 1063
C.4.1 STS-2 Measurements and Data Analysis 1064
C.4.2 Measurement Resolution 1064
C.4.3 Spatial Properties of the Vertical Ground Motion 1066
C.4.4 Application to the NLC 1069

C.1 Introduction

One of the major challenges for the NLC will be to keep the beams stable pulse-to-pulse so the collision luminosity is not degraded. Because the beam sizes are so small, motions of the quadrupole magnets on the nanometer scale can cause significant orbit changes. Hence, the quadrupole vibration caused by the ambient ground motion is a concern as is vibration due to man-made sources, such as cooling systems.

The sensitivity to vibration depends on both the time scale and correlation of the motion. Motion slower than about 0.1 Hz will be heavily suppressed by the trajectory feedback systems in the NLC and so it is less of a concern. Motion that is random will generally be much more of a problem than motion which is correlated. For example, the tolerance on the vertical rms quadrupole motion in the linacs is 6 nm if the motion is random quad-to-quad, but orders of magnitude larger if it is correlated over distances longer than the betatron oscillation lengths of the beams.

At frequencies above 0.1 Hz, the ambient rms ground motion is generally a few hundred nm or more, but the motion is correlated over long distances. The exact nature of this correlation has a big effect on whether the resulting beam motion in an accelerator will be tolerable or not. In the past year at SLAC, we have developed both the theoretical understanding and the measurement capability to address this question for the NLC. In this Appendix, we present the results of this learning process, including measurements of both the spatial and spectral characteristics of vertical ground motion in the SLAC linac tunnel.

To begin, we develop the theoretical framework that is needed to interpret ground motion measurements and to predict its effect on beam motion. Next, we describe some of the instruments that are used to measure ground motion, including their resolutions. In the last section, we discuss the measurements that were made at SLAC, and present a model to describe the data that is based on the dynamics of wave motion in the ground. We then derive a general equation that will govern the ground-motion effect on luminosity in the NLC in the framework of this model. However, the implications for the NLC are described separately for each region of the machine in the chapters that discuss these regions.

C.2 Theory

In this section we derive a general equation that describes the effect of ground motion on the separation of the beams at the interaction point (IP) of a collider. Specifically, we will show that the average square vertical displacement of the left-incoming from the right-incoming beam centroid can be written as an integral

$$\left\langle \Delta y_{IP}^2 \right\rangle_t = \int_0^\infty P(\omega,k)G(k)F(\omega)\frac{dk}{2\pi}\frac{d\omega}{2\pi} \tag{C.1}$$

where $<>_t$ indicates the average over time, $P(\omega,k)$ is a function describing the relevant aspects of the ground motion, $G(k)$ is a function describing the response of the lattice, and $F(\omega)$ describes the behavior of the feedback system. The definition, meaning, and properties of each of these functions are described in the next sections.

C.2.1 Ground Motion Spectra and Related Functions

The ground displacement along the accelerator may be described by a function of two variables $y(t, s)$, where s denotes position along the beam line and t denotes time. To minimize any possible ambiguity, imagine that at a point to be designated as $s = 0$, a horizontal line is constructed parallel to the machine orientation and perpendicular to a radius through the origin of the earth. The variable s will indicate the distance from the origin measured along this straight line. For simplicity we assume the radius of the earth at the origin of this line is chosen so the average displacement $< y(t, 0) >_t = 0$, and that $< y(t, s) >_t = 0$ for all s.

The Time Power Spectrum $P(\omega)$

A simple first question one might ask is "How large a ground displacement, at the point s, am I likely to experience in time τ starting from some unknown time t?" The answer can be given in a statistical sense as

$$\left\langle \Delta y(\tau; s)^2 \right\rangle = \left\langle [y(t + \tau, s) - y(t, s)]^2 \right\rangle \quad . \tag{C.2}$$

Assuming uniformity along the horizontal line, this should be independent of s, but for a local source this could conceivably be s-dependent. Squaring and expanding Eq. C.2:

$$\left\langle \Delta y(\tau; s)^2 \right\rangle = 2 \left[\left\langle y(t, s)^2 \right\rangle_t - \left\langle y(t + \tau, s) y(t, s) \right\rangle_t \right] \quad . \tag{C.3}$$

The last term will be an important function for us, and we will define $P(\omega; s)$ as its Fourier transform, namely:

$$\left\langle y(t + \tau, s) y(t, s) \right\rangle_t = \int_0^\infty P(\omega; s) \cos(\omega\tau) \frac{d\omega}{2\pi} \quad . \tag{C.4}$$

The $\cos(\omega\tau)$ is appropriate because the left-hand side is unchanged by the substitution of $-\tau$ for τ (followed by the substitution of $t = t' + \tau$ and noting that the average over t' is the same as the average over t). Next we will argue that $P(\omega) \geq 0$ for all ω.

Suppose that $y(t, s)$ has been Fourier decomposed. The time average of two terms with unequal frequencies will be zero, leaving only the sum of Fourier terms times themselves shifted by the time τ. That time average gives a $\cos(\omega\tau)$ times the amplitude of the term squared. The $P(\omega; s)$ must therefore be a constant times the square of a Fourier amplitude, and hence always positive.

Setting $\tau = 0$, we have the relationship

$$\left\langle y^2(t, s) \right\rangle_t = \int_0^\infty P(\omega; s) \frac{d\omega}{2\pi} \quad . \tag{C.5}$$

The quantity on the left-hand side is clearly positive for all t. According to the above paragraph, we can identify $P(\omega; s)$ as the contribution at frequency ω to the time averaged square of $y(t, s)$.

We will now make these statements more rigorous. Assume that we are averaging over a large time interval $[0, T]$. In this time interval, the function $y(t, s)$, for any fixed t, can, by imposing periodic boundary

conditions, be expanded in a Fourier series

$$\begin{aligned} y(t,s) = & \sum_{n=1}^{\infty} \left[a_n^T(s)\sin(\omega_n t) + b_n^T(s)\cos(\omega_n t)\right] \\ = & \sum_{n=1}^{\infty} c_n^T(s)\sin\left(\omega_n t + \psi_n^T(s)\right) \end{aligned} \tag{C.6}$$

$$\text{where} \quad \omega_n = \frac{2\pi n}{T} \quad .$$

Forming the product $y(t+\tau,s)y(t,s)$ and taking the time average (and using the orthogonality relationship of the sin and cos functions) we get

$$\langle y(t+\tau,s)y(t,s)\rangle = \sum_{n=1}^{\infty} \frac{1}{2} c_n^T(s)^2 \cos(\omega_n \tau) \quad . \tag{C.7}$$

As T gets large, this can be converted to an integral by using the relationship $dn/d\omega_n = T/2\pi$:

$$\langle y(s,t+\tau)y(s,t)\rangle = \int_0^{\infty} \frac{1}{2} \lim_{T\to\infty} \left[c_\omega^T(s)^2 T\right] \cos(\omega\tau)\frac{d\omega}{2\pi} \tag{C.8}$$

and we can identify

$$P(\omega,s) \equiv \frac{1}{2} \lim_{T\to\infty} \left[c_\omega^T(s)^2 T\right] \tag{C.9}$$

which is clearly a positive function for all ω. Since the left-hand side of Eq. C.8 is well defined, the product $c_\omega^T(s)^2 T$ must be independent of T for large enough T. This is equivalent to saying, that as T doubles, and the number of states in a given frequency range doubles, it is the energy of the states in that region that remains constant. For this to happen, c_n, the amplitude of a state in the shorter time interval, must decrease as $\sqrt{2}$ in the doubled time interval. (The sum of the squares of two modes equals the square of the previously single mode.)

Thus $P(\omega;s)$ can be interpreted as the power spectrum of $y(t,s)$ at s. Since we will be concerned primarily with seismic sources in this appendix, we will drop the s argument. We note that this s independence implies that the $c_n(s)$ is also independent of s. Only the phase, $\psi_n(s)$, of the motion at this frequency can retain an s dependence.

Since the left-hand side of Eq. C.5 is well-defined, and by experience known to be finite, the integral of the right-hand side should exist. However the left-hand side is quite large, since the daily changes in the earth's radius at any given location are about ± 40 cm. In other words, though $P(\omega;s)$ is integrable, it is expected to be huge for small ω. We expect that these low-frequency motions are coherent, that is to say, they are predominantly composed of long spatial wavelength motions. For these motions, the beam lines move as a whole, and we are led intuitively to believe they will not be a problem. Also there are beam-based feedback systems that receive information on beam position at 180 or 120 Hz, that will suppress the influence of low-frequency motions.

However at 0.15 Hz there is a micro-seismic peak due to ocean waves, with an amplitude as large as a micrometer, much larger than the beam spot size at the IP. To determine the impact of such motions, we must determine the wavelength decomposition of the ground motion as well as the frequency decomposition.

The Wavelength Power Spectrum $P(k)$

Proceeding as in the previous section, one might ask: "At any given moment how much are the vertical placements at the point s and $s+\ell$ likely to differ?" Again, the answer can be given in a statistical sense as

$$\left\langle \Delta y(\ell;t)^2 \right\rangle = \left\langle [y(t,s+\ell) - y(t,s)]^2 \right\rangle_s \quad . \tag{C.10}$$

Proceeding exactly as in the preceding paragraph would lead to

$$\langle y(t,s+\ell)y(t,s)\rangle_s = \int_0^\infty P(k;t)\cos(k\ell)\frac{dk}{2\pi} \tag{C.11}$$

where

$$P(k;t) \equiv \frac{1}{2} \lim_{L\to\infty} [c_k^L(t)^2 L] \tag{C.12}$$

when we analyze the motion at a fixed time t along an interval in s of length L, imposing periodic boundary conditions. $P(k;t)$ gives the wavelength power decomposition of $y(t,s)$ at time t. If the distance s is long enough, and external conditions are not changing with t, we would expect this quantity to be independent of t, at least for short intervals of time, and we drop the explicit time-dependence, $P(k)$.

As a practical matter, when measurements are made to determine $P(k)$, the average over all s is not taken. Two seismometers are placed at some s separated by a distance ℓ, and it is assumed that the average over s is equal to the average over t, for some time interval T. Thus in practice we assume

$$\langle y(t,s+\ell)y(t,s)\rangle_s = \langle y(t,s+\ell)y(t,s)\rangle_t \quad . \tag{C.13}$$

The 2D Power Spectrum $P(\omega,k)$

Because the feedback systems are operating in time (getting a piece of information at 180 or 120 Hz) the frequency of any disturbance will be of great importance. On the other hand the response of the lattice will depend on the wavelength of the disturbance, so the wavelength decomposition is of great importance. One is lead to the undeniable conclusion that one must have both, namely at each frequency one needs to know the wavelength decomposition of the motion at that frequency. One must find the answer to question, "Given two points separated by a distance ℓ how much does their relative vertical displacement change in time τ?" This quantity is

$$\left\langle \Delta y(\tau,\ell)^2 \right\rangle = \left\langle [\{y(t+\tau,s+\ell) - y(t+\tau,s)\} - \{y(t,s+\ell) - y(t,s)\}]^2 \right\rangle_{t,s} \quad . \tag{C.14}$$

Squaring the expression within the square bracket, there are two new terms: $< y(t+\tau,s+\ell)y(t,s) >_{t,s}$ and $< y(t+\tau,s)y(t,s+\ell) >_{t,s}$. The sum of these is a symmetric function of τ and ℓ, and as above, we can define its Fourier transform through

$$\begin{aligned} &\tfrac{1}{2}\left[\langle y(t+\tau,s+\ell)y(t,s)\rangle_{t,s} + \langle y(t+\tau,s)y(t,s+\ell)\rangle_{t,s}\right] \\ &\quad = \int_0^\infty \int_0^\infty P(\omega,k)\cos(\omega\tau)\cos(k,\ell)\tfrac{dk}{2\pi}\tfrac{d\omega}{2\pi} \quad . \end{aligned} \tag{C.15}$$

As with P(ω) and P(k), P(ω,k) ≥ 0 for all ω and k, and it is also a power spectrum.

Assume that we are averaging over a large time interval $[0,T]$ and space interval $[0,L]$. In this time and space interval the function $y(t,s)$ can be expanded in a double Fourier series

$$\begin{aligned} y(s,t) = \; & \textstyle\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[a_{n,m}^{T,L}\sin(\omega_n t)\sin(k_m s) + b_{n,m}^{T,L}\sin(\omega_n t)\cos(k_m s) + \right. \\ & \left. d_{n,m}^{T,L}\cos(\omega_n t)\sin(k_m s) + e_{n,m}^{T,L}\cos(\omega_n t)\cos(k_m s)\right] \qquad \text{(C.16)} \\ & \text{where } \omega_n = \tfrac{2\pi n}{T} \text{ and } k_n = \tfrac{2\pi m}{L} \quad . \end{aligned}$$

Forming the products $y(t+\tau, s+\ell)y(t,s)$ and $y(t+\tau,s)y(t,s+\ell)$, and taking the time average (and using the orthogonality relationship of the sin and cos functions) we get

$$\begin{aligned} & \tfrac{1}{2}\left[\langle y(t+\tau,s+\ell)y(t,s)\rangle_{t,s} + \langle y(t+\tau,s)y(t,s+\ell)\rangle_{t,s}\right] \\ & \qquad = \textstyle\sum_{n,m=1}^{\infty}\tfrac{1}{4}c_{n,m}^{T,L^2}\cos(\omega_n\tau)\cos(k_m\ell) \qquad \text{(C.17)} \\ & \text{where } c_{n,m}^{T,L^2} = a_{n,m}^{T,L^2} + b_{n,m}^{T,L^2} + d_{n,m}^{T,L^2} + e_{n,m}^{T,L^2} \quad . \end{aligned}$$

As T and L get large this can be converted to an integral by using the relationships $dn/d\omega_n = T/2\pi$ and $dm/dk_m = L/2\pi$ and we can identify

$$P(\omega,k) = \frac{1}{4}\lim_{T,L\to\infty}\left[c_{\omega,k}^{T,L^2}TL\right] \quad . \qquad \text{(C.18)}$$

By taking $\ell = 0$ or $\tau = 0$ in the defining expression for $P(\omega,k)$, we have the following relationships among the power spectrums so far defined:

$$P(k) = \int_0^{\infty} P(\omega,k)\,\frac{d\omega}{2\pi} \quad \text{and} \quad P(\omega) = \int_0^{\infty} P(\omega,k)\,\frac{dk}{2\pi} \quad . \qquad \text{(C.19)}$$

The Measure $\mu(\omega,k)$

The spectrum $P(\omega)$ is observed to vary according to weather conditions and can change dramatically if there is an earthquake. Excluding earthquakes, for which we have little data, the spectrum $P(\omega)$ may change by factors of up to ten, however for any given ω the distribution of wavelengths at that frequency appears to be quite constant. This has led us to define a function $\mu(\omega,k)$:

$$\mu(\omega,k) = \frac{P(\omega,k)}{P(\omega)} \quad . \qquad \text{(C.20)}$$

The function $\mu(\omega,k)$ is a probability measure in the variable k, since $\int_0^{\infty}\mu(\omega,k)\frac{dk}{2\pi} = 1$ for any ω. In words, $\mu(\omega,k)dk/2\pi$ gives the probability that a ground disturbance of frequency ω will contain wavelengths in the interval dk about wave number $k = 2\pi/\lambda$.

If at frequency ω one has simply a wave of wavelength $\lambda(\omega)$ traveling to the right or left along the direction s at velocity $v(\omega) = f\lambda = \omega/k$, then

$$\mu(\omega,k) = 2\pi\delta(k - k_0(\omega)) \quad \text{where} \quad k_0(\omega) = \frac{2\pi}{\lambda(\omega)} = \frac{\omega}{v(\omega)} \qquad \text{(C.21)}$$

and $\delta(k - k_0)$ denotes the Dirac delta function.

If at frequency ω one has only waves of wavelength $\lambda(\omega)$ traveling back and forth at some distinct angle θ to the direction s, then the wavelength of the disturbance along s will be $\lambda_s = \lambda/\cos\theta$ and

$$\mu(\omega, k) = 2\pi\delta(k - k_0(\omega)\cos\theta) \quad \text{where} \quad k_0(\omega) = \frac{2\pi}{\lambda(\omega)} = \frac{\omega}{v(\omega)} \quad . \tag{C.22}$$

If at frequency ω one has only waves of wavelength $\lambda(\omega)$ traveling at all angles θ to the direction s, then there will be a distribution of wavelengths along s. There will be none with $k > k_0$. Since for any given angle the wave number along s is $k = k_0\cos\theta$, and the distribution is presumed uniform in $0 < \theta < \pi/2$, the fraction at any θ in interval $d\theta$ will be

$$\frac{2}{\pi}d\theta = \frac{2}{\pi}dk\frac{d\theta}{dk} = \frac{2}{\pi}\frac{dk}{k_0\sin\theta} = \frac{2}{\pi}\frac{dk}{\sqrt{k_0^2 - k^2}} \tag{C.23}$$

and we have

$$\mu(\omega, k) = \frac{4}{\sqrt{k_0^2(\omega) - k^2}} \quad \text{for} \quad k < k_0 \quad . \tag{C.24}$$

It is possible that at any frequency there could be contributions from several different types of waves, for example surface waves, compression waves or Raleigh waves. And there is the possibility for motion due to stochastic processes that would have an entirely different kind of behavior. For example, much discussed in the literature is a relationship called the ATL law [Baklakov 1991]. If this type of motion is present then we will show later that

$$\mu(\omega, k) = \frac{A(\omega)}{k^2} \quad \text{as} \quad k \to \infty \quad . \tag{C.25}$$

The main task of ground motion measurements will be the determination of $\mu(\omega, k)$.

The Ratio $R(\omega, \ell)$

$P(\omega)$ gives the spectral decomposition of the rms motion of a single point. According to our definitions above

$$\langle y^2 \rangle = \int P(\omega)\frac{d\omega}{2\pi} \quad . \tag{C.26}$$

The power spectrum of the difference measurement of two points as indicated in Eq. C.10 can be written

$$\begin{aligned} \langle \Delta y(\ell)^2 \rangle &= 2\int P(k)\,[1 - \cos(k\ell)]\,\tfrac{dk}{2\pi} \\ &= 2\int\int P(\omega, k)\,[1 - \cos(k\ell)]\,\tfrac{dk}{2\pi}\tfrac{d\omega}{2\pi} \\ &= 2\int P(\omega)\int \mu(\omega, k)\,[1 - \cos(k\ell)]\,\tfrac{dk}{2\pi}\tfrac{d\omega}{2\pi} \\ &\equiv 2\int P(\omega)R(\omega, \ell)\tfrac{d\omega}{2\pi} \end{aligned} \tag{C.27}$$

from which we can conclude that the power spectrum for the difference of two points separated by distance ℓ is related to the power spectrum of single point by the ratio $2\,R(\omega, \ell)$ with

$$R(\omega, \ell) = \int \mu(\omega, k)\,[1 - \cos(k\ell)]\,\frac{dk}{2\pi} \quad . \tag{C.28}$$

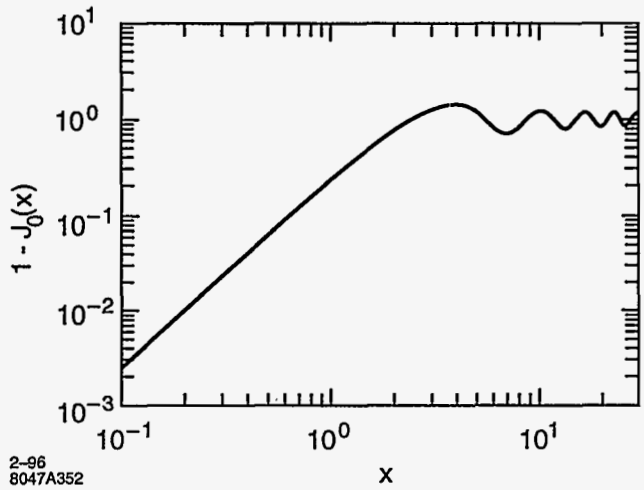


Figure C-1. *A log-log plot of the function $y = 1 - J_0(x)$. Note the asymptotic x^2 behavior for small x and the position $x = 2.4$ of the intercept at which $R = 1$ ($J_0 = 0$). The first maximum is about 1.44 at $x = 3.8$.*

We are guaranteed that $R(\omega, \ell) \leq 2$.

$R(\omega, \ell)$ can be determined directly in experiments, so we must try to extract the measure $\mu(\omega, k)$ given the ratio $R(\omega, \ell)$ for many ℓ. If one had data for many ℓ one could simply take a Fourier transform to find $\mu(\omega, k)$. Since we only have data at a small number of ℓ values, we will proceed differently by concentrating on the asymptotic behavior of R.

The Ratio $R(\omega, \ell)$ as $\ell \to 0$

If the coherence of ground motion is to be helpful in suppressing its effect on beam motion, then we will be interested to know when $R(\omega, \ell)$ is small. This typically occurs as ℓ becomes small, so the asymptotic behavior for small ℓ is of special importance.

For the measure given by Eq. C.24, corresponding to waves of a single wavelength coming from all directions, then the integral of Eq. C.28 gives

$$R(\omega, \ell) = 1 - J_0\left(k_0(\omega)\ell\right) \tag{C.29}$$

where $J_0(k_0\ell)$ is the zeroth-order Bessel function. In this case, $1 - J_0(k_0 l) \to (k_0\ell/2)^2$ as $\ell \to 0$.

This asymptotic quadratic behavior is quite general for $R(\omega, \ell)$ as can be seen by looking at Eq. C.28. Suppose the measure μ has the property that

$$\overline{k^2(\omega)} \equiv \int k^2 \mu(\omega, k) \frac{dk}{2\pi} < \infty \quad . \tag{C.30}$$

Then $1 - \cos(k\ell)$ in the definition of R can be expanded in a power series with the result that

$$R(\omega, \ell) \to \frac{1}{2}\overline{k^2(\omega)}\ell^2 \quad \text{as} \quad \ell \to 0 \quad . \tag{C.31}$$

This predicts a clear l dependence for very general circumstances, and can be a strong consistency check on the data.

There can be other kinds of asymptotic behavior as $l \to 0$ if the second moment of μ does not exist. One such case of interest is when $\mu(\omega, k)$ has the form $A(\omega)/k^2$. In this case the condition of Eq. C.30 is not fulfilled, and the conclusion of Eq. C.31 is not valid. We will see in the next section that then $R(\omega, \ell)$ will be proportional to ℓ as $\ell \to 0$, which we refer to as the ATL law behavior.

Finding the ATL Law in Ground Motion Measurements

If at a given frequency ω there is a small ATL law component together with a dominant wave motion, then its signature will be that as ℓ becomes small one first observes $R(\omega, \ell) \to \alpha\ell^2$, and then this behavior shifts to $R(\omega, \ell) \to \beta\ell$. In other words, on a log-log plot of $R(\omega, \ell)$ as a function of ℓ one should observe a change from a slope of 2 to a slope of 1. If only the ℓ^2 behavior is seen, one can find an upper bound on the strength of the ATL coefficient. Let us suppose that this behavior is seen and for small ℓ. Let us take

$$R(\omega, \ell) = \alpha\ell^2 + \beta\ell \quad . \tag{C.32}$$

The intersection of the two straight-line behaviors will occur at $\ell = \ell_x$ with $\ell_x = \beta/\alpha$. The value of R at this intersection will be $R(\omega, \ell) = 2\beta^2/\alpha$, twice the value of the extension of either straight line. If no departure from an ℓ^2 behavior can be seen, we can conclude that $\ell_x (= \beta/\alpha) < \ell_{min}$, where ℓ_{min} is the smallest ℓ for which data is available. Hence $\beta < \alpha\ell_{min}$. We will use this relationship to establish upper bounds on the magnitude of the ATL law coefficient. Now we show how this asymptotic behavior is related to the measure $\mu(\omega, k)$.

Suppose that $\mu(\omega, k)$ has a $1/k^2$ tail. This is allowed because dk/k^2 is integrable as $k \to \infty$. Let us decompose the measure into two parts, one which contains the $1/k^2$ tail, and another which possesses a second moment. We write $\mu = \mu_R + \mu_T$ and take

$$\mu_T(\omega, k) = \frac{a(\omega)}{k^2}\left[1 - e^{-\frac{k^2}{2k_1^2}}\right] \tag{C.33}$$

where k_1 is a cut-off parameter.

The measure μ_T contains the $1/k^2$ tail and is integrable at both large and small k. The integral of the measure (integrate by parts) gives the fraction of the measure which is in the tail:

$$\mu_T(\omega) \equiv \int_0^\infty \mu_T(\omega, k)\frac{dk}{2\pi} = \frac{1}{2\sqrt{2\pi}}\frac{a(\omega)}{k_1(\omega)} \quad . \tag{C.34}$$

$\mu_R = \mu - \mu_T$ is supposed to have a finite second moment. Hence for small ℓ

$$\begin{aligned} R(\omega, \ell) &= \int(\mu_R + \mu_T)\left[1 - \cos(k\ell)\right]\tfrac{dk}{2\pi} \\ &\to \tfrac{\ell^2}{2}\int \mu_R k^2 \tfrac{dk}{2\pi} + \int \mu_T \left[1 - \cos(k\ell)\right]\tfrac{dk}{2\pi} \\ &= \tfrac{\ell^2}{2}\left\langle k^2 \right\rangle_R + \tfrac{a(\omega)}{4}\ell - \tfrac{\ell^2}{4}\tfrac{a(\omega)k_1}{\sqrt{2\pi}} \\ &= \tfrac{\ell^2}{2}\left[\left\langle k^2 \right\rangle_R - \mu_T(\omega)k_1^2\right] + \tfrac{a(\omega)}{4}\ell \quad . \end{aligned} \tag{C.35}$$

The coefficient A in the ATL law results in a spectral coefficient $4A/(k^2\omega^2)$. Hence the relation between A and $a(\omega)$ is given by

$$A = \omega^2 P(\omega)\frac{a(\omega)}{4} \quad . \tag{C.36}$$

Typically A is thought to be in the range of $10^{-4} < A < 10^{-6}\,\mu\mathrm{m}^2/\mathrm{m/s}$. If there is no evidence of linear ℓ behavior, we can place a limit on A of $\omega^2 P(\omega)\beta < \omega^2 P(\omega)\alpha\ell_{\min} = \omega^2 P(\omega)\alpha\ell_{\min}$ where α is the coefficient of the ℓ^2 behavior and $\ell_{\min}$ is the smallest ℓ in the data set.

We note in closing that mathematically other behaviors at small ℓ are allowed, of the form ℓ^ϵ with $0 < \epsilon \leq 2$, corresponding to tail behaviors of $1/k^{1+\epsilon}$.

C.2.2 The Lattice Response Function $G(k)$

Sum Rules for Optical Functions

We will suppose that we have a lattice of elements with the quadrupole strength, $\kappa(s)$, given as a function of s. This function is defined by

$$\kappa(s) = \frac{B_T(s)}{a(s)(B\rho)} \tag{C.37}$$

where $B_T(s)$ is the quadrupole pole-tip field at s, and $a(s)$ is the radius of the pole tip. $B\rho = p/q$ is the particle momentum divided by the charge.

We will assume that there is some ideal line along which the magnets are arranged, and the displacement is given as a function of time by $y(t,s)$ with $< y(t,s) >_t = 0$. Then the displacement of the beam at the end of the beam line, which we will denote by $s = 0$, is

$$y_B(t) = \int_0^L \kappa(s)y(t,s)R_{34}(s)ds + R_{33}(L)y(t,L) \quad . \tag{C.38}$$

The assumption here is that the beam enters the beam line at a displacement equal to the ground level. It will be the function of the upstream beam line to carry out this task. Any failure to do this can be accounted for by calculating the deviation due to the upstream beam line and multiplying by $R_{33}(L)$. We will define a quantity $\Delta y_B(t)$ which is the difference $y_B(t) - y(t,0)$. Then

$$\Delta y_B(t) = \int_0^L \kappa(s)y(t,s)R_{34}(s)ds + R_{33}(L)y(t,L) - y(t,0) \quad . \tag{C.39}$$

If one translates the beam line by inserting a $y(t,s)$ which is independent of s, then $\Delta y_B(t)$ should be zero. According to Eq. C.39 this requires that

$$0 = \int_0^L \kappa(s)R_{34}(s)ds + R_{33}(L) - 1 \quad . \tag{C.40}$$

This is true, because $R_{34}(s)$ from $s \to 0$ is the same as the R_{34} from $0 \to s$. The latter is a sine-like ray coming from the origin, which we will denote by $\tilde{y}$ and has $\tilde{y}'(0) = -1$. Because of the equations of motion

$\kappa(s)\tilde{y}(s) = \tilde{y}''(s)$, the integral in Eq. C.40 is $\tilde{y}'(L) - \tilde{y}'(0)$. The derivative of the R_{34} function coming from the origin is just -1 times the R_{44} function coming from the origin, which equals the -1 times the R_{33} function going from s to the origin. Hence Eq. C.40 is found to be valid, as it must.

If one inserts $y(t,s) = \alpha s$ into Eq. C.39, then $\Delta y_B(t)$ should again be zero if L is $n\pi$ from the origin. (If the phase change from 0 to L equals $n\pi$, the slope of the incoming beam does not affect the position at the IP). This implies that

$$0 = \int_0^L s\kappa(s)R_{34}(s)ds + LR_{33}(L) \quad . \tag{C.41}$$

Inserting $\tilde{y}(s)$ for R_{34}, substituting $\kappa(s)\tilde{y}(s) = \tilde{y}''(s)$ as above, and integrating by parts, and using the fact that $\tilde{y}(L) = 0$, one can confirm Eq. C.41.

These intuitively valid relationships would imply that it is not the magnitude of a displacement, nor even its slope that will cause beams not to collide, but it is the curvature of the earth's displacement that is important.

We have gone into some detail to describe and confirm these optical rules, because often approximations used for optical functions do not satisfy these conditions, and numerically incorrect results are obtained. A high degree of suppression of motion will only be obtained numerically when these sum rules are obeyed.

Paired Lattices

For lattices that have the property that for every $-I$ section there is an identical $-I$ section adjacent to it, which we designate as a "paired lattice," we have the result that if the displacement function has constant curvature the effects of the two adjacent sections will cancel one another. This leads to the conclusion that for paired lattices it is not the displacement, the slope, or the curvature, but the derivative of the curvature that causes a displacement of the beam.

Definition of $G(k)$

We will first write the beam displacement from the ground motion at the IP as a sum,

$$\Delta y_B(t) = \sum \mu_i y(t, s_i) \tag{C.42}$$

where the μ_i are dimensionless numbers equal to $\kappa(s_i)R_{34}(s_i)\Delta s_i$ or $R_{33}(L)$ or -1. Now we square this to get

$$\begin{aligned} < \Delta y_B(t)^2 > &= \textstyle\sum_{i,j} \mu_i\mu_j < y(t,s_i)y(t,s_j) >_t \\ &= \textstyle\sum_{i,j} \mu_i\mu_j \int_0^\infty P(k)\cos\left[k(s_i-s_j)\right]\frac{dk}{2\pi} \\ &= \textstyle\int_0^\infty P(k)\frac{dk}{2\pi}\sum_{i,j}\mu_i\mu_j\cos\left[k(s_i-s_j)\right] \\ &= \textstyle\int_0^\infty P(k)\frac{dk}{2\pi}\left[\left(\sum_i \mu_i\cos ks_i\right)^2+\left(\sum_i \mu_i\sin ks_i\right)^2\right] \\ &\equiv \textstyle\int_0^\infty P(k)G(k)\frac{dk}{2\pi} \quad . \end{aligned} \tag{C.43}$$

The lattice properties are completely contained in the sum, and therefore we have defined a lattice response function $G(k)$ by

$$G(k) = \left(\sum_i \mu_i \cos ks_i\right)^2 + \left(\sum_i \mu_i \sin ks_i\right)^2 \quad . \tag{C.44}$$

It is precisely Δy_B^2 for a cosine wave plus Δy_B^2 for a sine wave.

We note that in the special case were we are looking at the displacement of the two beams from one another at the IP, then if we assume that the NLC is symmetric, the cosine wave gives no displacement of the two beams, and the sine wave effect is doubled. Hence, when considering the effect on luminosity due to ground motion in both halves of a symmetric NLC, we have

$$G(k) = 4\left(\sum_i \mu_i \sin ks_i\right)^2 \tag{C.45}$$

where the sum is only over one half of the machine.

Behavior of $G(k)$ as $k \to 0$

As $k \to 0$, $\cos(ks_i) \to 1$. As shown above the constant term can give no net displacement. The next term is $1/2k^2s_i^2$. This term has a constant curvature, so for a paired lattice it will also be zero. The first term surviving from the cosine sum squared will go as k^8 in this case.

As $k \to 0$, $\sin(ks_i) \to ks_i$. As shown above the constant slope term also gives no net displacement. The first term surviving from the sine sum squared will go as k^6.

Some examples of $G(k)$ for various lattice sections are shown in Figures C-2, C-3, and C-4.

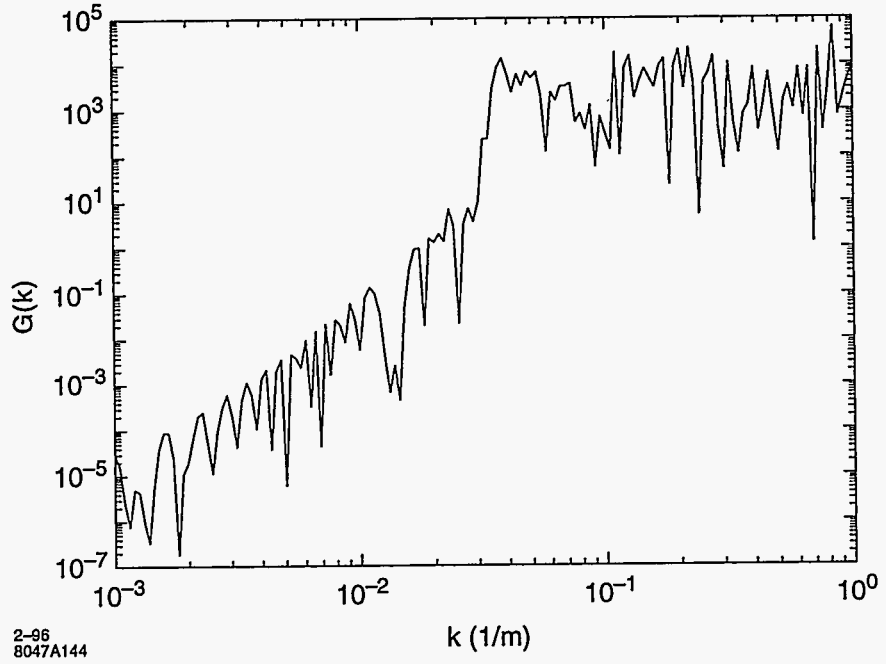


Figure C-2. *NLC-IIb linac lattice response function assuming symmetric linacs (i.e., Eq. C.45) that are separated by 10 km.*

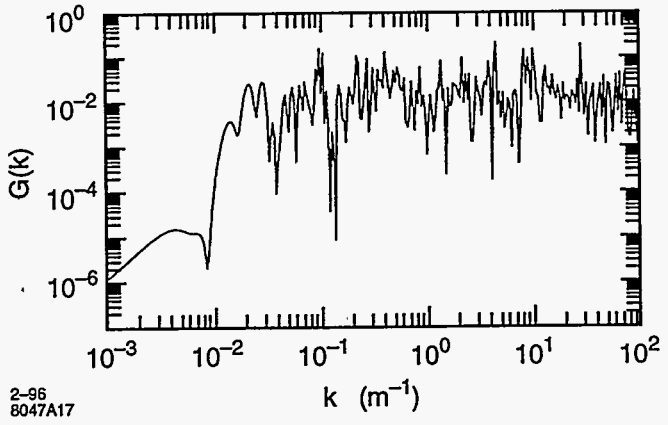


Figure C-3. *Collimation system lattice response function.*

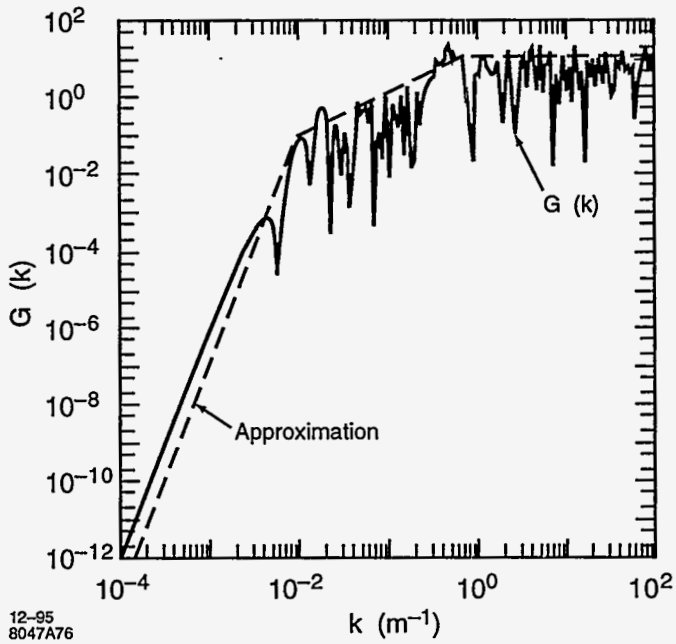


Figure C-4. *Final-focus lattice response function.*

C.2.3 Combining Spatial and Frequency Dependence

$F(\omega)$

The particulars of a feedback response function will vary according to the exact feedback algorithm employed. In each case however there is a function $F(\omega)$ which will give the effect of the feedback for each input frequencies. As a result Eq. C.43 can be generalized, in the presence of feedback, to be

$$< \Delta y_B(t)^2 >_t = \int P(\omega, k) G(k) F(\omega) \frac{dk}{2\pi} \frac{d\omega}{2\pi} \quad . \tag{C.46}$$

$G_\mu(\omega)$

Once the measure $\mu(\omega, k)$ has been determined, $P(\omega, k)$ in Eq. C.46 can be replaced by $P(\omega, k) = P(\omega)\mu(\omega, k)$ and the integral over k in Eq. C.46 can be carried out. The result of doing the k integral we will denote by

$$G_\mu(\omega) \equiv \int_0^\infty \mu(\omega, k) G(k) \frac{dk}{2\pi} \quad . \tag{C.47}$$

This gives the response of the lattice to the combination of wavelengths present in the ground motion at the frequency ω. Doing this integral first, Eq. C.46 becomes

$$\left\langle \Delta y_B(t)^2 \right\rangle_t = \int_0^\infty P(\omega) G_\mu(\omega) F(\omega) \frac{d\omega}{2\pi} \quad . \tag{C.48}$$

C.3 Ground Motion Measurement Devices

C.3.1 Ground Motion Levels

Seismic ground noise varies greatly with time and place. Down to about 1 Hz, a large variety of instrument types can resolve ground noise at noisy sites where cultural noise dominates the spectrum. At the quietest sites though, only the most sophisticated instruments have broad enough bandwidth and a low enough noise floor to accurately measure ground motion from 0.1 Hz to 100 Hz. An informal ground noise standard has been developed at the USGS for bench marking seismic instruments. The power spectral densities for ground noise at some of the most quiet sites in North America have been averaged into a low-noise model which is now included in the instrument noise-floor plots of the most sensitive commercial instruments. Figure C-5 shows the acceleration and displacement amplitude spectra for this low-noise model.

These plots show that at 1 Hz the seismic instrument must detect nano 'g's of acceleration per $\sqrt{\text{Hz}}$ which corresponds to 1/10ths of nanometers of ground displacement per $\sqrt{\text{Hz}}$. Ground motion at most accelerator sites is 10 to 100 times larger than the low-noise model for frequencies $\geq$1 Hz where cultural noise dominates.

C.3.2 Accelerometers

The simplest, least expensive detectors are piezoelectric crystals coupled to sensitive charge amplifiers. These instruments are high-sensitivity versions of industrial machine vibration detectors. Their frequency response is nearly flat up to the first resonance of the crystal (100 Hz–1000 Hz). Their noise floor is dominated by electrical noise in the high-gain pre-amp. Figure C-6 shows information from the data sheet for the Wilcoxon Model 731 piezoelectric seismic accelerometer.

At 1 Hz, noise in the Model 731 is equivalent to an acceleration spectral density of 50 ng/$\sqrt{\text{Hz}}$ or a displacement density of 13 nm/$\sqrt{\text{Hz}}$. For noisy sites this detector might show ground noise signals above self-noise at frequencies above 5 Hz, but it does not reach the low-noise model noise floor at any frequency. Its most useful application is the direct measurement of local sources of mechanical noise such as cooling system flow turbulence and pulsed power conversion equipment.

C.3.3 Geophones

Moving coil velocity meters or geophones transduce earth motion by electromagnetic induction. Figure C-7 shows a simple geophone model. They contain a permanent magnet and a moving coil attached to a spring suspended mass. A voltage is produced across the coil that equals the transduction (G) times the relative velocity ($\dot{x} - \dot{x}_g$) of the mass measured with respect to the permanent magnet which is fixed in the housing of the transducer.

A typical geophone suitable for microseismic ground motion studies is the L4C Geophone made by Mark Products. The 1-kg-proof mass and suspension system has a 1-Hz resonant frequency. In addition to its intrinsic internal mechanical damping, this instrument is usually brought to a .707 critically damped response by shunting the output by a damping resistor $R_s = 8905\Omega$. The Laplace transform transfer function of output

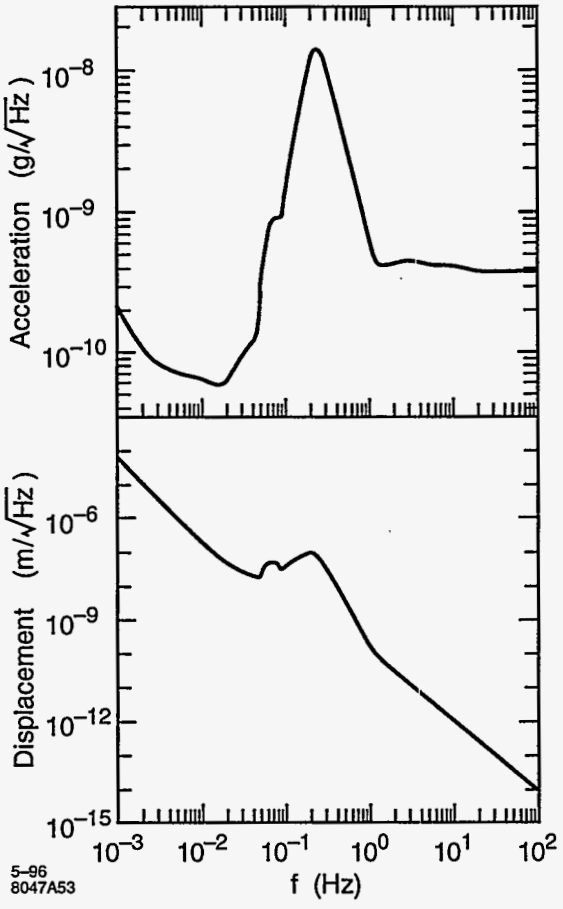


Figure C-5. *USGS low-noise model.*

voltage *vs.* input ground velocity is:

$$\frac{V_{\rm out}(s)}{\dot{x}_g(s)} = \frac{R_s}{R_c+R_s} G \left(\frac{s^2}{s^2 + 2\omega_0 \zeta s + \omega_0^2} \right) \quad \text{where} \quad 2\omega_0\zeta = \left(b + \frac{G^2}{R_c+R_s} \right) / M \quad . \tag{C.49}$$

The amplitude and phase of the complex function $V_{\rm out}(i\omega)/\dot{x}_g(i\omega)$ are plotted in Figure C-8. Above resonance ($s >> i\omega_0$), the transducer conversion factor from Table C-1 for a damped L4C geophone is:

$$\frac{V_{\rm out}}{\dot{x}_g} = \frac{8905\ \Omega}{5500\Omega + 8905\Omega} (276.4\,{\rm V\,s/m}) = 170.9\,{\rm V\,s/m} \quad . \tag{C.50}$$

One-nm rms motion at 10 Hz, for example, would produce a 11-μV signal at the geophone's output if a 8905-Ω external damping resistor is used.

Specifications

Sensitivity, ±10%, 25° C	10 V/g
Accleration Range	0.5 g peak
Amplitude Nonlinearity	1 %
Frequency Response	
± 10 %	0.10 – 300 Hz
± 3 dB	0.05 – 500 Hz
Resonance Frequency, mounted, nominal	950 Hz
Transverse Sensitivity, max	1 % of axial

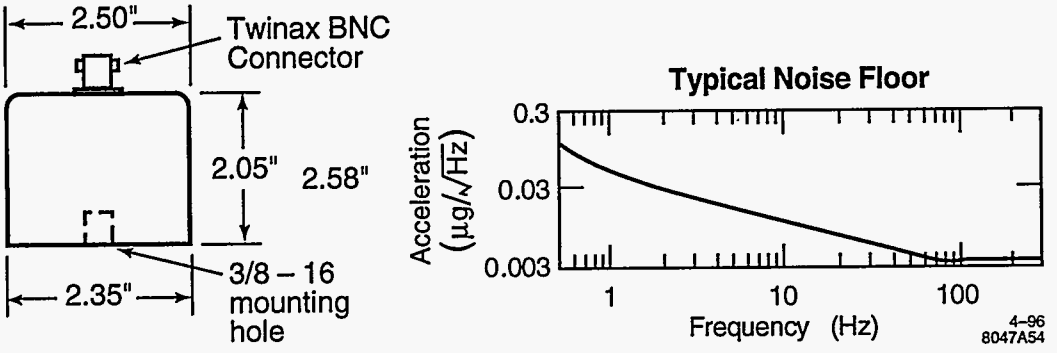


Figure C-6. *Wilcoxon Model 731 accelerometer specifications.*

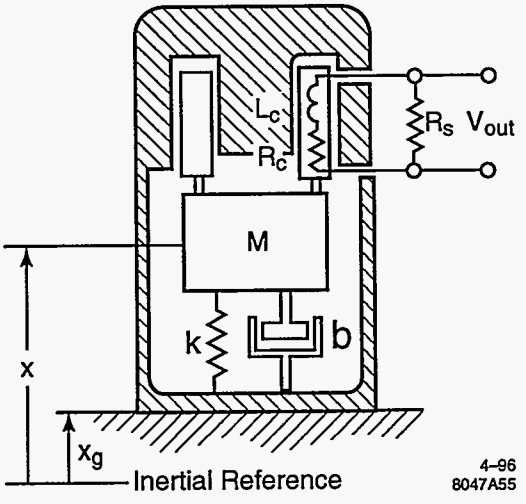


Figure C-7. *Geophone Model.*

Transduction	G	276.4	V s/m
Natural Frequency	ω_0	2π	rad/s
Coil Resistance	R_c	5500	Ω
Shunt Resistance	R_s	8905	Ω
Damping	ζ	.707	

Table C-1. *L4C parameters.*

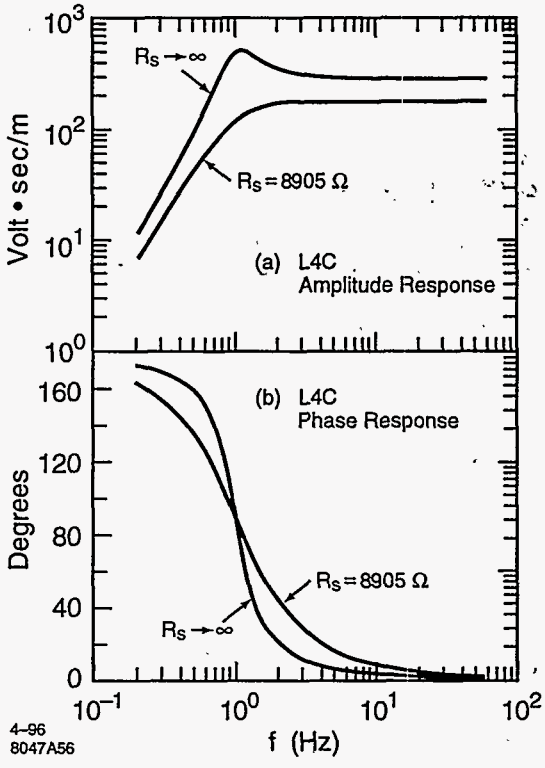


Figure C-8. *L4C transfer function.*

C.3.4 Sensor Internal Noise Sources

All inertial instruments sense the relative motion of a proof mass with respect to ground. There are three major sources of instrument noise: Brownian motion of the proof mass which is in thermal equilibrium with the outside world through various damping mechanisms [Gabrielson 1993], Johnson noise in the sensor output circuit and electrical current and voltage noise in the input amplifier [Rodgers 1992, Riedesel 1990]. Thermal noise in the mechanical proof mass/suspension system depends on the damping ratio ζ and the ratio of resonant frequency to proof mass. The power spectral density of acceleration noise above resonance is

$$S_{\text{mech}} = 16\pi kT\frac{\zeta f_0}{M} \quad (\text{m/s}^2)^2/\text{Hz} \quad . \tag{C.51}$$

For a .707 critically-damped 1-Hz oscillator with a 1-kg mass at room temperature, S_{mech} is $1.46 \times 10^{-19}(\text{m/s}^2)^2/\text{Hz}$. Integrating the corresponding displacement density above one Hz yields 6-picometer rms motion. For comparison, the low-noise model integrates to 60-picometer rms above 1 Hz. While thermal mechanical noise is not a serious limit for large proof masses with critical damping, it does set a limit for smaller sensors with proof masses measured in grams.

Johnson noise in the sensor output circuit generates a white-noise voltage signal:

$$S_{\text{johnson}} = 4kTR \quad (\text{V}^2/\text{Hz}) \quad . \tag{C.52}$$

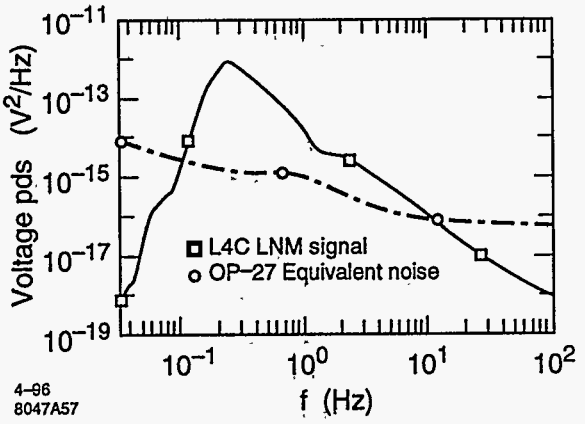


Figure C-9. *L4C low-noise model signal and amplifier noise.*

A .707 critically-damped L4C has a total parallel coil+damping resistance of 3.4 kΩ which generates a white-noise spectral density of $5.63 \times 10^{-17}\,\mathrm{V^2/Hz}$. When integrated over frequencies above one Hz, this contributes noise equivalent to 7-picometer rms motion.

The third and largest source of electronic noise comes from the various amplifiers used with the sensor. Figure C-9 [Rodgers 1992] shows the performance of the L4C with the OP-27 amplifier. The amplifier noise is below the low-noise model signal from about 0.1 Hz to 10 Hz.

C.3.5 Feedback Seismographs

The most sensitive, low-noise, broad-band modern seismographs employ electromagnetic feedback to stabilize the proof mass to the instrument housing. A model for a feedback seismograph is shown in Figure C-10. Proof-mass motion relative to the instrument housing is detected by a capacitive displacement transducer. This signal is used as the instrument output as well as for feedback control. The feedback elements produce a response equivalent to an oscillator with damping.

The Streckeisen STS-2 is a portable instrument employing feedback to detect seismic signals at the low-noise model level or better from .01 Hz to 10 Hz. Over a wide range of frequencies, the output voltage is proportional to ground velocity $\dot{x}_g$. The instrument contains three identical independent orthogonal seismographs oriented 45° to vertical (Figure C-11). Outputs for vertical and horizontal motions are appropriate vector sums of the three seismograph signals.

To fully model the instrument's response at frequencies around 50 Hz, it is necessary to include the dynamics of the housing's three support legs. The instrument response is the product of the feedback seismograph response and the response of the instrument's support structure. The Laplace transform transfer function for the STS-2 is

$$\frac{V_{\mathrm{out}}(s)}{\dot{x}_g(s)} = \left(\frac{\omega_0^2 + 2\omega_0\zeta s}{s^2 + 2\omega_0\zeta s + \omega_0^2}\right)\left(\frac{Ms^2/k_1C}{\frac{M}{k_1k_2C}s^3 + s^2 + \frac{1}{R_1C}s + \frac{1}{R_2C\tau}}\right) \tag{C.53}$$

and its parameters are listed in Table C-2. The amplitude and phase of this transfer function evaluated with $s = i\omega$ are shown in Figure C-12.

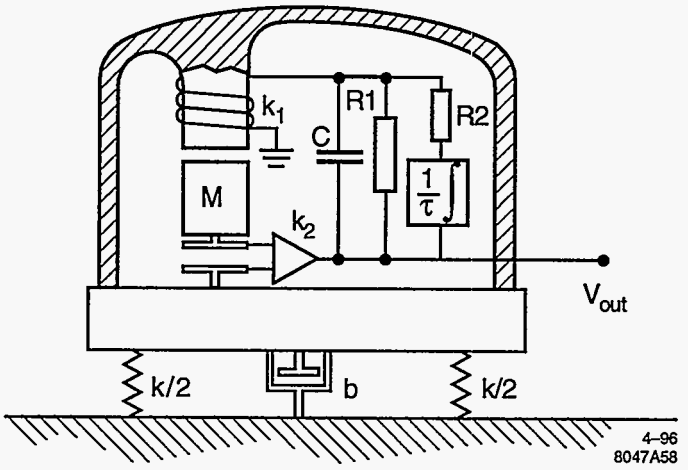


Figure C-10. *Feedback seismograph model.*

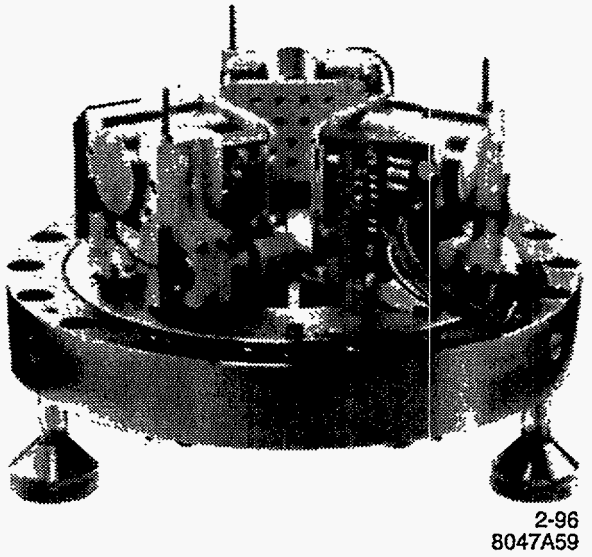


Figure C-11. *STS-2 with its cover removed.*

Proof mass	M	0.3	kg
Transduction	k_1	50	Newtons/amp
Displacement transducer	k_2	8×10^4	V/m
Integration time	τ	80	s
Feedback resistor	R_1	1.7×10^6	Ω
Integrator resistor	R_2	0.6×10^6	Ω
Derivative capacitor	C	7.8×10^{-6}	farads
Support frequency	ω_0	$2\pi(50)$	rad/s
Support damping	ζ	0.2	

Table C-2. *STS-2 parameters.*

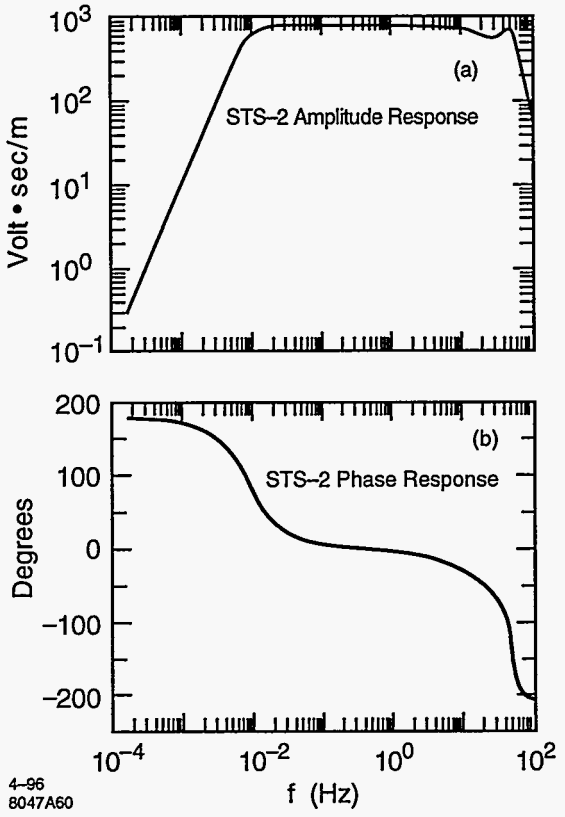


Figure C-12. *STS-2 frequency response.*

C.3.6 Other New Sensors

New methods of measuring inertial forces have recently started development. Optical strain sensor in the form of glass fiber interferometers wrapped around elastic proof-masses have recently been proposed as the basis of a sensitive accelerometer [Gardner 1987]. Field effect tunneling currents are sensitive to Å changes in the tunneling gap. Integrated circuit seismographs have recently been built using tunneling current to monitor proof-mass motion [Kenny 1994]. Because of their small size these instruments are subject to thermal noise but in the future larger transducers using this principle may be developed.

Over the past decade the needs of integrated circuit manufacturing and optical fiber communication have made commercial transducers and actuators available with sub-nanometer resolution. Capacitance position measuring sensors can be purchased with a better than 0.01-nm/$\sqrt{\text{Hz}}$ noise floor and 50-μm range.

C.4 SLAC Ground Motion Measurements and Analysis

As discussed in Section C.2, one needs to know both the ground motion spectrum and correlations to compute the its effect on beam motion in the NLC. While much data exists on ground motion spectra,

little is available on its spatial characteristics. One exception is a set of correlation measurements between points separated by 0 to 3 km in the LEP tunnel [Juravlev 1993]. However, the limited resolution of the seismometers used for these measurements does not allow one to fully examine the quantities of interest. To see how typical the CERN results are, and to better understand the dynamics of ground motion, we did a similar set of correlation measurements in the SLAC linac tunnel using higher-resolution seismometers. Because of time constraints, only vertical ground motion was studied in detail. In this section, we describe these measurements, the spectral and spatial analysis of the data, and how the results can be used to compute the effect on beam motion in the NLC.

C.4.1 STS-2 Measurements and Data Analysis

The seismic measurements were made in the SLAC linac tunnel during a period when the magnet and accelerator cooling water was shut off. Two Streckeisen STS-2 seismometers were borrowed for this purpose and were placed on the tunnel floor with Styrofoam boxes around them for thermal insulation and to shield them from air currents (see Section C.3.5 for a description of these devices). The basic measurement consisted of simultaneously digitizing the vertical ground motion signals from the two seismometers at an effective rate of 128 Hz for about a two-hour period using a PC-based data acquisition system. The actual acquisition rate was 6.4 kHz but the digitized results were decimated by 50 to reduce the data volume to a manageable size. Before decimation, the data were convoluted with a Gaussian function (3 dB point = 96 Hz) to filter out high frequency components.

Over a period of a few weeks, measurements were taken with one seismometer at the end of linac Sector 4, and the other downstream from it by distances (Δz) of 0, 6, 12, 24, 50, 100, 200, 400, 1000, and 2000 m. After any change in position, the STS-2s were allowed to stabilize for at least 10 hours before data were recorded. Typically, the configuration was kept fixed for one or two days, and three measurements were made each day beginning at 10 PM, 2 AM and 6 AM. The results presented here are based on the 2 AM data.

After completing a measurement, the data were divided into 50 128-s samples and the complex Fourier transform of each sample was computed with a cosine-like windowing function. Since the response function of the seismometers is fairly flat in the frequency (f) range of interest, no corrections were applied. However, the seismometers measure velocity, so the Fourier results were divided by $2\pi fi$ to obtain the vertical displacement components, $A_1(f)$ and $A_2(f)$, of the ground motion at the two STS-2 locations. From these, the average of the 50 power spectra were computed, $< A_1^2(f) >$ and $< A_2^2(f) >$, as well as the cross power $< A_1(f)A_2^*(f) >$, from which the correlation was obtained,

$$\text{Correlation} \equiv Re\,\langle A_1(f)A_2^*(f)\rangle \,/\sqrt{\langle A_1^2(f)\rangle\,\langle A_2^2(f)\rangle} \quad . \tag{C.54}$$

Also, the difference spectrum, $\left\langle (A_1(f) - A_2(f))^2 \right\rangle$, was calculated.

C.4.2 Measurement Resolution

The performance of the measurement system is best characterized by the data that were taken with the seismometers placed side-by-side ($\Delta z = 0$). Figure C-13 shows the two power spectra for this configuration together with the difference spectrum, and the quantity,

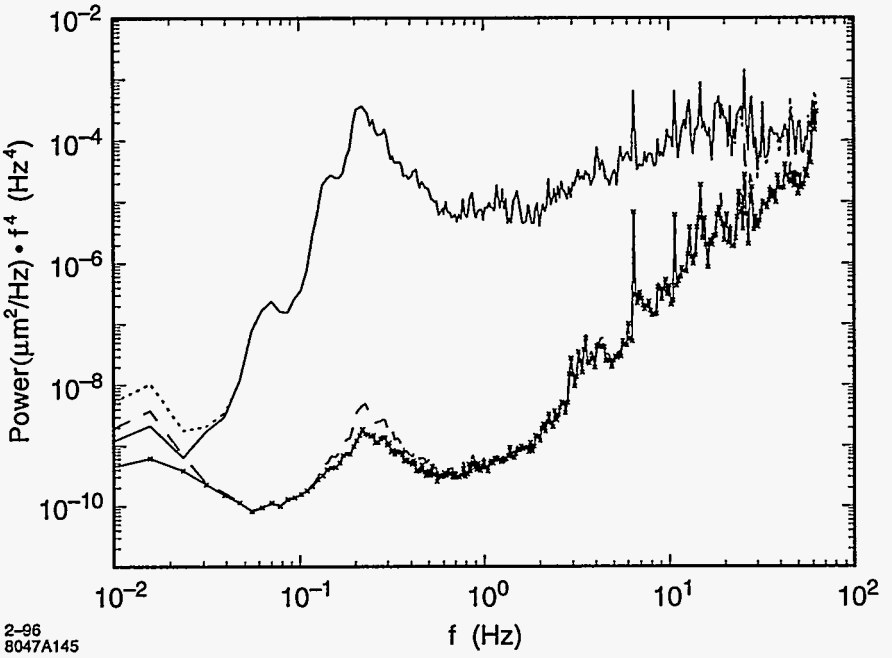


Figure C-13. *Vertical power spectra (weighted by f^4) measured by two STS-2 seismometers placed side-by-side (solid and dotted lines). Also shown are the difference (dashed line) and corrected difference (line with crosses) power spectra.*

$$\text{Corrected Difference} \equiv 2(1 - \text{Correlation})\sqrt{\langle A_1^2(f)\rangle \langle A_2^2(f)\rangle} \tag{C.55}$$

which would be the difference spectrum if the calibration of the two seismometers were identical and equal to the geometric mean of the actual calibrations (note that the spectra are weighted by f^4 to make the plot more compact). One sees that the difference spectrum is generally close to the corrected spectrum except at low frequency where the deviation corresponds to a calibration difference of about a factor of two. Also, the curves deviate near the micro-seismic peak ($f = .2\,\text{Hz}$) where the corresponding calibration difference is only a few parts in a thousand. While this latter deviation is consistent with the expected accuracy of the seismometers, the reason for large low frequency difference is not clear. However it occurs in a region where power spectrum contributes little to the integrated motion so it does not have a big effect on the results.

The difference spectrum in Figure C-13 is roughly that expected from the noise characteristics of the seismometers. To better characterize this noise floor, we computed the integral of the difference spectrum which yields the vertical rms ground motion difference that would be measured if only frequencies above the frequency plotted contributed to the motion. Figure C-14 shows this integrated spectrum together with that for the corrected difference spectrum. Thus, the STS-2s can discern ground motion differences of about 0.5 nm for $f > 0.1\,\text{Hz}$, which is adequate for most NLC studies. As a measure of how sensitive these devices are, we note that this difference level increases by about 100 if the Styrofoam boxes are not used to isolate the seismometers.

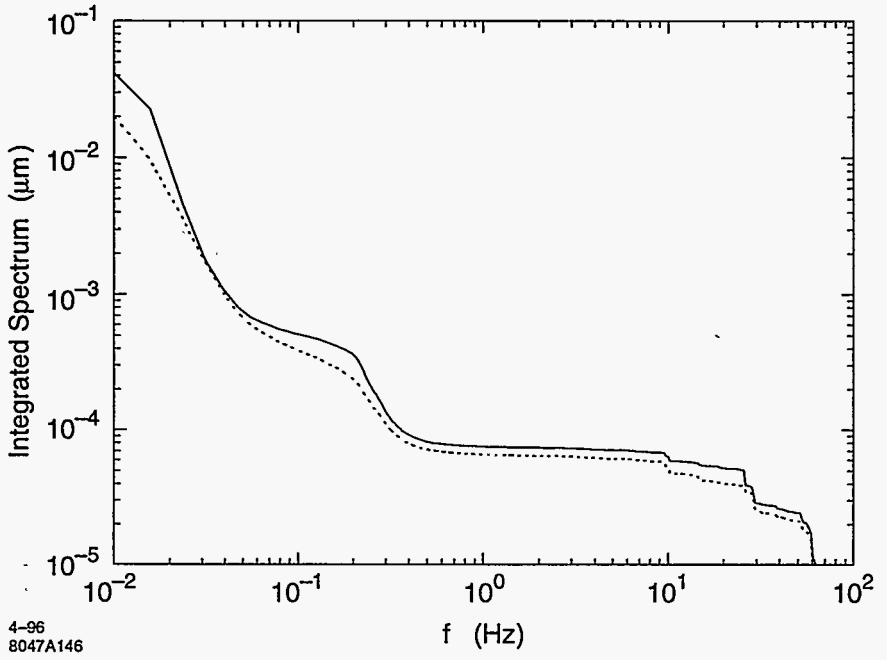


Figure C-14. *Integrated difference (solid line) and corrected difference (dotted line) power spectra computed from the difference spectra in Figure C-13.*

C.4.3 Spatial Properties of the Vertical Ground Motion

The data that were taken with the two seismometers separated by various distances were analyzed in the same manner as described above. One sees that the difference spectra rise as the seismometer separation increases. The differences in these cases are best characterized by the correlation spectra since these ratios tend to be more stable than the power spectra, which can vary by as much as ±10 dB day-to-day. Also, they are more useful for understanding the dynamics of ground motion and for computing its effect on beam transport.

As examples of correlation measurements, Figures C-15 and C-16 show the results obtained with the seismometers separated by 100 m and 1000 m, respectively. As with the CERN results, the correlation is nearly unity at low frequencies (the dip near .01 Hz is due to the electronic noise of the seismometers). Just how the correlation varies with both frequency and separation is needed to compute the effect of ground motion on the beams in the NLC. In examining the data, it was found that it can be reasonably well described by a model in which the ground motion consists of horizontally traveling waves that are isotropically distributed in direction. This model predicts that the correlation equals $J_0(2\pi f \Delta z/v)$ where J_0 is the zero order Bessel function, f is the frequency of the waves, Δz is the separation of the seismometers, and v is the wave phase velocity, which we find depends on f (*i.e.*, the ground is dispersive).

We determined v as a function of f by dividing the data into narrow frequency bins and then fitting the correlation *vs.* Δz curves for each set to the expected Bessel function dependence with velocity as the only free parameter. For this procedure, the log of (1− correlation) was actually fit to better match how the data deviates from unity, which is particularly important for modeling ground motion (note, 1− correlation = R in Section C.2). As an example, Figure C-17 shows the $0.8 < f < 0.9$ Hz data, which have been averaged

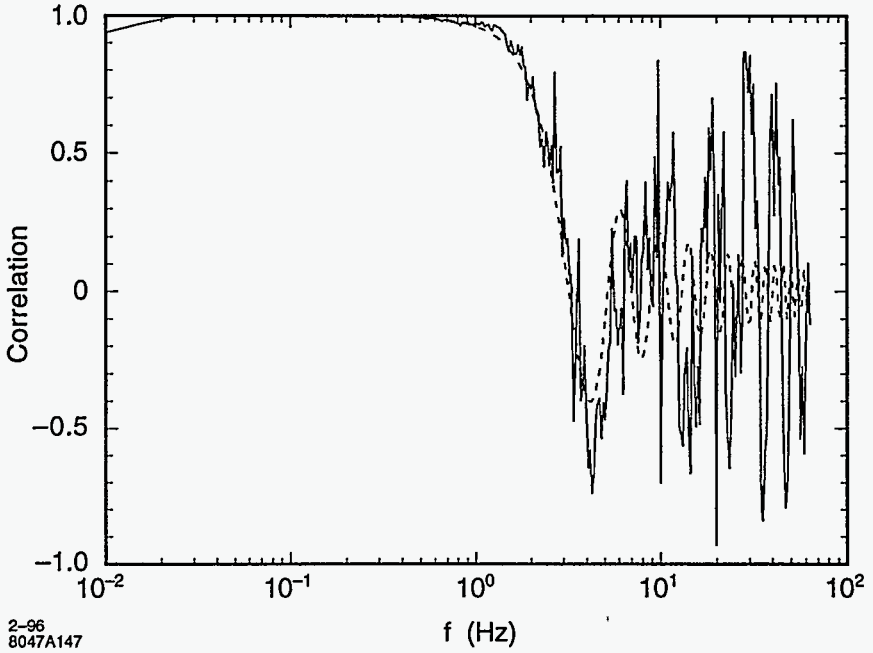


Figure C-15. *Correlation spectrum (solid line) measured with the seismometers separated by 100 m. The dashed line is described in the text.*

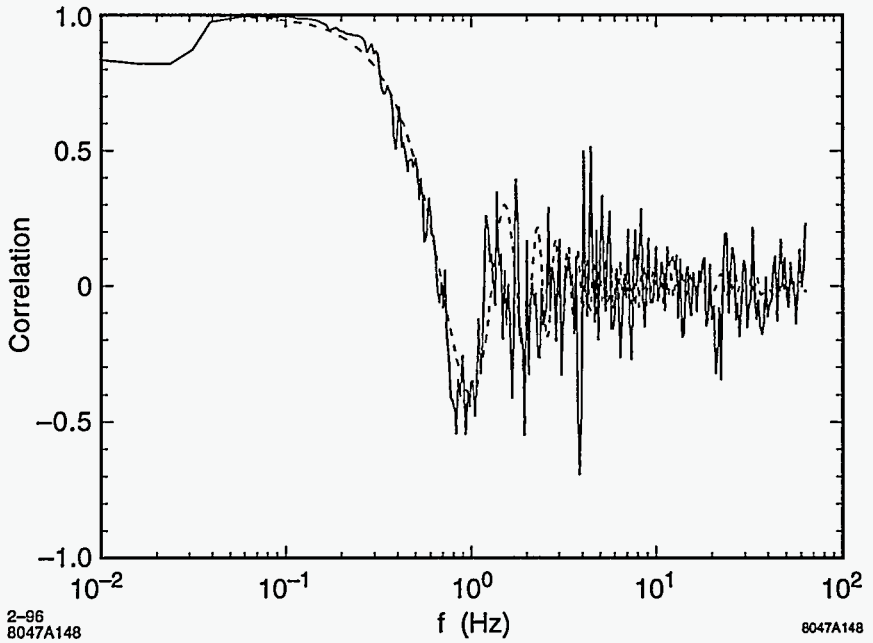


Figure C-16. *Correlation spectrum (solid line) measured with the seismometers separated by 1000 m. The dashed line is described in the text.*

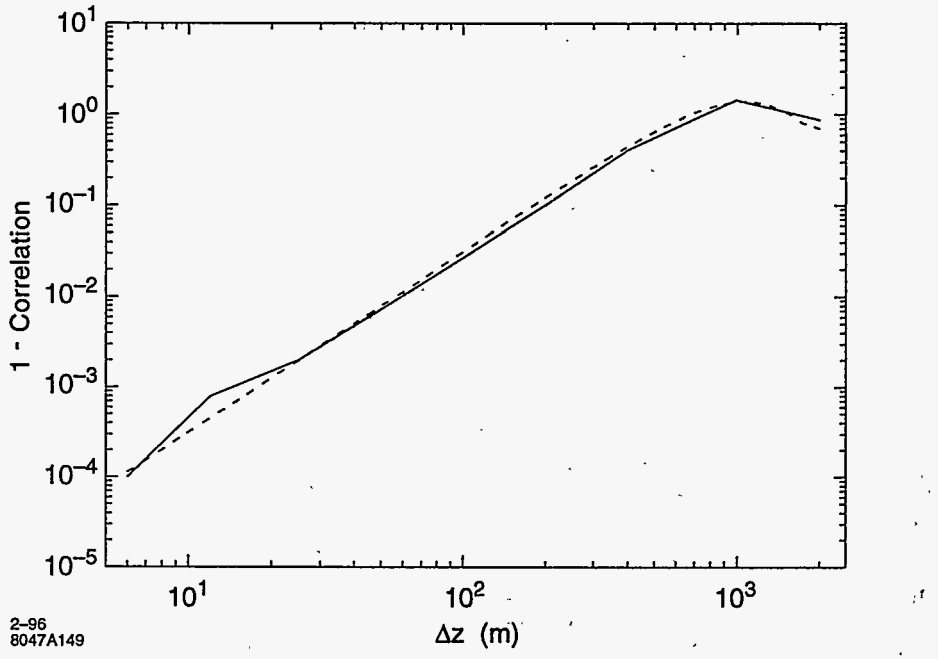


Figure C-17. *Dependence of (1− correlation) on the separation (Δz) of the seismometers for $0.8 < f < 0.9$ Hz. The data in this frequency range has been averaged (solid line). Also shown is the fit to the data (dashed line) which is described in the text.*

together in this plot, and the fit function, $1 - J_0(2\pi f \Delta z/v)$, which is evaluated at the fit velocity, 1510 m/s, and the mean frequency, 0.85 Hz.

The velocity fits were done only for the data in the $0.15 < f < 12$ Hz range. At higher frequencies, the correlations are small even at the lowest values of Δz so the velocity dependence is hard to discern. At lower frequencies, the decreasing ratio of ground motion to seismometer noise limits the minimum value of (1− correlation) that can be measured.

The values of velocity that result from the fits are plotted in Figure C-18 as a function of f. To characterize the dispersion, the data were empirically fit: an exponential function provides a good match and yields,

$$v(\text{m/s}) = 450 + 1900\exp(-f(\text{Hz})/2.0) \tag{C.56}$$

which is shown as a dashed line in Figure C-18. To check our results, we have searched geophysics journals for other measurements of dispersion in this frequency range. Although we have not yet found direct measurements, theoretical predictions of the dispersive properties of surface waves at lower frequencies would suggest that our results are not unreasonable [Bolt 1972]. Also, a simple analysis of the CERN data, using only the values of Δz at which the correlations first cross zero, yields a similar dispersion relation although the velocities are about 30% larger.

Using the measured dispersion function, the correlation versus frequency measurements were compared to predictions. The dashed lines in Figures C-15 and C-16 are examples. In regions where the correlation is high, the match is good as expected. The fact that the predictions do not match the data in regions of low correlation is less important since it is the Δz dependence of transition from high to low correlation that needs to be well modeled to characterize the spatial properties of the ground motion at any frequency.

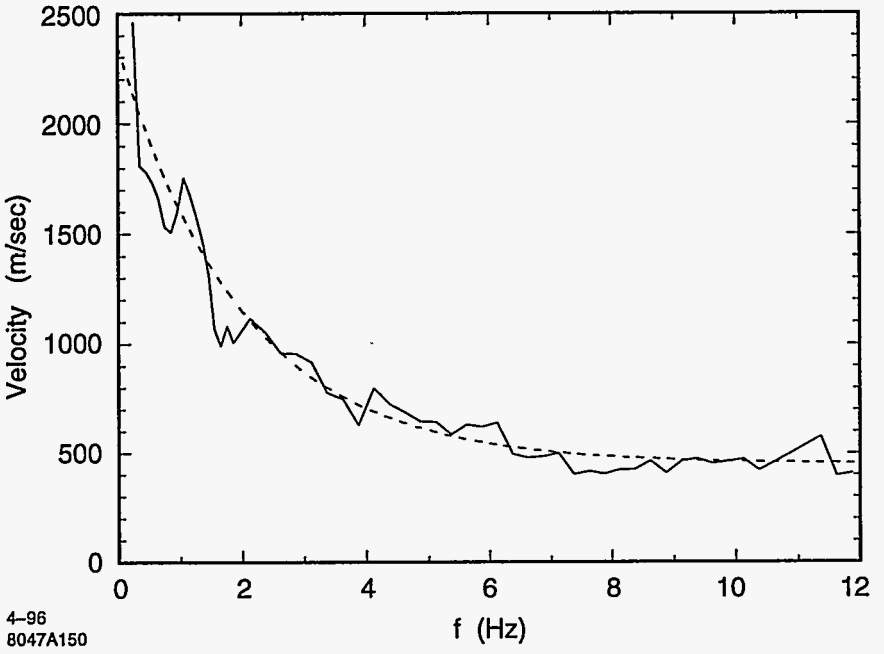


Figure C-18. *Velocity derived from fits to the correlation data (solid line). The dashed line is an empirical fit to the data:* $v = 450 + 1900\exp(-f/2.0)$.

So far we have considered only the real part of the measured correlations. If the ground motion is wave-like, and the waves are isotropic in direction, then the imaginary part of the correlations should be zero. What is observed, however, is roughly a $J_1(2\pi f\Delta z/v)$ dependence for frequencies below about 1 Hz. This dependence corresponds to a $(1+\cos(\phi))^2$ distribution of incoming waves where ϕ is the angle with respect to the linac axis. This result is not too surprising given that the SLAC linac is oriented nearly perpendicular to the Pacific coastline, and most seismic activity in this frequency range originates from ocean waves. If we redo our analysis of the real component of the correlations assuming a $(1+cos(\phi))^2$ distribution, the velocities that result are systematically about 10% larger, and the goodness-of-fits are about the same. Since only the real part of the correlation is needed for the beam motion predictions, and the J_0 dependence adequately matches the data, we continue to use this simple functional dependence and the corresponding dispersion results.

C.4.4 Application to the NLC

With the correlation function described above, the effect of ground motion on beam transport in the NLC can be derived using the formalism given in Section C.2. In particular, the luminosity reduction, $\Delta\mathcal{L}/\mathcal{L}$, due to vertical ground motion with frequencies $> f_0$ is

$$\Delta\mathcal{L}/\mathcal{L}(f_0) = \frac{\langle\Delta^2_{\rm IP}\rangle_{f_0}}{4\sigma^2_{y,\rm IP}} = \frac{1}{2\pi\sigma^2_{y,\rm IP}}\int_{f_0}^{\infty} P(f)\int_0^{k_0}\frac{G(k)}{\sqrt{k_0^2-k^2}}dk\,df \qquad \text{(C.57)}$$

where: $\sigma_{y,\rm IP}$ = rms vertical beam size at the IP, $k_0 = 2\pi f/v(f)$ with $v(f)$ from Eq. C.56, $P(f)$ = power spectrum, and $G(k)$ = lattice response function, which is defined in Section C.2.2. The effect of trajectory

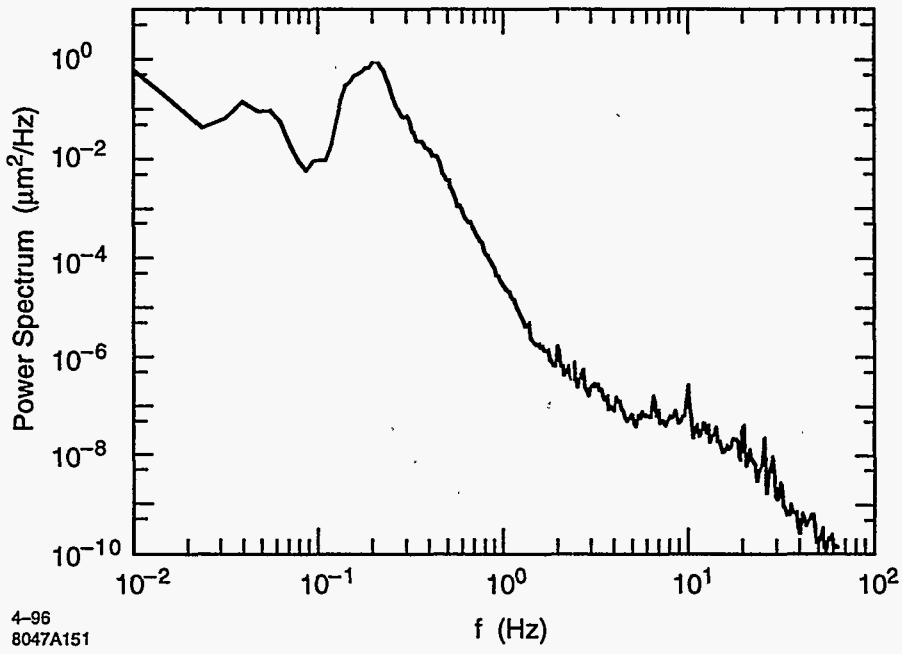


Figure C-19. *Average of 10 power spectra of vertical ground motion that were measured on different days in Sector 4 of the SLAC linac at 2 AM.*

feedback loops in the NLC can be included by factoring into the integral the frequency response function of the feedback system.

To a good approximation, one can split the function $G(k)$ for the NLC into a sum over its values in different regions of the machine (this generally yields a worst-case estimate). Only one half of the NLC needs to be considered since $G(k)$ for the opposing half is simply related by symmetry. In this approximation, $G(k)$ for a region is computed for the case where the beam initially follows the ground motion, and the difference in the beam and ground motion that develops by the end of the region remains unchanged relative to the beam size as the beam propagates to the IP. Using the $G(k)$ functions given in Section C.2.2 and the power spectrum that was measured in the SLAC linac tunnel (Figure C-19) and elsewhere, the luminosity losses due to ground motion in most regions of the NLC have been computed.

To be conservative, estimates have also been made which account for the limits that the electronic noise of the seismometers place on our measurement of correlation. As a worst case, the electronic noise contribution to the data is assumed to be ground motion that is uncorrelated, quadrupole-to-quadrupole. That is, at each frequency it is assumed that a certain fraction of the measured power is due to waves, and the remainder is due to uncorrelated motion. Combined, the two sources account for the measured power and correlation spectra except for $\Delta z = 0$ where the predicted correlation is unity but measured is not since the uncorrelated component is actually noise.

The noise power was obtained from the $\Delta z = 0$ difference data and is plotted in Figure C-20 as a fraction ($\equiv \rho_n$) of the average total power (note that ρ_n is the minimum value of (1 - correlation) measurable if the ground motion spectrum equals that in Figure C-19). Including the noise as uncorrelated motion, the

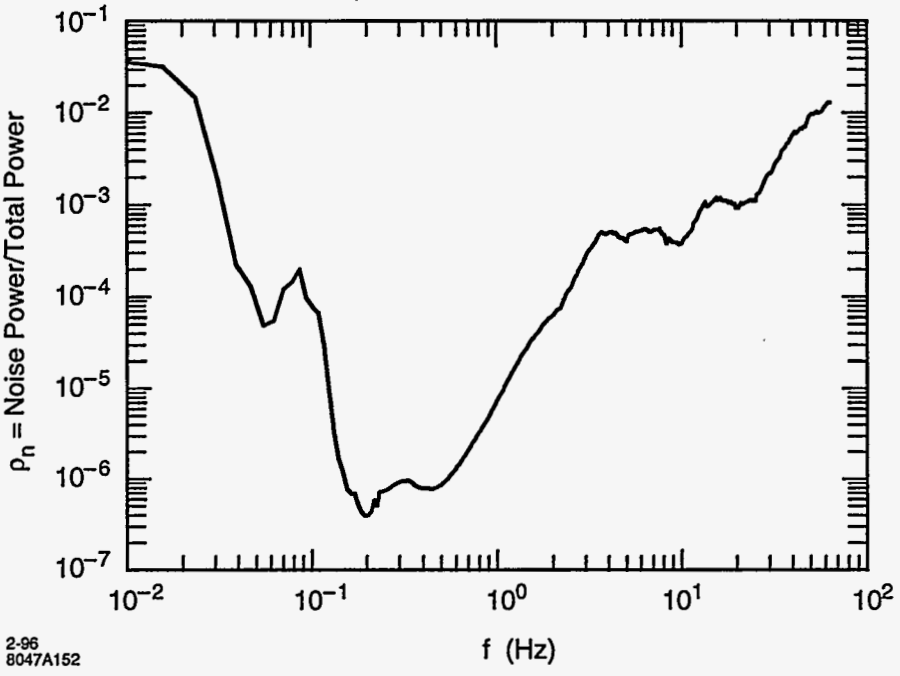


Figure C-20. *Fraction (ρ_n) of the power in Figure C-19 that is due to STS-2 electronic noise.*

expression for the luminosity reduction becomes

$$\Delta\mathcal{L}/\mathcal{L}(f_0) = \frac{1}{2\pi\sigma_{y,\mathrm{IP}}^2}\int_{f_0}^{\infty} P(f)\left[\int_0^{k_0}\frac{(1-\rho_n)G(k)}{\sqrt{k_0^2-k^2}}dk + \frac{\pi\rho_n G(\infty)}{2}\right]df \quad . \tag{C.58}$$

To predict the luminosity loss due to lower-frequency ($f < .01\,\mathrm{Hz}$) ground motion requires some assumption about the spatial characteristics of the motion since we do not have measurements in this regime. There is evidence at long time scales (days to years) that ground motion has an "ATL" behavior where the variance of relative displacements grow linearly with time and separation (*i.e.*, $\langle\Delta y^2\rangle = A \times T \times L$ where A is a constant) [Baklakov 1991]. The frequency scale at which the motion changes from predominately wave-like to ATL-like is not known. However, as a worst case, one can assume that this occurs at 0.01 Hz and use the value of A obtained from our data as an upper limit. To compute A, we note that the power spectrum for ATL-like motion is

$$P(\omega,k) = 4A/\omega^2k^2 \tag{C.59}$$

so the rms squared difference of the motion for $f > f_0$ at two locations separated by Δz is

$$\langle\Delta y^2\rangle = \int_{2\pi f_0}^{\infty}\int_0^{\infty}\frac{4A}{\omega^2k^2}2\left(1-\cos(k\Delta z)\right)\frac{dk}{2\pi}\frac{d\omega}{2\pi} = \frac{A\Delta z}{2\pi^2 f_0} \tag{C.60}$$

For $\Delta z = 200\,\mathrm{m}$ and $f_0 = .02\,\mathrm{Hz}$, we measure a 10-nm-rms vertical deviation, yielding $A = 2\times10^{-7}\,\mu\mathrm{m}^2/\mathrm{m/s}$, which is on the low end of that measured on long time scales (another method to compute an upper limit on A is discussed in Section C.2.1). Using this value in the ATL power spectrum above, one can thus

make worst-case estimates of the low frequency contribution to luminosity loss assuming that A is indeed frequency-independent.

The results on luminosity loss due to ground motion in the various regions of the NLC are presented in the chapters that discuss these regions. It should be noted that the general formalism presented here can also be used to estimate beam-emittance growth that depends quadratically on quadrupole position, such as that due to dispersion.

References

[Baklakov 1991] B.A. Baklakov, P.K. Lebedev, V.V. Parkhomchuk, A.A. Sery, A.I. Sleptsov and V.D. Shiltsev, "Investigation of Seismic Vibrations and Relative Displacement of Linear Collider VLEPP Elements", *Proc. 1991 Part. Accel. Conf.*, San Francisco, CA, 3273 (1991).

[Bolt 1972] B.A. Bolt (editor), "Seismology: Surface Waves and Earth Oscillations", *Methods in Computational Physics*, Vol. 11 (Academic Press, NY, 1972).

[Gardner 1987] D.L. Gardner and S.L. Garret, "Fiber Optic Seismic Sensor", SPIE Vol. 838, *Fiber Optic and Laser Sensors V*, 271–277 (1987),

[Gabrielson 1993] T.B. Gabrielson, "Mechanical-Thermal Noise in Micromachined Acoustic and Vibration Sensors", *IEEE Trans. Electron. Devices* **40**, 903–909 (1993).

[Juravlev 1993] V.M. Juravlev *et al.*, "Investigations of Power and Spatial Correlation Characteristics of Seismic Vibrations in the CERN LEP Tunnel for Linear Collider Studies", CERN Note CERN–SL/93–53 and CLIC–Note 217 (December 1993).

[Kenny 1994] T.W. Kenny, W.J. Kaiser, H.K. Rockstad, J.K. Reynolds, J.A. Podosek, and E.C. Vote, "Wide-Bandwidth Electromechanical Actuators for Tunneling Displacement Transducers", *J. Micro. Elec. Sys.* **3**, 97–103 (1994).

[Riedesel 1990] M.A. Riedesel, R.D. Moore, and J.A. Orcutt, "Limits of Sensitivity of Inertial Seismometers with Velocity Transducers and Electronic Amplifiers", *Bull. Seismological Soc. of Am.* **80**, 1725–1752 (Dec 1990).

[Rodgers 1992] P. Rodgers, "Frequency Limits for Seismometers as Determined from Signal-to-Noise Ratios", *Bull. Seismological Soc. of Am.* **82**, 1071–1098 (April 1992).

Contributors

- Chris Adolphsen
- Gordon Bowden
- John Irwin

D

Beam-based Feedback: Theory and Implementation

Contents

D.1	Introduction	1076
D.2	Planned NLC Feedbacks	1076
D.3	Feedback System Design	1077
D.4	Performance Questions	1077
D.5	Adaptive Cascade	1078
D.6	Rate Considerations and Corrector Speeds	1078
D.7	Calibrations and Modeling	1079
D.8	Global Performance Characterization	1079
D.9	Summary	1080

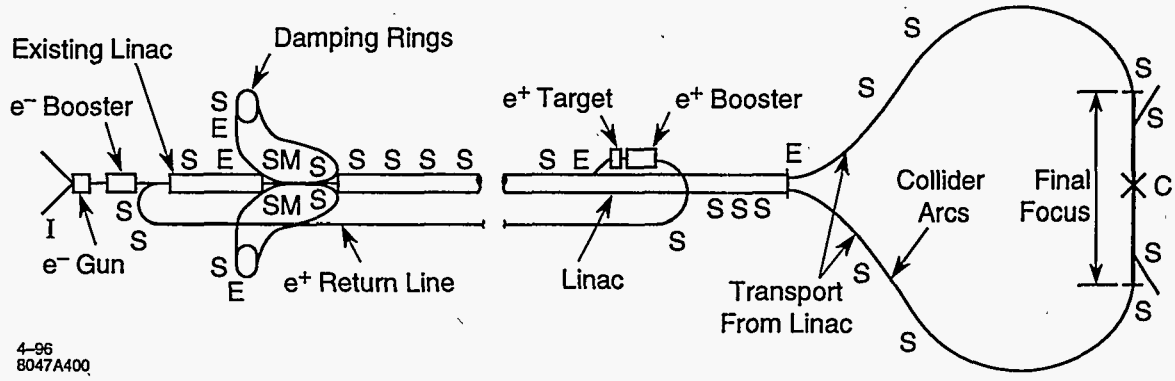


Figure D-1. *SLC schematic with feedback locations. S=steering; E=energy; I=intensity/gun; C=maintains beam collisions; M=minimization*

D.1 Introduction

In planning for feedbacks in the NLC, it is useful to study the existing SLC feedback system as a model. While the control system architecture for the NLC is likely to be different from the SLC, the basic feedback algorithms should be similar. The SLC feedback system is the first of its kind for accelerators. It is a generalized, database-driven system that applies linear closed-loop control. The mathematical design goal is to minimize the RMS of the beam jitter by applying state space techniques. In practice in the SLC, the feedback serves to stabilize the beam and decouple different areas of the machine, facilitating smoother startup after machine outages and easier machine tuning. This system has been essential to the operation of the SLC, and it is assumed that a similar system will be equally necessary for the NLC.

D.2 Planned NLC Feedbacks

For the NLC, feedback loops are planned for each major area of the machine, similar to the SLC feedbacks. At the guns, intensity and timing parameters will be controlled. In addition, feedback controls for each of the laser wire scanners are expected. Energy and steering (position and angle) parameters will be regulated for damping ring injection and extraction. For each linac, five diagnostic sections will each have energy and steering controls. For each final focus, feedbacks will stabilize the beam steering at two locations. Finally, an interaction-point feedback loop will regulate the beam-beam deflections to keep the beams in collision.

The feedbacks are planned to run at the full beam rate of 120 or 180 Hz. The steering loops control the average trajectory of the bunch train rather than individual bunches, so the "Q" BPMs are used, which measure the average of the train. Dipole corrector magnets are used, for which the feedback controls the magnetic field by setting a DAC (digital to analog converter) which alters the current from a power supply. Correctors and other actuators need to respond (to make 90 percent of a requested change) in a single 180-Hz period.

If SLC experience is taken as a model, many additional feedbacks are likely to be added to those originally planned. For the SLC, the original eight loops have been expanded to over 50 control loops, many of which are shown in Figure D-1.

D.3 Feedback System Design

The SLC feedback system is designed in a generalized, database-driven fashion, which contributes greatly to the flexibility of the system [Hendrickson 1995]. The system uses standard control-system hardware. As a result of this design, unplanned control loops can often be added with only database work, without requiring hardware or software changes. In addition to the beam position monitors and correctors used for steering control, the feedback system is capable of measuring and controlling a wide variety of devices. For example, it is equally trivial to add a steering loop in the linac as it is to stabilize the laser gun timing in the injector. Special-purpose extensions to the linear feedback system have been added to accommodate non-linear cases, such as optimization feedbacks in which the measurement responds parabolically to actuator movement. The system also provides built-in diagnostic and analysis capabilities, and the many sample-only monitoring loops provide a wealth of diagnostic information. These design features have been key to the success of the SLC system, and the NLC design should be equally flexible and extensible in order to support unplanned controls needs.

The feedback control algorithm is based on state space formalism, with an LQG (Linear Quadratic Gaussian) controller. Matrices are designed and calculated offline in advance, with inputs including a model of the beam transport and the expected beam and BPM noise characteristics. The mathematical design minimizes the rms of the beam states over time, given the noise inputs [Himel 1991]. By modifying the input beam noise design assumptions, it is possible to tune the feedback performance response characteristics. The initially-proposed feedback algorithm does not adapt to modeling or noise spectrum changes, consistent with the current SLC design. In the future, adaptive methods may provide improved performance. The beam transport characteristics within a single loop may either be obtained from the accelerator model, or be measured by an invasive online beam-based calibration procedure.

An extension to the basic feedback system, "cascade" is designed to allow multiple linac loops to communicate with each other, avoiding overcorrection problems when a perturbation is induced upstream of the chain of feedback loops. With cascade, each feedback loop receives the calculated positions and angles from the next upstream loop, mathematically transports them to the downstream location, and subtracts them from the states calculated using the local BPMs. The resulting adjusted states are then used for the local feedback corrections, so that each loop should correct only perturbations which were not seen by an upstream loop.

D.4 Performance Questions

There are a variety of performance questions which have been investigated in the SLC system, in particular with the chain of linac steering loops. These include concerns about the stability of the SLC model, speed of the steering magnets, and functions of the cascade system. In the SLC, gain factors can be used to slow the feedback response and reduce sensitivity to suspected feedback imperfections. Several types of imperfections are discussed in the following sections. Where quicker feedback response is not needed, the lower gain factors have been successful in improving the stability of the feedback system.

D.5 Adaptive Cascade

Problems were observed with the cascade system, which is intended to allow each of a string of feedback loops to correct only the perturbations initiated immediately upstream of it. The cascade system relies on a linear beam transport which is independent of the source of the perturbation. The adaptive feature of the cascade system enables each feedback loop to learn the transport from the upstream loop, using the beam jitter to calculate the beam transport. The adaptation has the assumption that perturbations immediately upstream of a feedback loop are uncorrelated with upstream perturbations. In the current SLC cascade design, each loop obtains beam information only from the adjacent upstream loop; the assumption is that if the loop upstream saw a perturbation, either it or any loops further upstream will eventually fix it.

In several tests, it was observed that the linac loops did not exhibit perfect cascade response. In particular, the feedback response to perturbations induced in the middle of the linac is different from the response to perturbations from the beginning of the linac. The SLC design assumes that the beam transport is independent of the source of a perturbation. At high currents this is not valid because of the effect of transverse wakefields which cause oscillations to propagate differently depending on their source. This is a fundamental problem with the SLC architecture, but for the low currents of the NLC, this effect should be less significant. If wakefield effects are expected to be a problem for the NLC, an alternate cascade design should be developed to provide downstream feedback loops access to beam information from all upstream loops instead of just the single nearest upstream neighbor. This would require more complicated algorithms and more communications paths than those available in the present SLC system.

Tests of cascade performance during low-current operation uncovered additional information. A test was done to confirm that the cascade system responds well when the transport is invasively measured (instead of adaptively calculated), and the resulting cascade response was between 95 and 100% effective (nearly perfect). A design flaw in the adaptive beam transport calculation, associated with exception handling for broken BPMs, was found and fixed. In the most recent test, large betatron jitter was induced upstream of the chain of linac feedback loops. The feedbacks adapted to the induced beam noise, and the resulting cascade response was between 90 and 100% effective. However, without either measurement or induced noise, the performance was much poorer. This may be explained by an additional design flaw, in which poor BPM resolution during low-current SLC running results in incorrect adaptive transport calculations. It is hoped that when this design flaw is fixed or better BPMs are available, low-current cascade adaption performance will be improved.

Further concerns include questions about the nature of the incoming beam jitter, which is the source of the adaptive transport calculations. If the incoming jitter is not dominated by betatron jitter, adaptive transport calculations would get the wrong answer. For the NLC, a conservative initial plan is to omit the cascade adaption to calculate the interloop beam transport matrices, and instead measure the transport semi-invasively by stealing beam pulses and perturbing the beam once an hour. With this method, the cascade should be able to work well at low intensities.

D.6 Rate Considerations and Corrector Speeds

Ideally, feedbacks should operate at the full repetition rate of the machine. At the SLC, where the rate for many loops is limited to 20 Hz by CPU and other hardware constraints, aliasing problems associated with

the partial sampling rate have been observed. Full-rate operation requires sufficient processing power and hardware response time.

A source of poor performance which was investigated at the SLC is the sensitivity to corrector speeds. Feedback simulations showed that, when the corrector speeds are slower than expected, the feedback performance is degraded. In particular, performance is extremely sensitive to the relative speeds of correctors within a single loop. If some correctors are slow, it is better to have them all the same speed as the slowest one. This effect is exaggerated when there are modeling errors or other imperfections in the loop design.

The correction system needs to be designed such that the response of the correctors is within one interpulse period, considering the speed of the power supply and the field propagation through the beam pipe. For the latest SLC run, a new feedback linac steering loop was implemented which is capable of 120 Hz response. This feedback loop has been commissioned and, under some noise conditions, has decreased the RMS of the beam jitter by up to 40%.

D.7 Calibrations and Modeling

The correct functioning of a feedback loop depends on knowledge of the model and transport between the steering magnets and position measuring devices. For the SLC, this transport is derived from the online machine model. A calibration procedure is provided to check and possibly update the transport matrices. Because such a procedure is invasive, it is rarely used. For the future, it may be desirable to use a fully adaptive feedback algorithm which is capable of responding online to machine changes. This extension of the state-space formalism is currently being developed at CEBAF, and may be useful for both the SLC and the NLC.

Recent SLC studies indicate that poor modeling is not currently as significant a problem for the linac feedbacks as are slow correctors and imperfect cascade performance. However, sensitivity to the model is exacerbated by incorrect modeling of the corrector response or other errors. Furthermore, simulations show that minor modeling imperfections can have disastrous results when combined with aggressive noise designs such as notch filters.

D.8 Global Performance Characterization

In addition to studies of specific sources of feedback imperfection, measurements and simulations were performed to study the global feedback response of a series of linac loops. For the SLC, measurements were taken, comparing the response of a step function with the series of linac loops off versus on. The FFT of both sets of data were calculated; the ratio of the FFTs with feedback on and off provides a measured amplification curve. Simulations reproduced the measured results, including the effects of measured imperfections. These measurements and simulations included effects of imperfect cascade correction, imperfect modeling, slow correctors, multiple loops running at low rates, and different gain factors. Reproducing the measured SLC performance improved confidence in our ability to simulate imperfections realistically. Given some assumptions about NLC conditions and imperfections, performance was evaluated for the series of NLC linac loops.

The NLC simulations evaluate a series of five linac loops, each running at the full beam rate of 120 Hz with relatively fast correctors that can control in one machine pulse. For the NLC, slow correctors should

not be a problem. Imperfections in modeling and cascade response are assumed for the NLC. We assume that modeling/calibration imperfections are comparable to SLC conditions. We assume that cascade is 85% effective. While still imperfect, this is better than the current SLC performance; wakefield effects shouldn't be a significant problem, and we assume that the cascade transport for the NLC will be measured each hour. For the "SLC" noise design, which corrects a step function with an exponential time constant of six pulses, the results of the NLC simulation were very good. Response of the system to a step function is good and the system is able to damp very well at frequencies below 6 Hz. Above 6 Hz, beam noise is amplified somewhat, with a maximum amplification of 1.5. In order to damp more strongly at low frequencies, more aggressive noise designs were considered. While the more aggressive designs damp better at low frequencies and are able to damp noise up to 10 Hz, the jitter amplification at higher frequencies is increased, up to a factor of 4.5 for one design. Furthermore, when the imperfections are considered, the more aggressive designs have a poorer response to a step function. These studies indicate that with the more-conservative "SLC" noise design, modest imperfections have a minor effect on performance, but with a more aggressive design the same imperfections become significant. More work should be done to find an optimal design, but at this point the "SLC" noise design produces acceptable results while providing a robust system which is tolerant of minor imperfections.

D.9 Summary

Initial NLC simulations indicate that acceptable feedback performance can be obtained with an "SLC"-type noise design. More work should be done to characterize feedback performance and to study possibilities for improvements. Additional beam studies and simulations should be done at low current to insure that the cascade performance under NLC conditions will be acceptable. Also, more work should be done to determine an optimal noise design using the results of recent ground motion studies. Simulations show that any feedback system will amplify incoming jitter at some frequency, but in the design of the feedback system we have some control over the frequency and magnitude of the amplification. The goal would be a design which damps at high-jitter frequencies and minimizes amplification while providing a robust feedback system.

SLC experience has shown that operational considerations are just as important as noise response. Feedback systems which decouple different areas of the machine minimize the invasiveness of tuning procedures, allowing downstream programs to continue by automatically stabilizing the beam. Operators are freed from many routine responsibilities, allowing time for more subtle tuning. The generalized design of the feedback system has allowed extension to many unplanned applications. Feedbacks are integrated with optimization packages; for example, feedback set-points are controlled to minimize beam emittance. Finally, the system improves machine reproducibility, supporting easier startup after outages and improving machine efficiency.

References

[Hendrickson 1995] L. Hendrickson *et al.*, "Fast Feedback for Linear Colliders", *Proc. 1995 Part. Acc. Conf.* (1995).

[Himel 1991] T. Himel *et al.*, "Use of Digital Control Theory State Space Formalism for Feedback at SLC", *Proc. 1991 Part. Acc. Conf.* (1991).

Contributors

- Linda Hendrickson