EXPANDED STUDIES OF LINEAR COLLIDER FINAL FOCUS SYSTEMS AT THE FINAL FOCUS TEST BEAM

Peter Gregory Tenenbaum

SLAC-Report-95-475

December 1995

MASTER

STANFORD LINEAR ACCELERATOR CENTER

.

EXPANDED STUDIES OF LINEAR COLLIDER FINAL FOCUS
SYSTEMS AT THE FINAL FOCUS TEST BEAM*

Contents

Peter Gregory Tenenbaum

Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

* Ph.D. thesis, University of California, Santa Cruz

DISCLAIMER

List of Tables

List of Figures

Abstract

In order to meet their luminosity goals, linear colliders operating in the center-of-mass energy range from 350 to 1,500 GeV will need to deliver beams which are as small as a few nanometers tall, with x:y aspect ratios as large as 100. The Final Focus Test Beam (FFTB) is a prototype for the final focus demanded by these colliders: its purpose is to provide demagnifications equivalent to those in the future linear collider, which corresponds to a focused spot size in the FFTB of 1.7 microns (horizontal) by 60 nanometers (vertical).

In order to achieve the desired spot sizes, the FFTB beam optics must be tuned to eliminate aberrations and other errors, and to ensure that the optics conform to the desired final conditions and the measured initial conditions of the beam. Using a combination of incoming-beam diagnostics, beam-based local diagnostics, and global tuning algorithms, the FFTB beam size has been reduced to a stable final size of 1.7 microns by 70 nanometers. In addition, the chromatic properties of the FFTB have been studied using two techniques and found to be acceptable. Descriptions of the hardware and techniques used in these studies are presented, along with results and suggestions for future research.

Acknowledgements

First, I would like to thank the forces of light at UC Santa Cruz for their assistance, moral and financial support, problem-set wisdom, and all-around studliness (and studentliness). Hartmut Sadrozinski gave me a job when I had no skills, and taught me the meaning of "experimental technique" (as well as the meaning of pain when he had me test 700 CMOS transistors). Bill Rowe imbued me with a small fraction of his vast hardware knowledge and with his relaxed philosophy of life. Abe Seiden and Georgia Hamel waved the appropriate magic wands which allowed me to work on the FFTB despite the fact that UCSC is not actually a member of the collaboration. Marylin Stevens and Judy Hurst kept the bureaucratic forces of darkness at bay, and listened patiently to countless prolix tales of woe from me. Thank you all, everyone.

A special thanks goes out to the "Grad Student Mafia:" Mark Nockelby, John Hiser, Richard Schnee, Robert Bork, and Glenn Reuschling, with whom I shared the equivocal joy of countless problem sets and studying for the Quals.

No discussion of my years at UC Santa Cruz could be complete without a bow of gratitude to Pat Burchat and George Brown, my UC advisors during the FFTB era. Pat and George introduced me to accelerator physics and SLAC, and Pat especially was a crucial part of my early beam-optics education. Both George and Pat have worked to maintain the shadowy pipeline which allowed me to eat and pay rent while preparing FFTB experiments as well, which was no small part in the ultimate preparation of this tome.

Many forces around the SLAC/FFTB continuum have contributed invaluably to

my knowledge of Physics in general and Accelerator Physics in particular. In the FFTB Collaboration itself, my education has benefited immeasurably from the availability of Katsunobu Oide and John Irwin, whose patience at explaining their approaches and innovations I have yet to exhaust. Rick Iverson brought home to me the importance of considering the operational side of accelerators: ensuring that the right control tools were available at the right time, and that you actually knew in advance what they did. Ghislain Roy prepared the most comprehensive theoretical guide to the FFTB, which I have referenced and quoted liberally in this document, and also guided me in adapting my slowly-growing general understanding of beam optics to the very specific case of final focus systems. Pantaleo Raimondi's matchless energy and limitless adaptability to accelerator problems has been an inspiration, and much of what I have learned and done in the FFTB experimental regime has built on foundations laid by him. Lee Ann Yasukawa, Greg Sherwin, Terri Lahey, and Nancy Spencer wrote most of the special-purpose control software for FFTB, and confidently assured me that no code was too wacky to implement in the SCP. Steve Wagner and Mike Hildreth provided Physics and computer support of too broad a band to even describe, and friendship as a side benefit. Special kudos to Mike, without whose help the nuts and bolts of producing this document would have been beyond me.

A special acknowledgement is due to David Burke, my thesis advisor and leader of the FFTB Collaboration. In addition to the more mundane matters of taking on an unknown grad student from another university, Dave gave me unparalleled opportunities to involve myself in an exciting and forward-looking experiment; and imbued me with some fraction of his particular wisdom, most notably the importance of asking the right question: once you've nailed down what you really want to ask, digging up the answer is straightforward.

There are many others – too many to list here – who have taken time out of their work schedules to educate me in their particular areas of expertise. To the members of this group – they know who they are – I want to say, collectively, thank you.

On a practical note, I would like to thank the Northern California chapter of the ARCS (Achievement Rewards for College Scientists) Foundation for a fellowship which arrived at just the perfect moment to hand me victory over financial turmoil.

Away from the Physics front (nominally), there are the many personal friends and relations who made this research possible. At the top of the list is my Santa Cruz family, "The Banzai Institute:" Pete, Ken, Carolyn, jj, Paul, Lori, Pat, Grant, Chuck, Jeanine, Nicki, and especially Donald. These people helped me to laugh through good times, bad times, triumph, adversity, good turkeys, bad tacos, and the occasional roomful of chimney rubble. Without them, life would indeed have been a bummer these last six years.

No list of supporters could possibly be complete without reference to my family, whose financial and moral support enabled me to attend grad school, encouraged me not to leave when things got ugly, and generally made it possible for me to get here by forcing me to do my homework in the fifth grade. Mom, Dad, Lynn, Grandma and Grandpa Tenenbaum, Grandma and Grandpa Donkin: you were right. Thank you for everything.

For providing the soundtrack to my graduate career: my gratitude goes out to Fender Corporation, Andrew Eldritch, Daniel Ash, and Peter Murphy.

Last but not least, thanks to those who brought romance to my life: Malaika, Julia, and most especially Martha. Thank you for reminding me to sleep, and encouraging me to live.

Introduction

This thesis presents the results of a series of accelerator physics experiments performed at the Stanford Linear Accelerator Center between 1993 and 1995. The experiments were performed at the Final Focus Test Beam, a facility constructed to act as a prototype final focus system for a future linear collider. In this capacity, the FFTB's design goal was to demonstrate in an existing facility the tremendous demagnification expected in such a future facility; more specifically, the FFTB was designed to demagnify the incoming electron beam from the SLAC linear accelerator by a factor of 380, to a final size of 60 nanometers. Prior to the construction of the FFTB, such a demagnification of a charged- particle beam had never been achieved.

The experiments reported here are a series of measurements of the properties of the FFTB and the incoming electron beam, and a set of corrections intended to drive those properties towards their design values. The ultimate goal of the entire experiment is to demonstrate a set of measurements and corrections which is necessary and sufficient to achieve the design demagnification, with the expectation that the same set will be necessary and sufficient to achieve the equivalent demagnification in the future machine. Because the design goals of the FFTB were unprecedented, many of the tuning algorithms were similarly unprecedented, while others extended existing techniques to more demanding limits. The subject of this thesis is the design and implementation of the tuning algorithms, an analysis of their performance, and recommendations for future improvements.

Chapter 1 is a general description of linear colliders: their advantages and dis-

advantages relative to conventional circular colliders, and brief description of their subsystems. This chapter also details the functions of the final focus system, and presents the only existing example – the Stanford Linear Collider (SLC) Final Focus.

Chapter 2 provides a detailed description of the Final Focus Test Beam: its optics, aberrations, and hardware. This sets the scale of the problem to be solved, and gives some information about the tools available to solve it.

Chapters 3, 4, and 5 essentially present the tuning of the FFTB from upstream to downstream: measurement and matching of the incoming beam, local diagnostics which attempt to pinpoint problems and correct them at their source, and global diagnostics which are used to eliminate aberrations distributed over the beamline with a small number of correction elements. The FFTB ultimately achieved a 70 nanometer vertical beam size, which is also presented in Chapter 5; this chapter describes the beam size expected from the measurements listed in Chapters 3, 4, and 5, and discusses the discrepancy between expected and achieved beam sizes.

Chapter 6 examines the performance of the FFTB as a system, considering two overall issues: the time needed to tune the beam size, and the energy-dependent aberrations remaining in the system. Finally, Chapter 7 summarizes the results of the FFTB experiments and offers recommendations towards the design of future linear colliders.

Chapter 1

General Considerations of Linear Colliders

For many years, electron-positron (e^+e^-) colliders have been widely recognized as a valuable tool for understanding fundamental particle interactions. Because of the point-like, non-composite nature of electrons and positrons, such colliders are ideal for high-precision studies which cannot be undertaken in hadronic colliders.

More recently, there have been studies [1] which demonstrated the usefulness of an extremely high- energy (0.5 – 1.5 TeV CM), high-luminosity ($10^{33} - 10^{34} cm^{-2} sec^{-1}$) e^+e^- collider. While neither of these requirements, in and of itself, presents an impossibly daunting design goal, the particular combination of such energies and e^+e^- collisions mandates the choice of a linear collider rather than a conventional circular one.

1.1 Circular versus Linear Colliders

As the name implies, the purpose of a *collider* is to *collide* bunches of accelerated charged particles which are moving in opposite directions; at the collision point, the energy of the particles can be transformed into matter, specifically into (hopefully)

interesting massive particles. The probability of such a collision can be enhanced by creating a particle "traffic jam" at the Interaction Point (IP), cramming the maximum number of particles into the smallest possible space. The severity of the "traffic jam" can be expressed quantitatively as the instantaneous luminosity of a collider, which is given by:

$$\mathcal{L} = \frac{N^+ N^- f}{4\pi \sigma_x \sigma_y} \cdot H_D. \tag{1.1}$$

Here the numerator represents the effects of having more particles at the IP: N^+ and N^- represent the charge per bunch of the colliding beams, f represents the collision frequency, in Hertz. The denominator represents the effects of packing the particles into a smaller space: σ_x and σ_y represent the horizontal and vertical beam sizes, usually reported in microns (note that $\mathcal{L}$ is usually reported in $cm^{-2}sec^{-1}$, for comparison to particle physics cross-sections, reported in cm^2) H_D is the "disruption enhancement" factor, which arises from the fact that bunches of oppositely-charged particles will tend to focus one another at collision, and is dimensionless. The simplest mechanism to increase the luminosity of a collider is to increase the bunch charges and/or the collision frequency. Because a circular collider stores and re-uses the same bunches for several hours, even a modest number of stored bunches can yield collision frequencies of several megahertz, and luminosities of over $10^{32} cm^{-2} sec^{-1}$ have been achieved at $e^+ e^-$ circular colliders [2].

The difficulty with extending circular collider technology to higher energy $e^+ e^-$ collisions is the dramatic increase in synchrotron radiation losses. It is well known that the energy losses of an electron beam which is bent in a magnetic field scales as the fourth power of the electron's energy. This has two immediate consequences on an electron storage ring. First, the energy loss experienced by the beam grows dramatically with the desired CM energy, requiring ever more robust acceleration systems to re-accelerate the bunches on every turn. Second, the radiation power incident on the accelerator itself grows as the fourth power of the CM energy. In order to compensate these two effects, the collider's bending radius (and thus its

total size) must increase. A full optimization including construction and operation costs yields a canonical scaling law for $e^+ e^-$ colliders: the circumference must increase as the square of the CM energy [3].

The highest-energy $e^+ e^-$ circular collider in existence is LEP, at CERN. LEP's circumference of 27 kilometers is based on a final CM energy, after the LEP II upgrade, of 200 GeV. Scaling this to a 1 TeV CM energy, and assuming no dramatic improvements in technology, dictates a circumference of 675 kilometers. This is clearly not a feasible machine to build and operate.

Linear colliders, by contrast, do not attempt to re-use bunches. Each bunch is collided with a single bunch of opposite charge, and both are then discarded. Because there are no synchrotron losses to negotiate, the total size of the collider is dictated by the length of linear accelerator (*linac*) needed to achieve the desired CM energy, which scales approximately linearly with that energy. This results in a machine which is smaller, contains fewer components, and is easier to upgrade.

While linear colliders are simpler to design in terms of energy, they are genrally quite a challenge in terms of luminosity. As we shall see below, it is not feasible to build a linear collider with a bunch charge of more than a few times 10^{10} particles (compared to $1.5 - 2.1 \times 10^{11}$ for LEP and CESR [4, 2]), or a collision frequency of more than 200 Hz (compared to 90 kHz for LEP [4] and 2.7 MHz for CESR [2]). By contrast, the only existing linear collider, the Stanford Linear Collider (SLC), typically collides bunches with 3.5×10^{10} particles with a collision frequency of 120 Hz. Because of these limitations, linear colliders are forced to boost luminosity through dramatically reducing the bunch transverse size at the IP. Circular colliders are unable to arbitrarily demagnify their beams due to beam-beam focusing effects which degrade the quality of the beams after many collisions, an effect known as *tune-shift* or *disruption*. Because linear colliders do not re-use bunches, they are free to operate in a disrupted regime.

Figure 1.1 shows a schematic diagram of a typical linear collider. In order to

understand the mechanisms which determine the bunch size and luminosity of such a facility, it is instructive to consider briefly each of its susbystems.
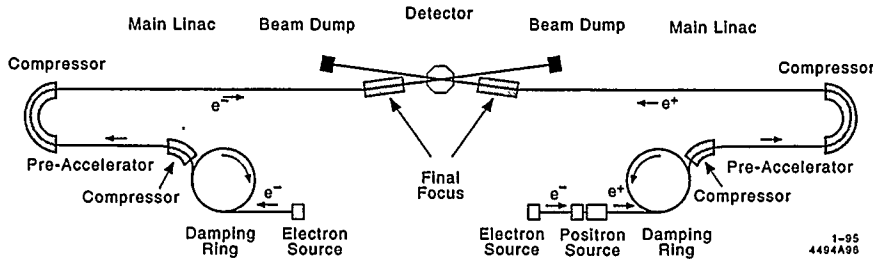
Figure 1.1: Schematic diagram of a linear collider.

1.2   Components of a Linear Collider

1.2.1   Electron and Positron Sources

The electron and positron sources for a linear collider can in principle be quite straightforward. The only demand upon these systems which differs from the sources for a circular collider is reliability. Because a linear collider is constantly injecting, the sources need to be available to deliver bunches at the full machine rate (up to 200 Hz) at all times. The sources are mentioned here primarily for completeness.

1.2.2   Damping Rings

The charged bunches from the electron and positron sources are then injected into separate storage rings, whose purpose is to reduce the transverse phase space volume, or "emittance" (see Appendix A) of the two bunches. This reduction is accomplished through synchrotron damping: particles passing through the bend magnets of the damping ring will emit synchrotron photons in their direction of travel, which reduces their total momenta; RF cavities along the ring straight sections accelerate the

particles back to their original momenta, resulting in a net reduction in transverse momentum. Damping rings are characterized by two parameters in each transverse plane: *damping time*, which is the characteristic time needed to reduce the emittance, and *equilibrium emittance*, the minimum emittance which the ring is able to produce.

The characteristics of the damping ring impact the final luminosity in several ways. Most obvious is that the equilibrium emittance of the damping ring will be the minimum emittance achievable at the IP, and hence is one of the limiting factors in the spot size. Additionally, the finite damping time of the ring limits the overall collision frequency at the IP, for each bunch must be stored long enough to reduce its emittance to the desired value. Finally, it is worth noting that damping rings typically produce beams whose horizontal emittance is 1 to 2 orders of magnitude larger than the vertical. For this reason, high-luminosity linear collider designs are generally expected to collide extremely flat beams.

1.2.3   Linear Accelerator

Once extracted from the damping rings, the charged bunches are then accelerated to collision energies by one or two linear accelerators. The acceleration is in the form of RF power provided by a series of klystrons, which is then transmitted to the beam as it passes through a disc-loaded wave guide. Because of the high acceleration gradients needed to reach the desired energy in a reasonable length of beamline, the klystrons are typically pulsed, emitting a high-intensity burst of RF power only when the beam is present. The limitations on the pulse rate attainable in the klystrons is a further limit on the collision frequency of the linear collider.

In addition to accelerating the bunches, the linear accelerator must also maintain the extremely small emittances from the damping rings to the final focus regions. This is made difficult by transverse wakefield effects in the disc-loaded wave guide, in which the electric field of the head of an electron bunch reflects off the discs and causes a transverse deflection of the tail. This effect causes the effective transverse
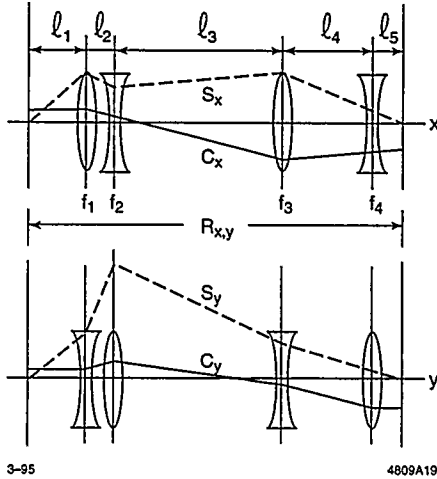
Figure 1.2: Schematic of a four-quadrupole telescopic system. The solid line through the center of the lenses is the reference trajectory. A particle which is initially travelling parallel to the reference but with a unit offset from it follows the cosine-like trajectory (C_x and C_y); while a particle which crosses the reference at the entrance of the system, with a unit angle, follows the sine-like trajectory (S_x and S_y).

emittance to increase; because wakefield emittance dilution is a strong function of bunch intensity, transverse wakefields in the linac limit the per-bunch charge of a linear collider.

Once accelerated to the desired energy, each bunch is then reduced in size by the final focus, and collided with a similar bunch of opposite charge at the interaction point. It is the final focus which is our main concern.

1.3 Linear Collider Final Focus Systems

The simplest possible formulation of a linear collider final focus system is a set of four quadrupoles which demagnifies the beam size and magnifies the beam divergence in each plane. Such a system has been described by Brown and Servranckx [5], and is shown diagramatically in Figure 1.2. Note that the cosine-like rays are parallel to the

reference trajectory at the beginning and the end of the system, and that the sine-like rays intersect the reference trajectory at the beginning and the end. A module with both such properties is said to have both *parallel-to-parallel* and *point-to-point* imaging. For a system with no aberrations, and given sufficient length and sufficiently strong quadrupoles, a four-quad telescope such as this is capable of reducing the beam size at the IP to any desired value.

The most immediate difficulty of such a system is that quadrupoles have inherent chromatic aberrations. The inverse focal length of a quadrupole, its integrated focusing strength, is given by:

$$K_q = \frac{B_0 L}{a B \rho},\tag{1.2}$$

where B_0 is the magnetic field at the pole-tip of the quad, L is the length, a is the aperture radius, and $B\rho$ is the familiar "magnetic rigidity":

$$B\rho(\text{T}\cdot\text{m}) = \frac{E}{q_e c}(\text{mks}) = \frac{E(\text{eV})}{c(\text{m/sec})},\tag{1.3}$$

which is linearly proportional to energy. We see from eq. 1.2 that the focal length of each quadrupole increases linearly with energy. Hence, at lowest order, particles which are low in energy will be over-focused (focal point upstream of nominal), while particles high in energy will be under-focused (focal point downstream of nominal). As a consequence of this phenomenon, if a bunch of electrons contains a finite energy spread, then the off-energy particles in the bunch will be out of focus at the nominal Focal Point (FP), leading to an increase in the spot size.

Let us define $K(s) \equiv \frac{B_0}{aB\rho}$ to represent the instantaneous quad strength at location s; the quantity K_q is the integral of $K(s)$ over the length of a given quadrupole magnet. We can use the formalism of Roy[6] and define the *chromaticity* of a beam line:

$$\xi_{x,y} = -\frac{1}{2}\int_{s_0}^{s_1} K(s)\beta_{x,y}(s)ds.\tag{1.4}$$

The precise definition of the betatron functions, $\beta_{x,y}$, in a telescopic transport system is discussed in Appendix A, along with a general review of adapting storage ring

notation to final focus optics. Suffice it to say here that the linear beam size scales with the square root of the betatron function. The chromaticity is a quantitative measure of the change in focusing strength of a beam line with the energy of the particles being focused. It can be shown (in Appendix A) that the focused size of a beam with finite energy spread, after passing through a chromatic transport line, is:

$$\sigma^{*2} = \sigma_0^{*2}\left[1 + \xi^2 (\frac{\sigma_E}{E})^2\right],\tag{1.5}$$

where σ^* is the measured beam size at the IP, σ_0^* is the linear monoenergetic spot size given by $\sigma_0^* = \sqrt{\epsilon\beta^*}$, and $\frac{\sigma_E}{E}$ is the rms energy spread of the beam.

Let us consider a linear collider final focus system modelled on the telescope of Figure 1.2, in which the Focal Point (FP) is on the left hand side of the figure. In this system, the offsets of particles which enter on the right are demagnified at the FP, and the angles of particles which enter on the right are magnified at the FP by the same factor; no mixing occurs (ie, initial angle is not converted into final offset, as required by the definition of point-to-point and parallel-to-parallel imaging). In a linear collider final focus, the FP linear beam size is reduced as much as possible, which is equivalent to increasing the beam divergence to as large a size as possible. This means that, on average, the amplitudes of particles at the FP will be small and their angles large. The offset of a particle at any point in the system is given by:

$$x(s) = C_x(s)x(0) + S_x(s)x'(0),\tag{1.6}$$

and since the values of $x(0)$ tend to be small and $x'(0)$ tend to be large, the beam size at the quadrupoles will be dominated by the product of the sine-like ray and the FP angular divergence:

$$\sigma_x(s) \approx \theta_x^* S_x(s).\tag{1.7}$$

Let us now replace the sine-like trajectories from the FP to other points, S_x and S_y, by their equivalents in the TRANSPORT formalism of Brown [7], R_{12} and R_{34}, respectively. We can replace the FP divergence and the quadrupole beam size with

expressions involving the Twiss parameters in Appendix A:

$$\sigma_x(s) = \sqrt{\epsilon\beta_x(s)}, \theta_x^* = \sqrt{\epsilon\gamma_x^*},\tag{1.8}$$

and equivalently for the vertical plane. Finally, let us assume that the positions and angles of particles in the bunch are uncorrelated, which corresponds to $\alpha = 0$ in both planes. In this special case, we may write:

$$\gamma_x^* = 1/\beta_x^*.\tag{1.9}$$

We may then substitute into Equation 1.7 to obtain the following relation between the FP betatron function and the betatron functions at the quadrupoles:

$$\beta_q = \frac{R_{12}^2(\text{IP} \rightarrow \text{quad})}{\beta^*}.\tag{1.10}$$

We may now introduce this expression for β_q and rewrite equation 1.4 as:

$$\xi_{x,y} = \frac{1}{\beta_{x,y}^*}\int_{s_0}^{s_1} K(s)R_{12,34}^2(s)ds \equiv \frac{1}{\beta_{x,y}^*}\mathcal{I}_{x,y}.\tag{1.11}$$

Note that the values of $K(s)$ and of $R_{12,34}(s)$ are functions of the optics only, and not of the incoming beam; therefore the entire integral becomes an invariant property of the optics. We can then substitute into equation 1.5 to find:

$$\sigma^{*2} = \epsilon\beta^* + \frac{\mathcal{I}^2}{\beta^*}(\frac{\sigma_E}{E})^2.\tag{1.12}$$

For a given lattice of pure quadrupoles, therefore, the chromatic aberrations of the quads cause the beam size to go through a minimum at some value of β^*, and then to increase again as β^* decreases. It is assumed in this computation that β^* is varied by changing the incoming beam with matching quads upstream of Figure 1.2, and not with the magnets within the figure itself.

Figure 1.3 shows the beam size, including chromatic contribution, as a function of β^* for a typical linear collider final focus system. It is clear from this figure that elimination of the chromatic aberration is crucial to the task of minimizing the spot size and maximizing luminosity.
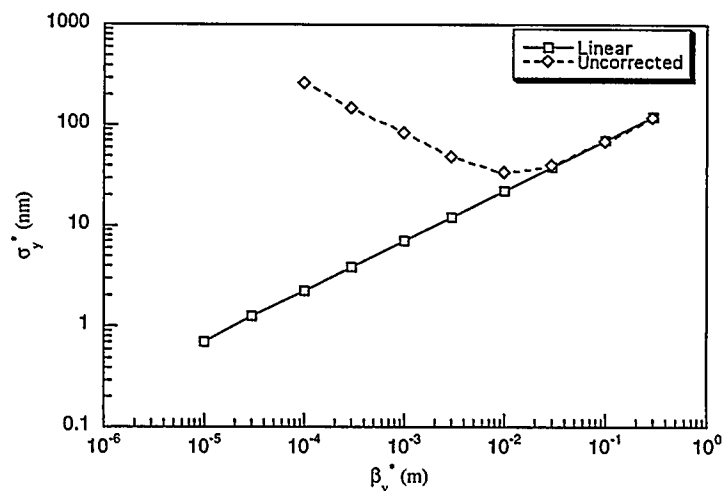
Figure 1.3: IP beam size as a function of β_y^* for a linear collider final focus. Both the linear and uncorrected chromatic sizes are shown. The other beam parameters are: $\epsilon_y = 5 \times 10^{-14}$ meter-radians, $\frac{\sigma_E}{E} = 0.003$.

1.3.1 Chromaticity Correction

In order to eliminate lattice chromaticity, it is necessary to introduce beamline magnetic elements which will produce energy-dependent focusing, and adjusting this effect such that it cancels the energy-dependent focusing of the quadrupole magnets. It is for this reason that sextupole magnets are introduced into the final focus lattice. Whereas pure quadrupole magnets have a magnetic field whose first derivative with respect to transverse coordinates is constant, pure sextupole magnets have a magnetic field whose second derivative is constant. The first derivative of the field varies linearly with transverse position across the sextupole aperture, and therefore the sextupole presents a linearly increasing quadrupole effect across its aperture. If the beam through the sextupoles can be manipulated such that the energies of the constituent particles and their transverse positions are also linearly correlated, then it is possible to globally cancel the energy-dependent focusing of the quadrupole magnets. The

high-energy particles are forced through a region of high gradient in the sextupole aperture, and the low-energy particles pass through a region of low gradient.

In order to produce a linear correlation between energy and transverse position at the sextupole magnet, it is necessary to introduce dipole bending magnets upstream of the sextupoles. Low energy particles are overbent by such magnets, and high energy particles underbent, resulting in the desired energy-position correlation. The trajectory of a particle with zero initial offset or angle, but unit energy offset, is called the *dispersive* ray, and its value through the beamline given by the *dispersion function*, denoted by $D_{x,y}(s)$ or $\eta_{x,y}(s)$. The dispersive ray must be carefully controlled and closed between the last sextupole and the FP, as a position-energy correlation at the FP will add in quadrature with the monochromatic spot and increase its size.

It is also worth noting that a monoenergetic beam which passes through such a system will also see differential focusing due to the sextupole. This geometric aberration of the sextupole can in practice be an even larger effect than the chromaticity of the quadrupoles. This geometric sextupole aberration is eliminated by placing the sextupole magnets in pairs separated by a $-I$ transform (a telescope with unity magnification, which inverts the image). By placing bend magnets upstream of the first sextupole and within the $-I$ module, the dispersion function is the same at the two sextupoles, and the chromatic aberrations add while the geometric aberrations cancel. By adjusting the dispersion and betatron functions, it is possible to produce a lattice in which the sextupole chromaticity ($\xi_s = \sum K_s \beta_s \eta_s$, where $K_s = \frac{B_0 L_s}{a^2 B \rho}$) cancels the quadrupole chromaticity. This scheme for chromaticity cancellation is also described by Brown and Servranckx [5]. This delicate cancellation of both the sextupole aberrations and the overall chromaticity results in tight tolerances on the alignment, strength, and field quality which are characteristic of linear collider final focus systems.

Figure 1.4 shows the beam size as a function of β^* for a linear collider final focus system with and without chromatic correction (essentially, the same system
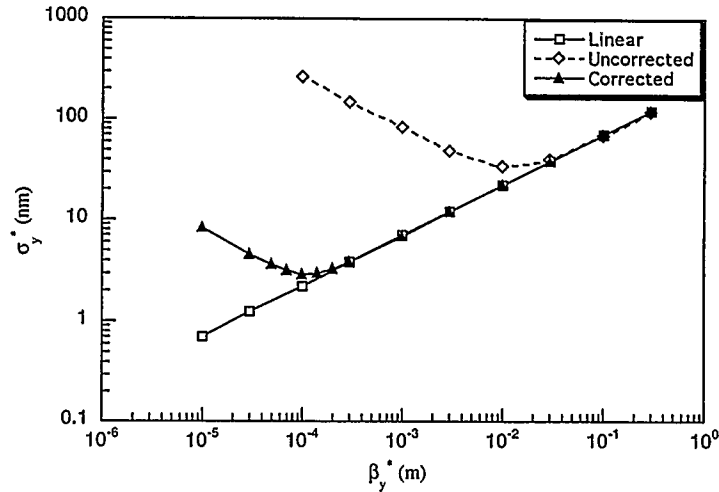
Figure 1.4: Beam size as a function of β_y^* for a linear collider, showing both chromatically corrected and uncorrected beam sizes. Beam parameters are the same as for Figure 1.2.

with sextupoles on and off). Note that even with chromatic correction, the spot size diverges from the linear monoenergetic size at some threshold. This is due to other aberrations which are not corrected.

The arrangement of strong quads, strong sextupoles, and bend magnets has been successfully implemented in one working collider, the Stanford Linear Collider (SLC). A brief examination of the optics of this system is instructive in understanding the trade-offs involved in final focus design.

1.4 The SLC Final Focus

The Stanford Linear Collider consists of a single linac in which both positrons and electrons are co-accelerated; a pair of achromatic arcs which steer the beams into a more or less head-on approach; and a pair of final focus systems which produce the small spots desired at the IP[8]. The optical arrangement of the SLC Final Focus is
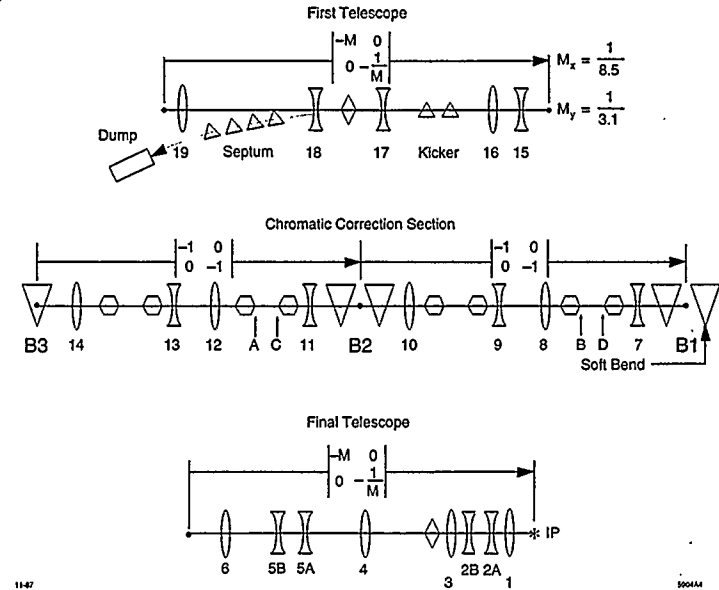
Figure 1.5: Optics of the SLC Final Focus. These are the original optics, which have since been upgraded in several ways; however, the basic principles of the system are unchanged.

shown schematically in Figure 1.5.

Note that the chromatic correction sextupoles are interleaved within the CCS section: while the sextupoles for correcting horizontal and vertical chromaticity are in pairs separated by a $-I$ transform, the first horizontal sextupole is placed between the two vertical sextupoles. This was necessary because of space constraints of the SLAC site, which limited the total length available to the final focus regions. In addition, the original operation of the SLC utilized equal horizontal and vertical emittances, in which situation the aberrations caused by the two sextupole pairs interacting with one another are not significant.

Currently, the SLC operates with unequal emittances in the two planes. In this configuration, the interleaved sextupoles have been identified as one of the aberrations
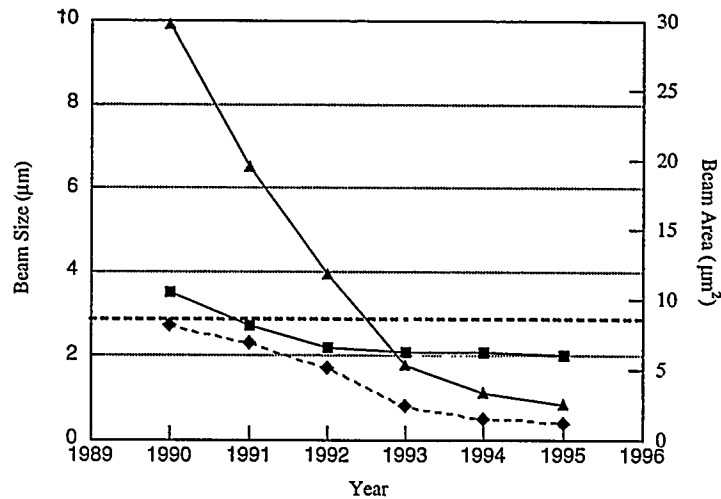
Figure 1.6: Beam size as a function of time at the SLC IP. Both horizontal size (squares) and vertical size (diamonds) are shown. The beam transverse area is also indicated (triangles); the dashed line represents the SLC design value for transverse area. Plot data courtesy of N. Phinney.

which limits the performance of the SLC Final Focus [9]. Nonetheless, in recent luminosity runs the SLC Final Focus have produced electron and positron beams with rms sizes of 2 microns in the horizontal and 400 nanometers in the vertical [10]. The time evolution of the SLC IP spot size is shown in figure 1.6.

In order to produce the desired luminosity, TeV-scale linear colliders will need to focus beams at the IP to a few hundred nanometers in the horizontal, and no more than a few nanometers in the vertical. Such beam sizes and aspect ratios will require final focus systems with the minimum achievable aberration content, and with specialized diagnostics capable of measuring and correcting errors in such a small, flat beam. These are the issues for which the Final Focus Test Beam has been constructed.

Chapter 2

Description of the Final Focus Test Beam

As we saw in Chapter 1, the difficulties inherent in making the small spots needed for high-luminosity linear colliders are formidable. Equation 1.10 shows that reducing the spot size at the IP requires enlarging it in the quadrupole magnets. This means that the tolerances on magnet strength, stability, alignment, and aberration content become arbitrarily tight as the beam size at the FP decreases. This becomes especially true as the sextupole strength needed for chromatic correction increases, and the rigid cancellation of sextupole aberrations and first-order aberrations from sextupole alignment come into play. To make matters worse, the extremely unequal beam sizes from emittance and beamstrahlung considerations places additional tolerances on all of these phenomena, as well as on incoming coupling and magnet roll (defined as rotation about the longitudinal axis).

In other fields of endeavor, when a new and challenging product is being contemplated, it is customary (and necessary) to build a prototype which addresses the new features and challenges of the product and exposes its potential weaknesses, in order that these weaknesses may be corrected prior to the start of production. This is such a useful approach that we have gone ahead and applied it to the "industry"

Parameter (units)	SLC FF Actual	FFTB Design	NLC FF Proposed
Beam Energy (GeV)	45.6	46.6	250–750
Energy Spread (%)	0.15	0.3	0.3
$\sigma_x^* \times \sigma_y^*(\mu m \times nm)$	2.0×400	1.7×60	0.25×2.5
$\beta_x^* \times \beta_Y^*(mm \times \mu m)$	6.7×2800	10.0×100	10×100
Demagnification	72	380	380
$\gamma\epsilon_x$(meter · radians)	6.0×10^{-5}	3.0×10^{-5}	5.0×10^{-6}
$\gamma\epsilon_y$(meter · radians)	6.0×10^{-6}	3.0×10^{-6}	5.0×10^{-8}
Aspect Ratio	5	28	100
Bunch Population	3.5×10^{10}	1.0×10^{10}	$(0.75 - 1.0) \times 10^{10}$
Repetition Rate (Hz)	120	30	120–180

Table 2.1: Comparison of IP beam parameters for SLC Final Focus, FFTB, and NLC Final Focus.

of linear collider final focus design. The Final Focus Test Beam (FFTB) is designed to be a scaled version of a TeV-scale linear collider final focus. The goal of the FFTB is to produce focused beams of size $\sigma_x^* \times \sigma_y^* = 1.7$ $\mu m \times 60$ nm, corresponding to $\beta_x^* = 10$ mm, $\beta_y^* = 100$ μm. Table 2.1 gives the relevant parameters of the SLC Final Focus, the FFTB design, and the expected NLC design. Note that the FFTB's vertical demagnification is identical to that of the NLC Final Focus, and that the betatron functions are also identical. This was done to ensure that the FFTB's aberration content and tolerances will match as closely as possible those of the NLC Final Focus. The design aspect ratio is an intermediate value between the SLC's (which was unity at the time of the FFTB design) and the NLC's (which is 100). This parameter determines the tolerances on cross-plane effects such as magnet roll, sextupole vertical alignment, and roll angles in diagnostic devices such as beam position monitors and wire scanners.
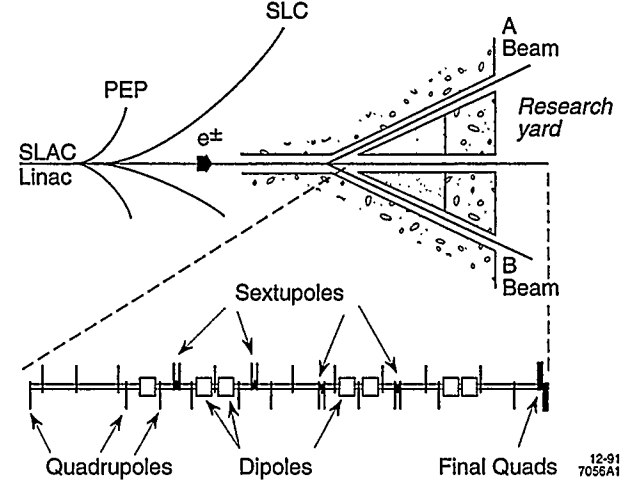
Figure 2.1: Location of the FFTB with respect to the existing SLAC beam lines: SLC arcs, PEP injection, and End Station A/B lines.

The original design for the FFTB indicated a value of β_x^* of 3 mm, which yields a σ_x^* of 1 μm and an aspect ratio of 16. The value of β_X^* was changed to 10 mm for two reasons. First, the expected IP parameters of the NLC, which originally indicated a 3 mm β_x^*, have evolved towards a larger horizontal spot. Second, as shown in Equation 1.10, smaller IP size requires larger size in the quadrupole magnets and consequently more danger of backgrounds generated by beam particles impacting ("scraping") on the magnet apertures. The FFTB's horizontal apertures were regarded as dangerously tight at the smaller FP size, and so the horizontal demagnification was relaxed in order to relieve backgrounds from beam scraping in the final doublet magnets.

Figure 2.1 shows the positioning of the FFTB in the straight-ahead ("C-channel") beamline, which is nominally colinear with the linac. The FFTB's total length exceeds that available in the original C-channel housing, and thus it was necessary to construct a shielded enclosure for the remaining beam line. This housing extends for 75 meters into the SLAC Research Yard (RSY), next to End Station A. The total length of the

FFTB is 350 meters, and the distance from the first optical element to the FP is 185 meters. The FFTB uses the electron beam as delivered to the SLC IP, with relatively minor adjustments. Because of synchrotron radiation losses in the SLC arcs, the linac energy is 1.0 GeV above the SLC IP energy, and thus the FFTB energy is 46.6 GeV and not 45.6 GeV.

2.1 Optical Modules

The Final Focus Test Beam can be divided into five optical regions, or modules, with distinct functions and optics. Each of these modules is described briefly below. The optics and tolerances of the modules have been examined in detail by Roy[6], and shall not be overmuch discussed here.

2.1.1 Beam Switch Yard (BSY)
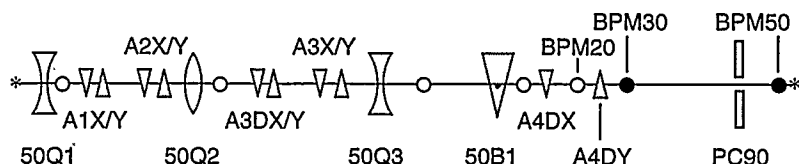
Figure 2.2: Schematic layout of the Beam Switch Yard. Normal quadrupoles are shown as lenses, bend magnets as large wedges, steering magnets as upright (xcor) or inverted (ycor) small wedges. Pre-existing BPMs are shown as open circles, FFTB BPMs as closed ones.

The SLAC Beam Switch Yard is not technically an optical module of the FFTB; however, the BSY contains both FFTB-specific and pre-existing components which have an impact on FFTB operations. These are shown in Figure 2.2. The elements of principal interest are: the quadrupole triplet which is common to SLC and FFTB; the SLC bending magnet 50B1, which must be degaussed at the beginning of FFTB runs; BPMs 30 and 50, which reconstruct the incoming position and angle of the beam in x

and y; corrector pairs A3DX/Y and A4DX/Y, which are used by the launch feedback to correct the orbit measured by BPMs 30 and 50; and PC-90, a fixed collimator which pre-dates the FFTB and which defines the *de facto* line from the BSY into the FFTB, usually called the FFTB's "Massimo Criterion" [1]. Each of these elements shall be discussed in greater detail in later sections, and are shown and described here to orient the reader.

2.1.2 Beta Match Region
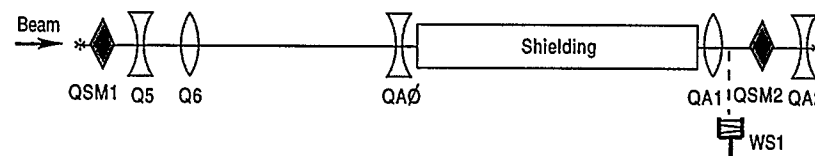
Figure 2.3: Schematic layout of the Beta Match region. Notation is as before, with the addition of diamonds to represent skew quadrupoles. Also shown are the locations of the beam reconstruction wire scanner, WS1, and the 16 meter muon shielding wall which permits access to FFTB during SLC running.

The first optical module of the FFTB proper is the beta matching region. The magnets in this region are used to match the parameters of the incoming beam to those of the desired beam at the FP, or elsewhere in the FFTB line. As shown in Figure 2.3, the beta matching region consists of five normal quadrupoles, and two quadrupoles rolled by 45° ("skew quadrupoles"). While primarily used for adjusting to changes in the incoming beam's Twiss parameters, these quads also allow considerable freedom in changing the overall demagnification of the system: betatron functions from 1 meter × 1 meter to the design have been produced at the FP. Because of the tremendous flexibility of the beta match region, it is possible to use these quads for all adjustment of the incoming beam needed for tuning. This in turn allows us to

[1]Named for M. Placidi of CERN, who resolved a problem with a similar aperture in the SLC Final Focus.

tune the magnets in the chromatic correction sections to their design values and leave them fixed, so that changing the beam size at the IP never requires changing these quads and potentially losing the cancellation of sextupole aberrations. In addition, the two skew quads allow independent adjustment of two of the coupling degrees of freedom of the incoming beam. As discussed in Appendix A, there are only two degrees of freedom which affect the vertical beam size at the FP; since the vertical size is so much smaller than the horizontal, it is much more susceptible to coupling dilution. The two coupling terms which affect vertical focused size are therefore the ones we typically seek to correct.

Finally, in order to match the incoming beam parameters properly it is first necessary to measure them. This is done by measuring the beam size on a wire scanner, known as WS1, which is located in the beta match. In order to perform the measurement (described in Chapter 3), it is necessary to focus the beam to an RMS size of a few microns in each plane on WS1. For this measurement, therefore, the beta match quads are tuned to a special optics, and the beam is stopped before the first CCSX sextupole.

2.1.3   Horizontal Chromatic Correction Region – CCSX
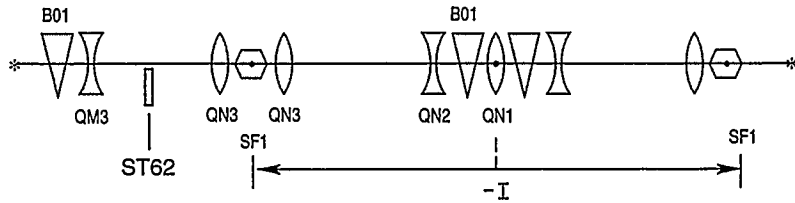
Figure 2.4: Schematic layout of the CCSX region. Chromatic Correction sextupoles are indicated by hexagons. Also shown is the movable stopper, ST62, which is inserted for incoming beam reconstruction.

Figure 2.4 shows the arrangement of magnets in the CCSX region. The first bend magnet and subsequent quadrupole provide the required dispersion at the first sex-

tupole; the five quadrupole magnets between the first and second sextupoles create the geometric $-I$ transformation, while the bend magnets on either side of QN1 are used to set the dispersion equal at the two sextupoles. During normal operations, the betatron functions in both x and y pass through a minimum in the center of QN1, and the horizontal betatron functions are at a maximum in the center of the SF1 sextupoles; $\alpha_x = \alpha_y = 0$ at both of these locations. In addition, the horizontal dispersion is at a maximum at the two sextupoles, corresponding to $\eta'_x = 0$.

In order to correct primarily the horizontal chromaticity, the optics are adjusted such that $\beta_x \gg \beta_y$ at the SF1's, and the phase advance from the sextupoles to the FP is $\frac{\pi}{2}$ in both planes. The phase advance between the center quadrupole, QN1, and the FP is an integer multiple of π. In order to correct the chromaticity without introducing unwanted sextupole aberrations, the stability tolerances on quadrupole and sextupole strengths in the CCSX are between 5 and 10 parts in 10^4. These tolerances, and many others, have been computed in exquisite detail by Roy [6]. A brief discussion on the meaning of tolerances is in Section 2.3.3.

2.1.4   Beta Exchanger
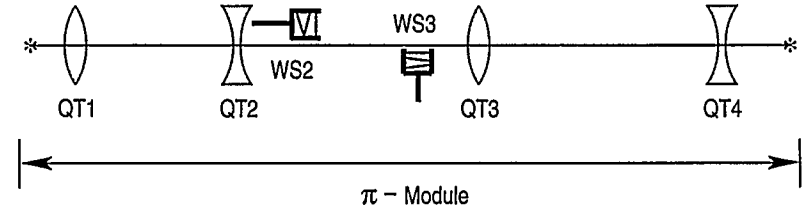
Figure 2.5: Schematic layout of Beta Exchanger region. Because of strength limitations, the "QT2" magnet is in fact a pair of quadrupoles set at the same strength with a separation of only a few centimeters. The optics contains a horizontal waist at the WS2 location and a vertical waist at the WS3 location.

In order to couple the CCSX sextupoles primarily to the horizontal chromaticity, the betatron functions were adjusted such that $\beta_x \gg \beta_y$ at the SF1 sextupoles. In order to correct primarily vertical chromaticity in the subsequent CCSY region, it is necessary to reverse this relationship. This reversal is the responsibility of the Beta-Exchanger, a canonical "four-quadrupole telescope" of the type shown in Figure 1.2. The phase advance from the beginning to the end of the module is π in both x and y, the horizontal magnification is 0.395, and the vertical magnification is 6.15. These magnifications are not possible for a strict four-magnet design, given the specifications of the FFTB standard quadrupoles (see Section 2.4) and the length available for the module; consequently, the "QT2" magnet has been implemented as a pair of magnets, QT2A and QT2B, which are placed a few centimeters apart and run at the same strength.

The beta exchanger is terminated at either end by the CCS sextupoles, which are $\frac{\pi}{2}$ out of phase with the FP, and the beta exchanger itself has $\phi_x = \phi_y = \pi$. Consequently, there must exist locations in the beta exchanger at which the beam is in phase with the FP, and $\frac{\pi}{2}$ out of phase with the sextupoles. In the case of the beta exchanger, the horizontal and vertical FP Images are separated by 2.85 meters longitudinally. As a result, the beam at WS2 is focused in x but large in y, while the beam at WS3 is large in x but focused in y. During normal running, with $\beta_x^* \times \beta_y^* = 10$ mm $\times$ 100 μm, WS2 $\beta_x = 8.3$ cm, while WS3 $\beta_y = 2.5$ cm. With the nominal emittances, this yields a beam size at WS2 of $\sigma_x, \sigma_y = 5.0$ μm, 99 μm, and at WS3 of $\sigma_x, \sigma_y = 171$ μm, 0.87 μm. These are rather extreme aspect ratios to measure, and we shall see in Section 2.5.2 how such measurements are made. Additionally, the spot sizes at WS2 and WS3 are potentially dominated by the chromaticity generated in the CCSX, and WS3 has non-zero horizontal dispersion; consequently, some care must be taken in using and interpreting the WS2 and WS3 beam size measurements. Note that WS2, the horizontal waist, is also a point of zero dispersion.
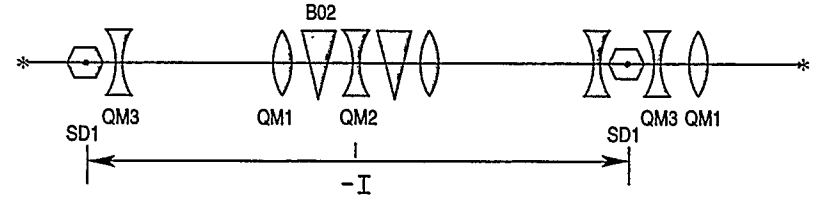
Figure 2.6: Schematic layout of CCSY region. The CCSY is optically identical to the CCSX, with the exception that the dipole and quadrupole polarities are reversed from one to the other.

2.1.5 Vertical Chromatic Correction Region – CCSY

After the beam has been adjusted in vertical and horizontal size by the beta exchanger, it enters the CCSY. Here the large vertical betatron functions ensure that the SD1 sextupoles will primarily affect the vertical chromaticity. This separation of the CCS sections allows the two chromaticities to be tuned independently by simply adjusting the strengths of one pair or the other. The CCSY is constructed identically with the CCSX, with the exception that all magnet polarities are reversed: horizontally focusing quads in the CCSX become horizontally defocusing quads in the CCSY, and vice versa, and the bend magnets bend to the geographic South in the CCSX and to the geographic North in the CCSY. The quadrupole polarity change has an interesting side effect: although $\eta_x' = 0$ at the SD1 magnets, the actual maximum of η_x occurs in the center of the QM1 magnets. Like the CCSX, $\alpha_x = \alpha_y = 0$ in the sextupoles and the center quadrupole.

Because the vertical betatron functions are so large in the CCSY sextupoles, the tolerances here are significantly tighter than in the CCSX. Magnet strength stability tolerances go as low as 1.7 parts per 10^4.

2.1.6 Final Transformer and Focal Point Region

Once the beam has passed through the last CCS sextupole, all that remains to be done is to restore the horizontal dispersion to zero and focus the beam down to the
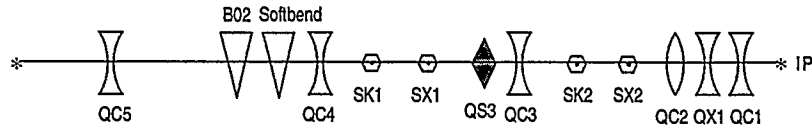
Figure 2.7: Final Transformer between the CCSY and the Focal Point. Small hexagons are sextupoles are for correction of residual geometric sextupole aberrations in the line.

the smallest spot possible. This is all accomplished in the Final Transformer. The dispersion is cancelled by an additional bend magnet identical to the CCSY bends, and a soft bend (B03 by name) which introduces a small angle between the beam and the synchrotron radiation from the preceding hard bend. This small angle ensures that the synchrotron radiation and the electron beam are spatially separated enough to mask off the synchrotron radiation before the FP. Three horizontally defocusing quadrupoles enlarge the beam in x before it enters the Final Doublet. The Final Doublet does the main work of reducing the beam size at the FP, and is also the primary source of chromatic aberration in the FFTB. Note that the FFTB's Final Doublet actually consists of three physical magnets. This is done because of aperture problems: in order to make the final lens strong enough, its aperture radius is a mere 6.5 millimeters; this is too small to give adequate clearance to the incoming beam, which has just passed through the QC2 magnet and is quite large in the vertical. Thus, a short, strong magnet (QX1) with a larger aperture is used to "pre-focus" the beam to a size which is capable of passing through the aperture of the main vertically-focusing quadrupole, QC1. The Final Transformer also contains 4 geometric sextupoles, which may be used to eliminate residual sextupole aberrations; such aberrations could come from imperfect cancellation of the CCS sextupole aberrations, or from small imperfections in the construction of the quads. There is also a skew quadrupole to compensate the effects of small roll angles in the Final Doublet.

Figure 2.8 shows the positioning of beam size measuring devices around the FFTB Focal Point. The design free length from the last quadrupole's downstream face to
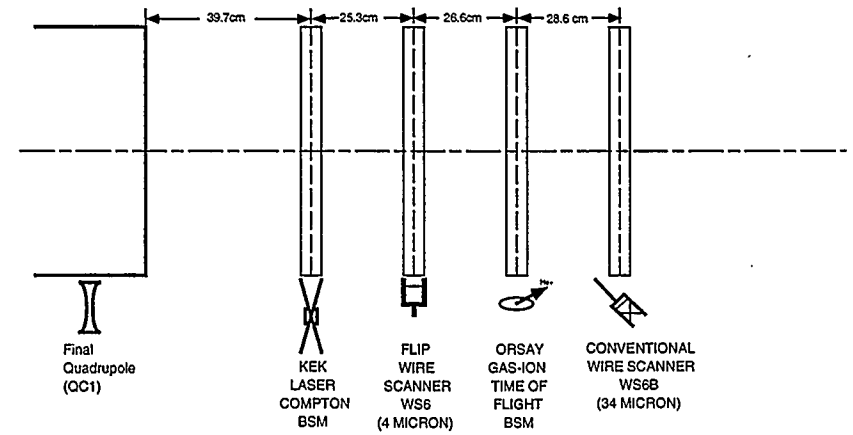
Figure 2.8: Arrangement of diagnostic devices at the FFTB Focal Point.

the nominal FP (l^*) is 40 centimeters, significantly smaller than the 1.5 – 2.5 meters planned for the NLC. The Laser-Compton Beam Size Monitor (Section 2.5.4) is set at this location. Approximately 25 centimeters downstream of this location is a wire scanner with 4μm carbon fibers, capable of measuring horizontal and vertical beam sizes. Following this after another 26 centimeter is the Gas-Ion Time of Flight Beam Size Monitor (Section 2.5.4). The optics of the Final Transformer can be adjusted to put the waist at any of these locations. Finally, there is a conventional wire scanner (1 fork, 3 wires at $0°, 90°, 45°$ from horizontal, 34μm wires). This location is sufficiently downstream of the FP that the beam size here is totally dominated by the FP angular divergences. Since this region is a drift space, WS6B gives a model-independent measure of the FP angular divergence.

2.1.7 Extraction Line

Once the beam has been focused and measured, the only remaining difficulty is to dispose of it safely. This is a potentially daunting problem: because the beam is
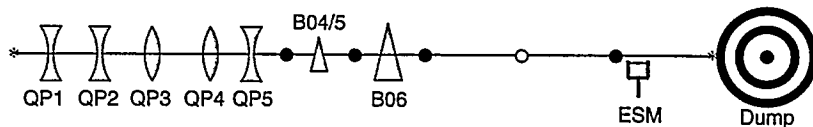
Figure 2.9: FFTB Extraction Line optics. The first three lenses are actually pairs of identical magnets. The vertical hard bends, B06, are permanent magnets which bend the beam down onto a water-cooled dump, providing vertical dispersion for energy measurement and energy spectrum measurement.

smaller at the FP than at any other point, its angular divergence is larger than at any other point. Consequently, the beam size changes quite rapidly as we move away from the Focal Point. The Extraction line captures the outgoing beam with an optical system pictured in Figure 2.9. The first three "lenses" are each made of a pair of identical quadrupole magnets, for reasons similar to the arrangement of the QT2A/B pair. The optics of the dumpline are designed to accept a beam whose focal point is at the Laser-Compton BSM (the first BSM in Figure 2.8), and focus this onto a dual waist at a vertical only wire scanner near the main dump. Between the last quad and the dump itself are a series of BPMs and vertical bend magnets: soft bends B04A/B and B05A/B, and permanent hard bends B06A–F. The vertical bend magnets bend the beam onto the dump, which is set below the level of the main beamline in order to back it up with a long strip of iron for muon attenuation. These bend magnets also provide significant vertical dispersion at the dump's location. This allows the BPMs in the dumpline to resolve energy variations of each pulse, and this information is used by the SLC linac to provide energy feedback. In addition, the wire scanner at the location of the last BPM is at a geometric waist; consequently, the beam size at this scanner is dominated by the product of energy spread and dispersion. Thus the scanner can be used as an energy spectrum monitor (hence the name ESM).
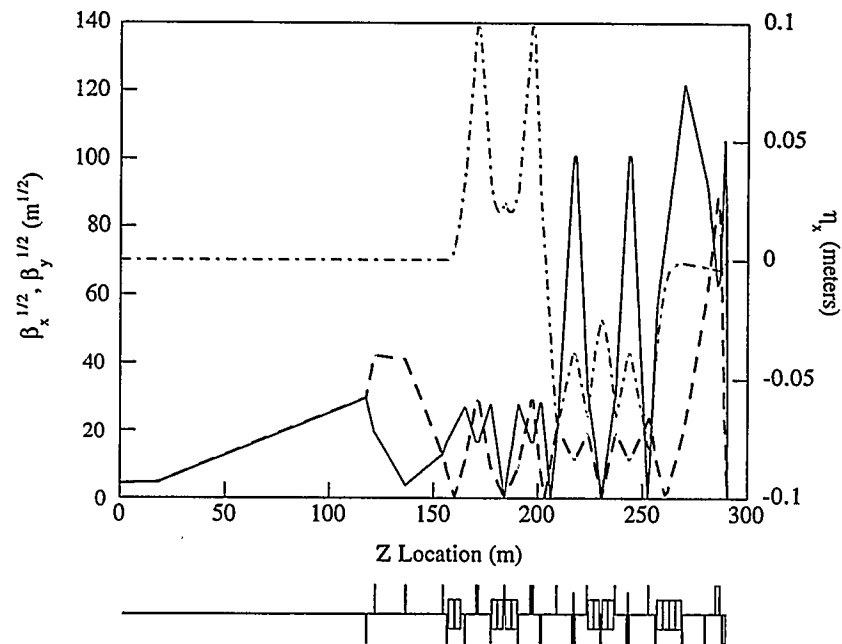
Figure 2.10: FFTB beam optical functions. Shown are $\beta_x^{1/2}$ (dashes), $\beta_y^{1/2}$ (solid), and η_x (dot-dash). The vertical dispersion function, η_y, has a design value of zero everywhere.

2.2 Optical Functions

Figure 2.10 shows the optical functions from the 50B1 magnet in the BSY to the FFTB FP.

2.3 Aberrations

As mentioned in Chapter 1, the first and most significant aberration encountered in a linear collider final focus system is chromaticity, which is corrected by the introduction of sextupoles and dispersion. Once this correction is applied, there are additional aberrations which come into play. These include tunable corrections due to misaligned

or mispowered elements in the beam line, and untunable corrections which come from high-order fabrication errors and other sources.

2.3.1 Chromaticity

As we saw in Figure 1.4, the chromaticity cancellation of a lattice like the FFTB's can be characterized by comparing the linear beam size, the uncorrected (chromaticity-dominated) size, and the corrected size. Figure 2.11 shows such a comparison for the FFTB, at the design energy spread of $\frac{\sigma_E}{E} = 0.3\%$.
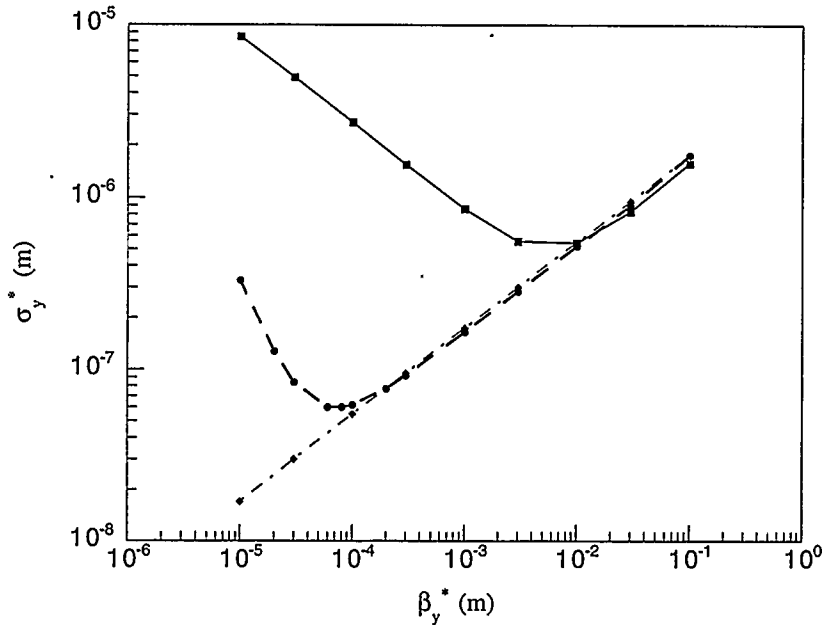
Figure 2.11: Chromatically-corrected (dashes) and -uncorrected (solid) beam sizes in the FFTB as a function of β_y^*. The linear monochromatic size is shown (dot-dash) for comparison.

Figure 2.11 shows that, at the design β_y^* of 100 μm, the linear spot size and the achievable spot size have begun to diverge, due to higher-order uncorrected aberra-

tions. The achievable spot size, approximately 60 nanometers, is approximately 10% larger than the linear size (54.7 nanometers).

2.3.2 Uncorrected Aberrations

There are several aberrations which are inherent to the FFTB design and cannot be eliminated through tuning of the existing magnets. These aberrations are: the long sextupole effect; the chromatic breakdown of the $-I$ transforms; synchrotron radiation losses due to the CCS bend magnets; and synchrotron radiation losses in the quadrupole magnets themselves, especially the final lenses ("Oide Effect"). These effects have each been developed in detail by Roy [11], and shall be covered only briefly here.

Long Sextupole Effect

The presence of strong, non-zero length sextupole magnets gives rise to third- order (octupole-like) aberrations. A thin slice of sextupole magnet gives a kick to the electrons which pass through it which is proportional to the square of their distance from the center of the magnet. This kick will slightly change the position of the electrons as they enter the subsequent slice of the sextupole, in a non-linear fashion. When these interactions between slices are summed over the total length of the real magnet, the resulting transformation to the electron's position and angle are equivalent to a thin sextupole coupled with a thin octupole magnet. This effect contributes a 6% increase in σ_y^{*2}, which corresponds to a 3% increase in the spot size.

Chromatic Breakdown

The cancellation of the geometric sextupole aberration due to the CCS sextupoles relies on the presence of a $-I$ transform between the sextupoles. For off- energy particles, however, the $-I$ transform is no longer rigorous, due to the chromaticity of the CCS quadrupoles themselves. While this term is itself small, and does not

give rise to a large sextupole kick, it damages the rigorous cancellation between the sextupole and quadrupole chromaticities. The result is that the spot at the FP is no longer chromatically corrected for all momenta. This aberration causes an increase in the spot size of about the same magnitude as the long sextupole effect.

Synchrotron Radiation in Bend Magnets

The bend magnets which provide the CCS sextupoles with dispersion also cause the electrons in the beam to emit synchrotron radiation. This results in the electrons losing energy between the entrance of the FFTB and the FP. While this is not crucial in and of itself, the amount of energy lost by each electron is determined by the stochastic process of synchrotron emission: there is no relationship (other than statistical) between a particle's value of δ at the entrance of the FFTB and its value of δ at the FP. As a result, the chromaticity introduced by the CCS sextupoles will not exactly cancel the chromaticity introduced by the final doublet, and some chromatic growth will occur. This effect has been estimated to add 6% to σ_y^{*2}, resulting in a further 3% increase in the spot size.

Note that the long sextupole and bend radiation effects must be traded off with one another to optimize the optics of the FFTB. Decreasing the strength of the bend magnets will decrease the synchrotron effect, but requires an increase in sextupole strength to maintain chromatic correction, thus exacerbating the long sextupole aberration. In this case, optimization of these effects within the constraints of the available site space was performed by Oide [12].

The Oide Effect

In addition to bend magnets, electrons undergo bending when they pass through strong quadrupole magnets. If the quadrupole is sufficiently strong and the beam sufficiently large in the magnet aperture, the bending of individual trajectories from focusing may be sufficient to measurably change the particle's energy offset from

Aberration	Source(s)	Knob Element(s)	Plane
x Waist	Quad Strength Sext x align	Final Doublet strengths	x
x Dispersion	Doublet Align Sext x align	Final Doublet orbit	x
x Chromaticity	CCSX Sextupole Strengths	CCSX Sextupole strengths	x
Geometric Sextupole	CCS Quad Strengths	FT sextupoles	x
y waist	Quad Strength Sext x align	Final Doublet strengths	y
y Dispersion	Doublet Align Sext y align	Final Doublet orbit	y
y Chromaticity	CCSY Sextupole Strengths	CCSY Sextupole strengths	y
x'y Coupling	Quadrupole Rolls	FT Skew Quad strengths	y
Geometric Sextupole(3)	CCS Quad Strengths	FT sextupoles	y

Table 2.2: FFTB aberrations and their primary dynamic sources.

nominal. This results in a spot dilution from the same process as the synchrotron radiation in the bend magnets, and is known as *the Oide Effect*. While not significant for the FFTB, this effect is likely to be a consideration in the design of future linear colliders at higher energies.

2.3.3 Corrected Aberrations

In addition to the four uncorrected aberrations listed in the previous section, there are a total of 11 aberrations which can dilute the FFTB spot size: four which affect principally the horizontal spot size, and seven which affect principally the vertical spot size. These aberrations are summarized in Table 2.2, along with their primary *dynamic* sources. The dynamic sources of aberration are those which are inclined to drift over time. This is a different category from *static* sources, which do not change. For example, machining errors in quadrupoles can give rise to geometric sextupole

aberrations, but these aberrations are virtually constant with time; consequently, any correction of this static source of sextupole aberration remains valid for all time. If, however, the CCS quadrupoles drift off the settings needed to maintain a $-I$ transform, this gives rise to an additional sextupole aberration which depends upon the time-stability properties of the power supplies.

Additionally, the aberrations listed in Table 2.2 include only those aberrations which arise primarily in the FFTB itself. The actual values of β_x^* and β_y^*, for example, could be set incorrectly, and this would result in changes in the FP beam size. Once this is set, however, the magnets of the FFTB are essentially unable to "drift" the magnification out of tolerance, because all the FFTB quadrupoles have tighter tolerances due to other strength-error aberrations (waist and $-I$ breakdown) than they do for total magnification. An additional aberration, xy coupling at the FP, can only arise from sources outside the FFTB itself.

Note on Tolerances

As shown in Table 2.2, the FFTB has been designed with global correction elements for each of the leading aberrations. These elements are able to cancel all contributions to their particular aberration, regardless of source, without introducing other aberrations in the process. As we shall see in Chapter 5, this orthogonality generally breaks as the knobs approach their full strengths. In addition, such knobs cannot be made arbitrarily powerful; as a result, a beamline whose elements are too badly misaligned and mistuned cannot be "beaten into submission" by repetitive application of the knobs.

For all classes of elements in the FFTB, therefore, we can define *capture tolerances*: these are the tolerances within which the elements must be tuned before global knobs are applied, in order to assure that the knobs will reduce the spot size within a finite number of iterations. These tolerances are generally a function of the design and arrangement of the knobs, and of the optics as well. The capture tolerances for the

Horizontal Alignment of Q and SX	100 μm
Vertical Alignment of Q and SX	30 μm
Strength Error of B, Q, and SX	0.1%
Roll Angle Error of B, Q, and SX	0.5 mrad
Incoming β mismatch	100%
Incoming α mismatch	100%

Table 2.3: Capture Tolerances for FFTB. Tolerances quoted are RMS values of gaussian-distributed errors. Bends (B), quadrupoles (Q) and CCS sextupoles (SX) have been included in the simulations.

FFTB have been determined by Oide through a series of simulations [12], and are summarized in Table 2.3.

Once the beamline has been tuned to the smallest possible spot size, another set of tolerances becomes operative. These are the *stability tolerances*, and they indicate how far a given element may drift from its post-tuneup value before enlarging the beam size by a given amount. The determination of what constitutes "a given amount" has a tremendous bearing on the difficulty of the problem posed by stability tolerances. For the FFTB, the permitted amount of beam size growth through drift of beamline elements has been set at 2% per aberration, through a process which is described in Chapter 5. Thus, at any given time after tuning the global knobs, we expect that the FFTB FP beam size may be as much as 14% above the design value in the vertical, and 8% above the design value in the horizontal.

The prescription for translating the "14% solution" [13] above into stability tolerances for individual elements, and then into classes of elements, has been a source of some confusion in the past. This is because, while we speak of growth due to individual aberrations, the aberrations are caused by drifts in magnet positions and power supplies, and a given drift may cause several different aberrations (for example, a CCS quad supply may cause a waist shift and a $-I$ breakdown) – essentially, the

aberrations and the magnets which cause them form two different bases which describe the same configuration space of the accelerator. It is important to remember, however, that the positions and strengths of the magnets are the elements which we can control directly, not the aberrations caused by changes in the magnets.

Every magnet in the FFTB has a finite set of parameters which may vary from the ideal: strength, horizontal and vertical alignment, roll, multipole content. Each of these aberrations may give rise to spot size growth through one of the aberrations listed in Table 2.2. It is possible to go through the FFTB optics and determine, for each magnet, how much of each deviation will result in a 2% spot size growth from each aberration which is generated, and this has been done for the FFTB [6]. However, since the magnets are assumed to all vary from the ideal in an incoherent manner, some way of combining the tolerances to form a group tolerance must be determined.

If we define the tolerance on a group of aberrations (for example, vertical vibrations of the standard quads, or power supply ripple of the standard quads) to be T_g; and we define the individual tolerances within the group to be t_i; then the amount of spot size growth caused by all of the magnets varying incoherently by T_g is given by:

$$\Delta\sigma^2 \equiv (\sigma^2 - \sigma_0^2) = 0.04\sigma_0^2 \sum_{i=1}^{N} \frac{T_g^2}{t_i^2}. \qquad (2.1)$$

If I require that the total spot size growth from this source be no more than 2%, then my defining equation for T_g is:

$$\frac{1}{T_g^2} = \sum_{i=1}^{N} \frac{1}{t_i^2}. \qquad (2.2)$$

At this point we have a set of values, T_g, which tells us the RMS tolerance for a class of objects (in which the objects in the class are assumed to be essentially identical, for example a set of identical magnets on identical power supplies and identical physical installations) for a given aberration. If we now have multiple classes which can give rise to the same aberration (for example, quadrupole vibration and bend power supply ripple), we can assign to each class a fraction of the 2% total allowed aberration, f_g.

This assignment is based on a judgement of the relative difficulty presented by the various T_g values. At this point, we have a set of tolerances, $T_g\sqrt{f_g}$, which, if met, will guarantee that the "14% solution" is also met. This is known as the beamline's *tolerance budget*.

2.4 Magnet Specifications

2.4.1 Normal Quadrupoles

The FFTB standard quadrupole is a warm iron magnet, 46.1 cm long with a 1.15 cm aperture radius. These magnets are capable of achieving an integrated gradient, $\int G \cdot dl$, of 388 kilogauss, which corresponds to a pole-tip field of 9.7 kilogauss. The maximum current required for this field is 240 amperes. All normal quadrupoles from the beginning of the line to QC4 are of this design, as are QP1A and QP1B in the extraction line.

Because the beam becomes quite large in the final telescope and the extraction line, it is necessary to have magnets with larger apertures in these locations. A variation of the standard quadrupole, with 46.7 cm length and 1.75 cm aperture radius, was used for QC3, QP2A, and QP2B. These magnets have a maximum $\int G \cdot dl$ of 273 kilogauss at 325 amperes, which corresponds to a pole-tip field of 10.2 kilogauss. A larger magnet, with 1 m length and 2.7 cm aperture radius, is used for QP3A, QP3B, QP4, and QP5. This magnet design has a maximum $\int G \cdot dl$ of 400 kilogauss at 500 amperes, for a pole-tip field of 10.8 kilogauss.

The group tolerance on power supply stability for the quadrupoles up to QC3 is $\frac{\sigma_I}{I} = 7.3 \times 10^{-5}$[15]. This tolerance is determined by shifts in the waist position for all magnets, as higher-order effects (such as geometric sextupole) do not appear at this level. The power supplies used for these magnets have achieved stabilities close to $\frac{\sigma_I}{I_{max}} = 3 \times 10^{-6}$ [16]. This implies that all quadrupoles which run at a current greater than 5% of the maximum will achieve their tolerance. In practice, most of the

Name	Length (m)	Aperture Radius (cm)	Max $\int G \cdot dl$ (kG)	B_{0max} (kG)	I_{max} (amperes)
QC2	2.026	2.60	880	11.3	500
QX1	0.310	1.00	456	14.7	250
QC1	1.120	0.65	2213	12.8	250

Table 2.4: Parameters of the FFTB Final Doublet magnets.

supplies run above 45% of maximum current, and all run above 7%. Each quadrupole upstream of the FP is driven by an independent power supply and controller. All of the quads in the extraction line are driven by independent power supplies, though some of these share controllers.

The group tolerance on alignment stability due to steering at the IP is approximately 750 nanometers in the horizontal and 200 nanometers in the vertical [17], while the group tolerance due to mis-steering through the CCS sextupoles and final doublet (causing additional quadrupole or dispersion, respectively) is approximately 300 nanometers in the horizontal and 200 nanometers in the vertical. The FFTB magnets are supported on ANOCAST artificial stone monuments coupled to the ground by concrete grouting, and no vibration measurement has shown an RMS motion greater than 100 nanometers [18]. The group tolerance on roll angle for all quads up to QC3 is 40 microradians [19].

The three quadrupoles which comprise the final doublet have such differing requirements that each of the three comprises a unique design [14]. Table 2.4 shows the specifications of these magnets.

The power supply stability requirements for each of QC2 and QC1 are 2 parts per 10^5, more stringent than even the group tolerance on all other quadrupoles, while for QX1 it is near 7 parts per 10^5 [15]. Because of the tight regulation capabilities of the standard 250 ampere supply used for the rest of the FFTB line, the standard supply was also used to power QX1 and QC1. As shown above, this supply easily meets the stability requirements for QC1. The QC2 magnet requires a 500 ampere power

supply, which is far beyond the capability of the standard (250 ampere) supply. The QC2 supply also demonstrates stability at the level of 3 parts per 10^6[16].

Vibration tolerances for the doublet are determined by the limits on allowed FP steering. The simplest model – that the three quads act as a single parallel-to-point lens in each plane – is adequate to determine that motions of the doublet translate to steering at the FP with a unity ratio; therefore, the vibration tolerance is equal to 1 σ in the plane of interest. The three magnets are mounted on a common, vibration-isolated table[20], with typical RMS vibrations measured to be between 35 and 70 nanometers [18]. While this is adequate in the horizontal, the vertical vibrations are not within the allowed tolerances, resulting in steering of the beam at the FP. The consequences of this are discussed in subsequent sections. The roll tolerance of the doublet as a system is 33 microradians [19].

The lowest-order multipole field above the quadrupole is the sextupole, and therefore this aberration is the first critical one for quadrupole magnets. The aberration content for sextupole field is usually quoted as $\frac{B_{sext}}{B_{quad}}$ at 70% of aperture radius. For all quadrupoles from Q5 to QC3, the RMS tolerance for this ratio is 5.6×10^{-4}, and for the doublet magnets it is about 3.7×10^{-5}. These tolerances, and the measurements of magnet sextupole content, are discussed at length in Appendix B. Note, however, that these tolerances are expected to be essentially static, in that the sextupole content of the quads is not expected to vary overmuch with time. Moreover, in the initial design phases, there was no provision for geometric sextupole magnets in the FFTB. The FFTB has, however, been constructed with such magnets. Consequently, even significant failure to meet this tolerance can be corrected with the use of the global sextupole magnets.

2.4.2 Bend Magnets

The six B01/B02 bend magnets displayed in section 2.1 have been fabricated as 12 distinct magnets, in order to reduce the length of individual units for ease of handling.

Each magnet has an effective length of 2.62 meters, with a pole-tip field of 2.144 kilogauss and a bend angle of 3.618 milliradians. Thus the six B01 bend magnets bend the beam 21.7 milliradians to the South, and the six B02 magnets bend it an equal amount North. The B03 magnet, also fabricated as a pair of magnets, are identical except for a slightly larger gap (to accomodate the divergent beam) and consequently a slightly larger effective length. The B03 pair each have a pole-tip field of 256 gauss, yielding a total bend angle of 434 microradians.

The current needed to drive the B01/B02 string at the design value is nearly 330 amperes. In this case, all twelve of B01/B02 magnets are driven by a single power supply which is identical to the QC2 supply, as the voltage needed to push 330 amperes through 12 long bend magnets is considerable. The power supply stability tolerance is determined by steering errors in the sextupoles, and is just under 1 part in 10^5[21]. As discussed in the previous section, the QC2 supply meets this tolerance. Note that the twelve magnets have different excitation curves, and thus a single current will not drive all of the magnets at the design value. Fortunately, each magnet has a trim winding which may be used to fine-adjust its delivered field. Each pair of bend magnets has a single 12-ampere trim winding power supply; these are discussed in Section 2.4.4. The B03 magnets require only 71 amperes to reach their design value. Because these magnets are located after the last CCS sextupole, their tolerance for strength is quite loose; the magnets are powered by a standard 240 ampere power supply, which has stability far superior to what is required for this purpose.

The rotation tolerance for the bend magnets, again determined by orbit transformations at the sextupoles, is approximately 13 microradians, which is the tightest roll tolerance in the FFTB [22].

2.4.3   CCS Sextupoles

The four sextupole magnets used for chromatic correction are each 0.25 m in length, with a 1.0325 cm bore radius. The sextupoles reach a maximum "sextupole strength"

($\int S dl$, where $S \equiv \frac{d^2 B}{dr^2}$) of 35,500 kilogauss/meter at a current of 325 amperes, for a pole-tip field of 7.6 kilogauss.

The FFTB sextupoles are powered by separate 333 ampere supplies, but each pair (SF1's and SD1's) shares a common controller. Thus while each sextupole is independently regulated and stabilized, the pairs are always set to the same current. This implies that the pairs of sextupoles must have quite identical excitation functions at the operating point(s), and this issue is discussed in Appendix B. The tolerances on sextupole strengh regulation are quite loose, over 1 part per 10^3 for each of the four sextupoles [15]. The sextupole power supplies are a variation of the standard FFTB supply (supplying 333 amperes with 30 volts, rather than 250 amperes with 40 volts), which provides a stability almost 3 orders of magnitude greater than required for this application. Because the sextupoles are placed on mounts identical with the quadrupoles, and because the primary spot-size effect of the quadrupoles is missteering in the sextupoles, the CCS sextupole alignment tolerances were included in the treatment of quadrupole alignments [23]. In other words, if the RMS vibration of *all* the quadrupoles and the CCS sextupoles is under 300 nanometers in x and 200 nanometers in y, the spot size growth will be 2%.

2.4.4   Other Magnets

The FFTB contains a total of 73 small magnets, which are driven by 12 ampere bipolar power supplies. These include: three skew quadrupoles (QSM1 and QSM2 in the Beta Match, QS3 in the Final Transformer); four sextupoles (normal sextupoles SX1 and SX2, skew sextupoles SK1 and SK2, all in the Final Transformer); an assortment of DC steering correctors in x and y; six bend trim winding supplies; the B04 and B05 magnets in the extraction line; several sets of quadrupole trim windings which are wired to act as dipole correctors (on QN1, QT3, QM2, and QC5; QP1A/B, QP2A/B, QP3); and independent trim windings on each pole of each doublet magnet (total of 12). The specifications for these power supplies require that each one be stable at

the level of 1 part per 10^4 at full current, which should be sufficient for most of these applications[2].

2.5 Instrumentation

Sophisticated diagnostics are an essential component of the tuning of the FFTB, and these in turn require reliable, flexible instrumentation. The FFTB's primary beam-measurement devices are a set of high-resolution beam position monitors (BPMs), and a cluster of wire scanners and profile montiors for measuring the transverse distributions of the beam. In addition to these, the FFTB makes use of a new generation of remote-controlled magnet movers, and two novel beam size monitors (BSMs)[3].

2.5.1 Beam Position Monitors

All of the normal quadrupoles between Q5 and QC2, and also QP1A, contain stripline BPMs. In addition, two FFTB-style BPMs and one older model are installed in the BSY for launch feedback; three FFTB-style units are installed in the Beta Exchanger for pulse-to-pulse jitter correction of the beam measured at WS2 and WS3; one FFTB-style unit is installed in the end of the CCSY for a similar purpose (the wire scanner in question was not installed); and the extraction line contains four BPMs similar to the standard design (and one low-resolution device salvaged from elsewhere) for pulse-to-pulse energy reconstruction and feedback. Figure 2.12 shows the cross-sectional profile of the standard BPM. Note the indentations in the vacuum extrusion necessary to fit the BPM into a quadrupole aperture.

The specifications for the BPMs require: pulse-to-pulse resolution of the beam position of 1 μm at a bunch charge of 1×10^{10}; linearity of 1% at a beam offset of

---

[2]A possible exception is the trim windings on the bend magnets.
[3]These should not be confused with *beamstrahlung monitors*, which are also abbreviated BSM.
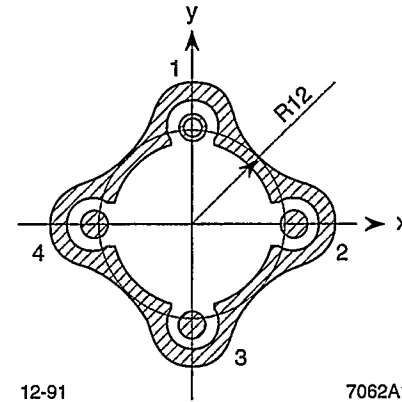
Figure 2.12: Cross-section of the standard FFTB BPM. The stripline radius, R12, is 1.15 cm, identical to the quadrupole aperture radius.

2 mm [24]; and a scale factor accuracy of 0.1%. Bench-test calibrations and beam-based measurements of compliance to these requirements is presented in Appendix C and Chapter 4, respectively.

The energy feedback BPMs have a larger aperture and shorter stripline length; consequently, their resolution is reduced from the standard design. These are required to resolve 6 μm pulse-to-pulse [36].

2.5.2 Wire Scanners and Profile Monitors

The FFTB Wire scanners are described in detail elsewhere [26], and their physical designs are summarized in Table 2.5.

The first scanner, WS1, is primarily used in reconstructing the incoming beam, and for this purpose must measure beam spots which are as small as a few microns. The scanner uses 3 separate yokes, one for each direction of measurement. The scanner can measure x, y, and v[4], and thus allow the beam's major and minor axes, and roll angle, to be reconstructed by geometry considerations.

---

[4]the u and v axes are rotated 45° clockwise from, respectively, the x and y axes.

Name	Purpose	Wire Size	Wire Angle
WS1	Incoming Beam Reconstruction	34 μm and 7 μm carbon	0°, 90°, 135° 3 separate yokes
WS2	β_x, α_x	4 μm tungsten	0°, ±3°, 0°, ±6°
WS3	β_y, α_y	4 μm tungsten	0°, ±0.7°, 0°, ±1.4°
WS6A	FP tuning (initial)	4 μm carbon	0° × 10, 90° × 10 (2 yokes)
WS6B	FP divergences	34 μm carbon	0°, 90°, 45°
ESM	Energy Spectrum	75 μm tungsten, 175 μm Be-Cu (24)	90°

Table 2.5: Arrangement of FFTB Wire Scanners. Note that WS3 wire angles are angles with respect to the horizontal, for measurement of the vertical.

The Beta Exchanger scanners, WS2 and WS3, are required to measure small spots with large aspect ratios. In this case, the primary problem is signal: the charge density along the long axis of the beam is so low that a wire stretched to measure in this direction never intercepts a significant amount of beam, and the signal is quite low. In order to remedy this situation, the scanners at these locations are equipped with wires set at shallow angles relative to the short-axis wire. This allows each wire to produce a readable signal, although as a result the odd-angle wires produce beam sizes which are comparable to the minimum size. This makes distinguishing the wires problematic, especially if one breaks or cannot be found. Furthermore, note that the WS3 angles are on the order of 1°. This requires that the yoke's rotation angle (relative to the symmetry planes of the magnets) be known to about 0.1°. In the case of WS3, an additional problem arises: the beam size at WS3 with nominal emittances is $\sigma_y = 800$ nm. A wire scanner whose wire is diameter d can reasonably measure a beam size whose RMS is $d/4$ before the wire size dominates the measurement [27], which for WS3 yields 1.0 μm.

The FP tuneup wire, WS6A, uses a pair of stationary forks with 10 wires each, horizontal or vertical (depending on the fork). These are inserted into the beam and scanned by stepping an upstream corrector. This has the action of moving the beam across the wire (as opposed to all other scanners in the FFTB, which

scan the wire across the beam via a stepping motor). The conversion of corrector strength in kilogauss-meters into position at the FP in microns is dependent upon the model loaded into the control system, which is subject to errors. Fortunately, the wires on a given yoke are known to be 50 μm apart. Thus, as part of any experiment involving WS6A, the beam is scanned across several wires to cross-check the calibration constant. Finally, it is worth noting that WS6A wires are at risk of breakage by the beam. SLC experience indicates that 4 μm carbon wires will be broken by a beam of 1×1^{10} population if $\sigma_x \times \sigma_y \geq 3$ μm^2[28]. During early FFTB experiments, several WS6A wires were in fact broken (including a sequence of 6 wires on a yoke broken in a single scan). Since then, however, a technique has been developed which allows use of WS6A wires for tuning, even when the beam is capable of breaking wires. This technique is described in Chapter 5.

The FP divergence wire, WS6B, is used to measure beams which are dominated by FP angular divergence. They are consequently quite large, and 34 μm wires are adequate for measurement. The wires are on a single yoke, a design which is in ubiquitous use in the SLC linac.

The beam in the dumpline is dominated by vertical dispersion (60 cm at the ESM, resulting in a beam size of 600 μm per 1×10^{-3} energy spread). The Energy Spectrum Monitor can therefore use a large wire and step size, which completely eliminates any danger of wire damage by the beam. In addition to the scanner, a set of 24 foils below the wire can be inserted into the beam path indefinitely; the signals from these foils allow continual monitoring of the energy spectrum.

2.5.3 Magnet Movers

Because alignment of FFTB quads and sextupoles is a critical aspect of tuning and maintaining a small spot, it is necessary to have the magnets installed in a fashion which permits them to be moved to aligned positions quickly, precisely, and repeatedly. Conventional accelerator designs require the magnets to be moved to aligned

positions by work crews. This in turn requires access to the accelerator housing and deactivation of the magnets, which invites thermal changes to shift the magnet unpredictably. The alternative to this procedure is to develop remote-controlled magnet movers which may be used to align the accelerator under run-time conditions.

A first-generation magnet mover was used in the SLC Collider Arcs to position the combined function arc achromat magnets [29]. These movers were single degree-of-freedom devices, which depended upon a rotary potentiometer to determine the magnet position. For the FFTB, however, both x and y positioning are required (roll is also desired), as well as a position-detection system which couples directly to the magnet, not to the stepper motors or camshafts.

The FFTB Magnet Mover is described in [30] and [31]. The mover uses 3 stepper motors, each attached off-center to a camshaft. The magnet is supported by a pair of V-blocks which ride on the camshafts. When the motors turn the shafts cause the blocks to rise and fall, and shift left and right. While no single motor drives a single degree of freedom (x,y,roll), it is possible to combine the stepper-motor actions to produce pure motions in these coordinates. In addition, the magnet position is determined by a set of three LVDTs which are in physical contact with the magnet. The LVDT resolution is approximately 0.3 μm, allowing a position reconstruction resolution of better than 1 μm at all times.

When the magnet mover is commanded to move, both the LVDTs and the rotary potentiometers are read out, to determine the current locations of both the cams and the magnet. The necessary movement of each cam is then computed and implemented, and the LVDTs are read again. If the magnet has not reached its target within 1 μm in each of x and y, and 1 μrad in roll, the computation and motion is iterated until convergence. In this way, the mover provides positioning accuracy of better than 1 μm over its full range of ± 2.0 mm (x), ± 1.5 mm (y), ± 5.0 mrad (roll). The testing and calibration of the movers is described in [31]. Each normal quadrupole upstream of the final doublet, each CCS sextupole, and QP1A in the extraction line

are mounted on such movers.

2.5.4 Beam Size Monitors

As indicated in section 2.5.2, beams which are micron-sized in both planes at FFTB intensities will destroy 4 μm wires through thermo-mechanical stresses. The FFTB's FP beam will be on the order of tens of nanometers. Logic indicates that, even if 10 nm wires could be installed on a wire scanner, no such wire could stand up to the power density of the FFTB beam. Consequently, new technology is required to measure the focused beam at the end of the FFTB. Two such monitors have been designed and installed in the FFTB: a Laser-Compton Beam Size Monitor, developed at KEK, and a Gas-Ion Time-of Flight Beam Size Monitor, developed at LAL Orsay.

Laser-Compton BSM[32]

The Laser-Compton BSM splits a Nd:YAG laser pulse ($\lambda = 1.064\mu$m) and recombines it at the electron FP with a crossing angle θ. The laser then produces an interference pattern, with characteristic modulation wavelength $d \equiv \frac{\lambda}{2\sin(\theta/2)}$, as shown in Figure 2.13.

An electron beam which strikes the interference pattern will produce a burst of photons through Compton scattering, which can be detected downstream of the laser. The number of photons, N, which are thus scattered by a beam with vertical size σ_y is given by Equation 2.3:

$$N = \frac{N_0}{2}\left[1 + \cos 2k_y y \cdot \cos\theta \cdot \exp\left\{-2\left(k_y \cdot \sigma_y\right)^2\right\}\right], \tag{2.3}$$

where $k_y = \frac{\pi}{d}$, N_0 is an overall normalization constant dependent upon the laser and electron beam intensities, and y is the vertical distance between the electron beam and a 0° reference in the interference pattern.

Once the interference pattern described above has been established at the electron FP, the electron beam is measured by scanning the beam vertically across the
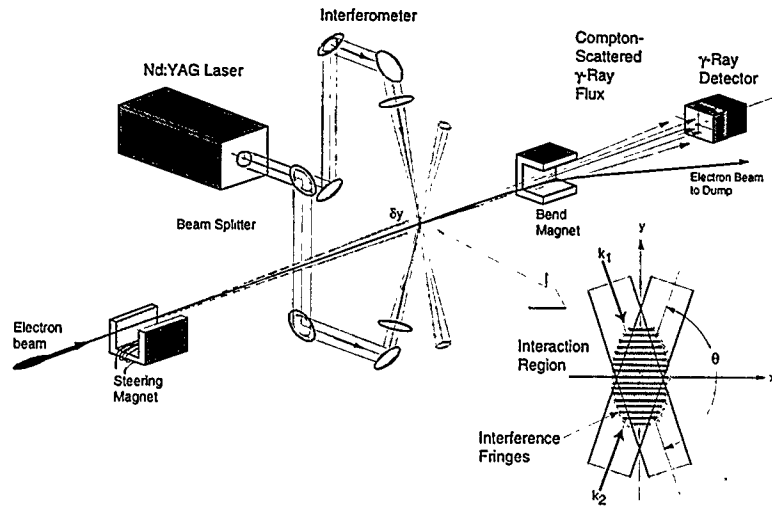
Figure 2.13: Operating principle of the Laser-Compton BSM. A laser is split and the two laser beams crossed at the FP, producing an interference pattern with a characteristic pitch determined by the laser wavelength and the crossing angle.

interference pattern, via an upstream corrector magnet. Equation 2.3 shows that N has a sinusoidal dependence on y, and that the amplitude of the oscillation is given by $\cos\theta \cdot \exp\left\{-2\left(k_y \cdot \sigma_y\right)^2\right\}$. By correlating the signal from the downstream photomultiplier tube with the electron beam position at the FP, and fitting the resulting sinusoid, the beam size can be extracted from the sinusoid's fit parameters. This is shown in Figure 2.14.

The Laser-Compton BSM is equipped to produce three different interference patterns. The mode of primary interest crosses the beams at an angle of 174°, resulting in a 533 nm fringe spacing. This mode can resolve beam sizes from 40 – 180 nm. The second mode ("Big Y") uses a 30° angle to produce a 2.1 μm spacing, for vertical beam sizes from 160 – 720 nm. The third mode uses a 6° angle transverse to the first two, producing a 10.2 μm spacing for measuring horizontal beam sizes from 0.76 – 3.4 μm. Although the attainable electron beam size at the FP is larger than the
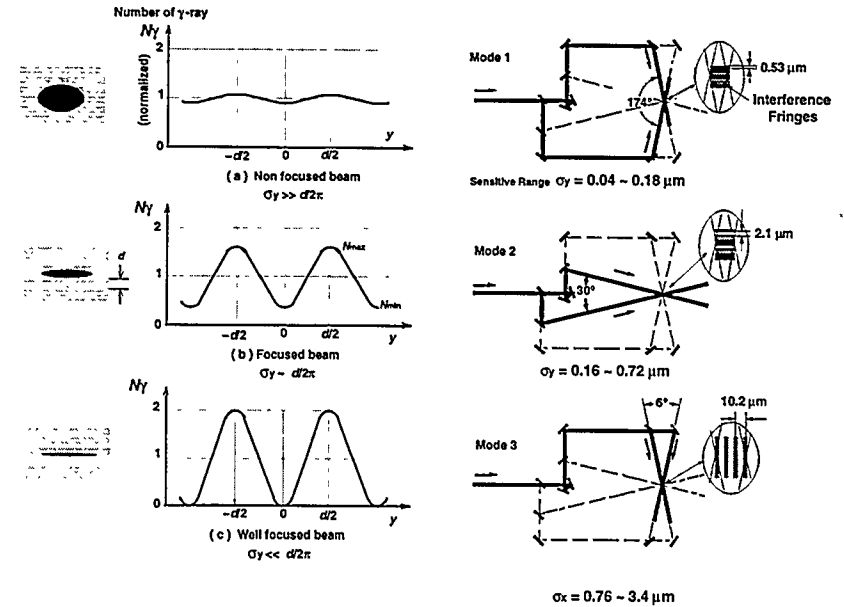
Figure 2.14: By correlating the number of Compton-scattered photons with the beam position, the beam size can be determined: smaller beams produce sinusoidal patterns with greater modulation depth than large beams (left). The Laser-Compton BSM is equipped to produce three interference patterns, in order to measure both horizontal and vertical beam sizes (right).

minimum size measurable by the 4 μm carbon fibers of WS6A, it is still necessary to measure the horizontal size with more exotic technologies, since the horizontal wire would still be destroyed by the high energy-density of the focused beam. The modulation depth as a function of crossing angle and beam size is shown in Figure 2.15. Note that the "Big-Y" mode achieves a maximum modulation of $\cos 30° = 0.866$.

Gas-Ion Time of Flight BSM[33]

The Gas-Ion Time of Flight BSM injects a small amount of Argon or Helium gas into the path of the electron beam just before it arrives. The high energy electron beam ionizes the gas, resulting in the presence of positively-charged ions in the vicinity of

Spot Size Response
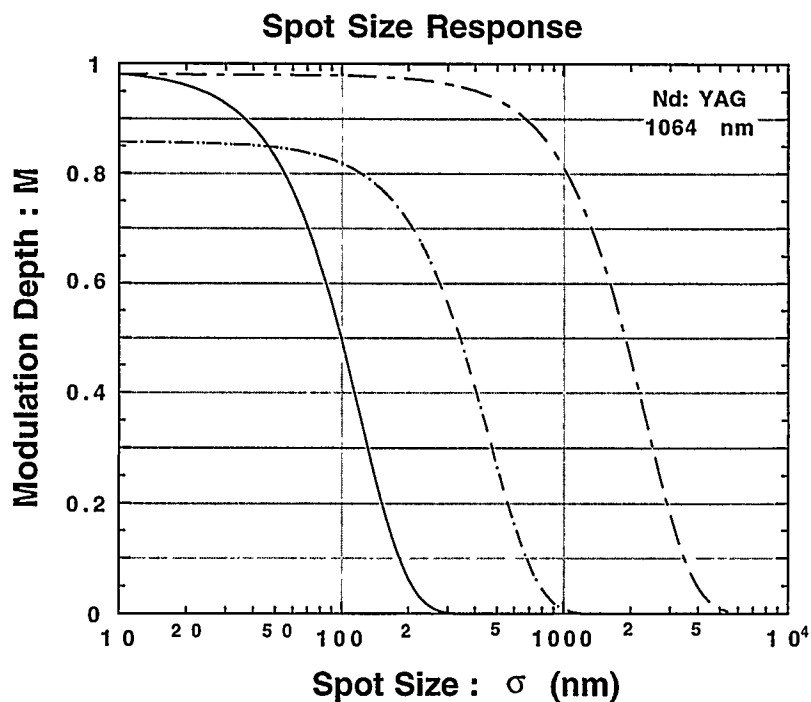
Figure 2.15: Modulation depth as a function of beam size for the three Laser-Compton BSM operating modes. Modes 1 (solid) and 2 (dashed) are the vertical measurement modes, while Mode 3 (dot-dash) is the horizontal measurement mode.

the negatively-charged electron beam.

The effect of the beam on the ions is twofold. For the heavier (Argon) ions, the ions are accelerated through the electron beam by its transverse electric field. The amplitude of the acceleration is proportional to the maximum electric field, which is inversely proportional to the major axis of the electron ellipse (in the xy plane). This results in a distribution of ions whose maximum velocity is inversely proportional to the major axis of the beam. The velocity, in turn, can be measured by measuring the time of flight of the ions from the FP to an array of microchannel PMT plates arranged in a ring around the FP, at a fixed distance, as shown in Figure 2.16. Thus,
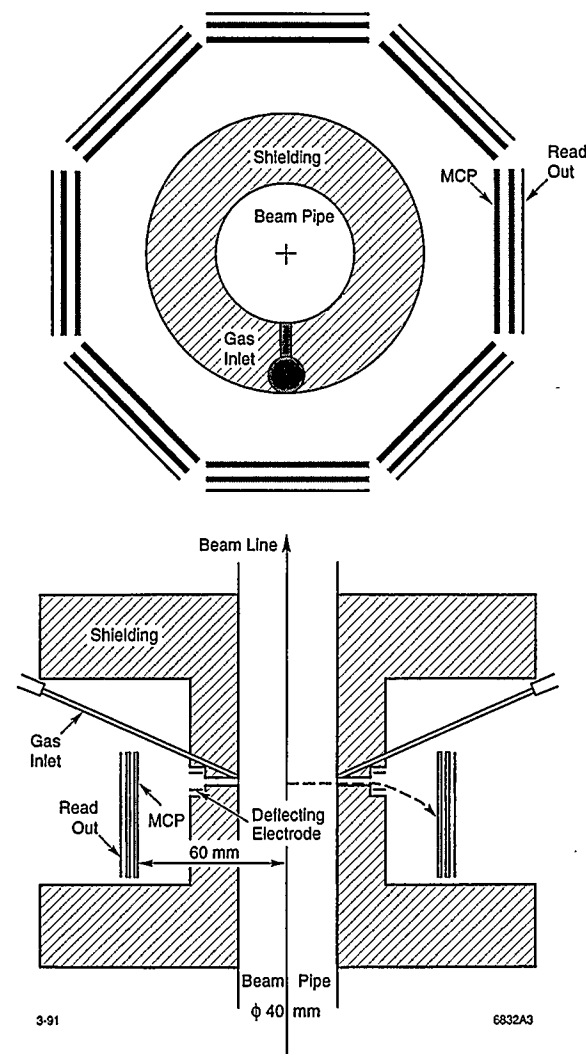
Figure 2.16: Two views of the Gas-Ion BSM: "beam's-eye" (top) and longitudinal section (bottom), showing the mechanisms for the injection of gas and the detection of ions. The deflecting electrode is for separation of singly- and doubly-charged Argon ions.
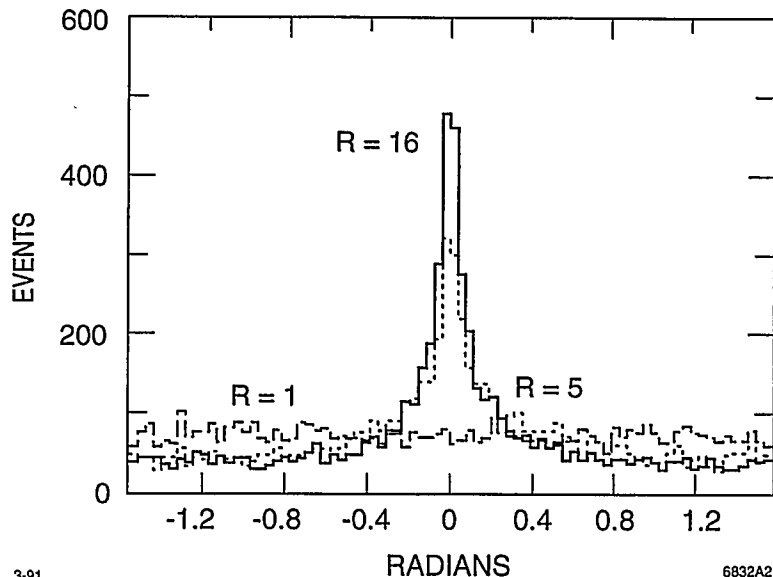
Figure 2.17: Simulated distribution of Helium ions for three different beam aspect ratios: R=1 (dashed), R=5 (dotted), R=16 (solid). In this system, an angle of zero corresponds to the horizontal direction. For all plots, $\sigma_x^* = 1$ μm, $\sigma_z = 500$ μm, and $N_e = 1.0 \times 10^{10}$.

the time of flight of the Argon ions provides a measurement of the larger of the two beam sizes.

The lighter (Helium) ions become trapped in the potential well of the electron beam and oscillate within it during the time it takes for the beam to pass through. Once the beam has exited the BSM, the ions will be emitted in the xy plane, with velocities proportional to their oscillation amplitudes. The larger horizontal amplitudes result in higher horizontal velocities, and consequently more ions reach the horizontal sensors than the vertical sensors. The aspect ratio of the beam is thus measured by the angular distribution of the Helium ions; this number, coupled with the major axis measured by Argon time-of-flight, gives the two sizes of the beam. In addition, the precise azimuthal angle of the peak ion detection determines the tilt angle of the beam

with respect to the BSM (information not available from the Laser-Compton BSM). Figure 2.17 shows the expected azimuthal distribution of Helium ions for beams with $\sigma_x^* = 1$ μm and different aspect ratios, as computed by a Monte Carlo simulation. Note that the horizontal axis of Figure 2.17 only covers half the circle; a full-circle plot is expected to provide two peaks, spaced halfway around the azimuth from one another, and of equal height.

Note that the Argon measurement is weakly affected by the aspect ratio of the beam, as the vertical size does influence the maximum electric field. However, the correction is a small one, and the Argon ions are far too heavy to oscillate in the potiential well of the electron beam – their total movement during the presence of the beam is quite small. A more significant correction to the Argon measurement is the separation of singly- charged ions from doubly-charged ones, accomplished by a deflecting voltage as shown in Figure 2.16. Finally, it is worth noting that the helium ions can also be used for the time-of-flight measurement; however, the precision of this measurement is limited, as the minimum time of flight for helium ions (60 nanoseconds) is only a few times the expected time resolution of the apparatus.

Chapter 3

Measuring and Matching the Incoming Beam

The Final Focus Test Beam is a magnetostatic system at the end of a 3 km. linear accelerator. A consequence of this arrangement is that many of the crucial beam parameters (emittance, energy, energy spread, etc.) cannot be varied by the FFTB itself, while others (coupling, Twiss parameters) must be measured and matched before the beam is introduced into the CCSX section. In this chapter, we examine the techniques by which the incoming beam is measured, and its parameters matched into those required by the FFTB.

3.1 Characteristics of the Incoming Beam

3.1.1 Emittance, Coupling, and Dispersion

The projected emittances (x and y) of the accelerated beam are measured by a series of multi-wire diagnostic stations in the linac itself: one just after the extraction from the damping ring (Sector 2); one near the center of the linac (Sector 11); and one near the end of the linac (Sector 28). Of the three stations, the Sector 28 measurement is

most crucial, as this is the measurement made closest to the FFTB itself.

The linac emittance stations use an advanced version of the standard linac multi-wire emittance algorithm [34]. Wire scanners are placed at several locations in the linac, with known transport matrices between them. By measuring the vertical and horizontal beam sizes at each of these locations, it is possible to extract the projected emittance and Twiss parameters in each plane. Because the beam has three degrees of freedom in each plane (ϵ, β, α), three wire scanner locations are sufficient to reconstruct the beam ellipse; the Sector 28 station contains 4 wire scanners, permitting an over-constrained solution.

During FFTB experimental runs (with bunch populations $N_e \sim 0.7 \times 10^{10}$), the measured normalized emittances at the Sector 28 station are typically $\gamma\epsilon_x = 2.5 \times 10^{-5}$ – 3.5×10^{-5}m · rad, and $\gamma\epsilon_y = 1.5 \times 10^{-6}$ – 2.5×10^{-6}m · rad. While the horizontal emittance has tended to be slightly larger than the design for FFTB, the vertical is almost always smaller. This leads us to expect an FP beam size smaller than the design value of 60 nm.

Note on Emittance Notation

SLAC currently uses two different standards for emittance notation, each of which is colloquially known as "SLC units." When *normalized* emittances ($\gamma\epsilon$) are quoted, the customary unit is 10^{-5}m · rad, while *laboratory-frame* emittances (ϵ) are quoted in units of 10^{-10}m · rad. Because normalized emittances are not changed by acceleration, measurements made in the accelerator itself are typically reported as normalized values. In beam-delivery areas, however, laboratory-frame emittances are more common, since these are the emittances which directly couple to beam size. These are the conventions which are used in this document as well. Note that the common "back of the envelope" conversion is to divide the normalized emittance by 10^5 to obtain the laboratory-frame values in the FFTB; however, this figure must be increased by 10%, as the relativistic γ for a 46.6 GeV electron beam is 91,400 and not 100,000.

Coupling and Dispersion in the Linac

While it is in general possible to use such a multi-wire system to measure xy coupling, in the case of the Sector 28 station a good coupling measurement cannot be made. This is because the phase advance between wire scanners is nearly equal in x and y, and consequently the different coupling phases cannot be unequivocally determined [35]. The coupling can be estimated using another device: a set of profile monitors in the last 100 meters of the linac (Sector 30). These profile monitors are equipped with fast kickers which "steal" pulses every few seconds, and their signals are digitized and transmitted continually to a set of dedicated monitors in the SLAC Main Control Center (MCC). The four "Decker screens"[1] provide single-pulse displays of the xy distributions of the beam at different phases, and this gives qualitative evidence that the linac beam is not coupled. As we shall see in Sections 3.2.3 and 3.2.4, the FFTB beam matching wires can fully resolve incoming coupling.

As discussed in Appendix A, there are experimental difficulties inherent in measuring the *residual dispersion* of a linear accelerator. This dispersion is estimated by comparing the projected emittance at the beginning of the linac with the same quantity near the end. As shown in Appendix A, the action of dispersion is to increase the in-plane beam ellipse terms, leading to a larger projected emittance. During FFTB running this growth is negligible, leading to the conclusion that all sources of dispersion in the linac are small. Of greater concern is dispersion arising from the BSY quadrupoles, and by any steering necessary to bring the beam from the linac into the FFTB. These effects were estimated during a 1992 experiment which brought the beam from the linac to D2 (a movable dump downstream of Q6) and found to be acceptably small [36].

---

[1]Named for F.-J. Decker, SLC physicist.

3.1.2 Position and Angle Jitter

Incoming position and angle jitter are translated through the FFTB optics to FP jitter. This raises problems for the Laser-Compton BSM, which relies on correlations between the beam position and the compton-scattered flux (see Section 2.5.4), and also implies difficulties for a future e^+e^- linear collider, which attempts to collide nanometer-sized beams at its IP.

Section 2.1.1 introduced the concept of the FFTB launch feedback. This system uses a pair of BPMs separated by 86 meters to reconstruct the incoming position and angle of the beam, and a pair of correctors in each plane to hold the BPM readings to some reference values. Assuming that the BPM errors are uncorrelated and that the intrinsic resolution of each BPM is 1 μm, the expected resolution of the position and angle at the first BPM is 1 μm and 0.016 μrad. Note that the primary function of the launch feedback is to correct slow drifts and decouple the FFTB steering from slow and/or DC effects upstream (such as quad strength changes, corrector failures, etc.). True compensation for stochastic variation on a pulse-by-pulse basis is not possible for a single-pass accelerator.

3.1.3 Energy Jitter, Feedback, and Spectrum

Because of the arrangement of the FFTB, the energy-related properties of the incoming beam are most easily measured in the extraction line, in which the vertical dispersion reaches over 60 cm. A cluster of BPMs and a wire scanner (ESM; see Section 2.5.2) provide information on the pulse-to-pulse jitter of the beam, and on the overall energy spectrum.

Because the energy-reconstruction BPMs are separated by drift spaces and/or vertical bend magnets, the $R_{12,34}$ matrix elements from one to another can be considered equal to the longitudinal separations between BPMs, and the $R_{11,33}$ terms are all unity. Such a system, with no quadrupole magnets between the BPMs, provides the simplest possible reconstruction of the beam energy, with the least possible de-

pendence on beam optics, but it is not the best. If we consider the first BPM to be the start of an optical section for reconstruction of the y, y', and δ of the beam in the extraction line, we see that the last BPM has the largest R_{34} and the largest R_{36} from the first BPM. This means that this BPM has the best resolution of both the incoming vertical angle and of the energy, and that the measurement of energy will be correlated to the vertical angle. In practice, this results in the energy feedback becoming "confused" when steering in the FP and extraction regions is changed, and altering the energy as well. This "confusion" can be avoided by halting the energy feedback program during such steering, or by using the BPMs in the CCSX to restore the energy after steering to its original value.

The energy jitter is measured by the extraction line BPMs to be on the order of 1.6×10^{-4} on a pulse-by-pulse basis, or approximately 7 MeV. Because the SLAC linac consists of 240 klystrons, a single klystron "cycling" results in an energy excursion 35 times the RMS quoted above. Such an excursion moves the beam by approximately 3 mm at the last BPM in the extraction line, and even larger excursions are possible due to various transient linac phenomena. Early experience indicated that such large excursions cause the last BPM to fail; since the other BPMs do not have a large dispersion, the energy measurement becomes poor and energy feedback fails to restore the nominal energy. Such a situation requires manual intervention to recover the beam and energy feedback. To alleviate this situation, an additional BPM with low resolution was placed at the midway point between the B06 bends and the last BPM. This BPM is lightly weighted in the energy fit, and only contributes significantly when the last BPM fails. Its job is to "rescue" the energy feedback, and bring the energy back close enough to nominal to allow the last BPM to read properly. Other adjustments to the feedback options (such as filtering and maximum/minimum allowed BPM values) have also enhanced the reliability of the energy feedback.

During nominal FFTB running, it was shown that the design configuration of extraction line quadrupoles resulted in unacceptable backgrounds on the detectors

for WS6A/B and the Laser-Compton BSM, due to "scraping" in the quad apertures themselves. The optics were changed to a lower-noise setting, which also has no geometric waist at the ESM. The beam size at the ESM is therefore a combination of the geometric and dispersive sizes, and the quoted "energy spread" is actually an upper bound. During normal running, the energy spread is found to have an upper bound of typically $\frac{\sigma_E}{E} \sim 4 \times 10^{-4}$. Energy spread measurements are discussed more fully in Chapter 6.

3.1.4 Collimation

Unlike a recirculating collider (in which electron and positron bunches quickly settle into an equilibrium distribution), the constant production and extraction of bunches in a single-pass system means that the beam halo is repopulated on each bunch, and this requires a vigorous collimation system to eliminate particles at large excursions from the average in position, angle, and energy.

The FFTB's primary collimation system is the SLAC linac collimators, which uses a set of movable jaws in the last 300 meters of the linac to scrape particles at large excursions [37], and which serve as the primary collimation for the SLC as well. The first set of collimators (Sectors 28-29) do the primary collimation, and the collimators cover both betatron phases in both planes. The second set of collimators (Sectors 29-30) are used to remove particles which scatter off the first set of collimators and other sources of repopulation in the collimation region itself.

The linac collimators do not eliminate large energy oscillations, nor are they adequate for regions with enormous betatron functions such as the FFTB. The FFTB itself has two movable collimators to address these issues. The first is a momentum slit (MSLT), which has a pair of horizontal jaws and is placed near the QM3A magnet upstream of the first sextupole. This collimator is adequate for eliminating particles for which the geometric and dispersive offsets from the reference orbit, when added together, are large enough to encounter the collimator jaws; particles for which this is

not the case (for example, if the geometric and dispersive terms are large but opposite, in which case they cancel) pass through. The second collimator (CXY) contains two movable horizontal jaws and two movable vertical ones, and is placed at a point with equal dispersion to the MSLT but separated by a geometric $-I$. Particles whose large geometric and dispersive offsets canceled at the MSLT will encounter the CXY horizontal jaws, while the vertical jaws provide a final collimation before the tight vertical aperture of the CCSY sextupoles.

In general, setting up the FFTB collimation for high-divergence optics is a tedious and difficult affair involving significant trial and error and small moves of the jaws, since each jaw will generally have an optimum position between allowing too much beam to pass through (causing backgrounds), and intercepting too much beam (which causes off-energy repopulation, and worse backgrounds). Beam-based alignment of the quadrupoles and sextupoles can generally ease the rigors of collimation setup, and also improves the final result.

3.2 · Emittance Measurement in the FFTB

Because the FFTB does not exercise full control over the 3 km SLAC linac and the BSY, it is not guaranteed that the beam parameters (β, α) will be equal to the FFTB design values at the entrance to the beamline. Furthermore, no reliable coupling measurement can be made at the Sector 28 wires. Therefore, the FFTB has its own wire scanners for reconstruction of the incoming beam phase space. These are the x, y, and v yokes of WS1, in the Beta Match.

Multi-wire beam ellipse measurements, such as those in Sector 28, are ideal for linear collider work because they are essentially non-invasive: the wires are scanned through the beam sequentially, requiring no alteration of the colliding- beam conditions. Unfortunately, such measurements cannot easily be carried out in the FFTB. During high-divergence conditions, the FFTB contains several images of the FP, and

the rest of the beamline is, to a good approximation, exactly $\frac{\pi}{2}$ away in betatron phase. The beam is only at a betatron phase between the FP's and the rest of the line within a few centimeters of the FP images. As a result, only one of the three parameters of phase space is easily accessible in the FFTB itself ($\frac{\pi}{2}$ from the FP), while the other two are very difficult to measure.

An earlier method for measuring the three parameters of the beam ellipse requires focusing the beam to a waist on a wire scanner, and measuring the beam size as a function of the strength of an upstream quadrupole [38]. This is the beam reconstruction technique used in the FFTB. The beam is focused onto a waist at WS1, via a special setting of the beta match quads. Because the outgoing beam is no longer matched to the FFTB, ST62 is inserted to prevent the beam entering the CCSX (see Figure 2.4). The beam size is then measured by the 7 μm wires of WS1 to extract emittances and Twiss parameters.

In addition to the original "Quad Emit" (or "Auto Emit") technique, two methods have been developed to permit measurement of the full 4×4 beam matrix via quad scans on WS1. Each of these techniques – Quad Emit, 4D Quad Emit, and Super Quad Emit – are discussed below.

3.2.1   Uncoupled Quad Emittance Measurements

Consider the situation depicted in Figure 3.1, in which a thin quadrupole magnet (originally at zero strength) is located upstream of a wire scanner; the R matrix from the quad's downstream face to the wire is R, the quad matrix is Q, and the product is $S \equiv RQ$. The incoming beam matrix $\sigma^{(0)}$, is transported to $\sigma^{(w)}$:

$$\sigma^{(w)} = S\sigma^{(0)}S^t. \tag{3.1}$$

Let us assume the scanner is an xy-only scanner, and therefore can only resolve $\sigma_{11}^{(w)}$ and $\sigma_{33}^{(w)}$, which are the squares of the RMS sizes measured by the wire scanner. Let us further assume that R is uncoupled and the thin quad is not rolled. We can
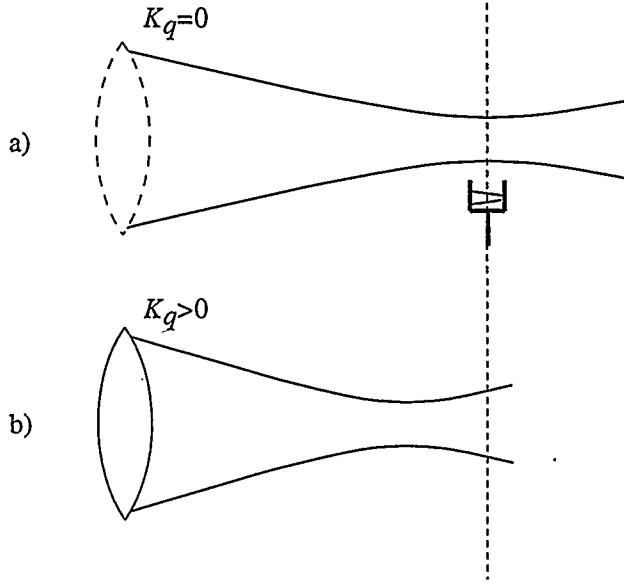
Figure 3.1: Quad Emittance measurement procedure. By changing an upstream quad strength from zero (as in (a)) to a non-zero value (as in (b)), it is possible to sample the beam size at different phase advances relative to the quadrupole entrance with a single wire scanner.

therefore express the measures values of $\sigma_{11}^{(w)}$ and $\sigma_{33}^{(w)}$:

$$\sigma_{11}^{(w)} = S_{1i}S_{1j}\sigma_{ij}^{(0)} = S_{11}^2\sigma_{11}^{(0)} + S_{12}^2\sigma_{22}^{(0)} + 2S_{11}S_{12}\sigma_{12}^{(0)},$$
$$\sigma_{33}^{(w)} = S_{3i}S_{3j}\sigma_{ij}^{(0)} = S_{33}^2\sigma_{33}^{(0)} + S_{34}^2\sigma_{44}^{(0)} + 2S_{33}S_{34}\sigma_{34}^{(0)}. \tag{3.2}$$

Q can be replaced by the thin-quad matrix[2]:

$$Q = \begin{pmatrix} 1 & 0 \\ -K_q & 1 \end{pmatrix}, \tag{3.3}$$

where K_q is the integrated quadrupole strength, $K_q = \frac{B_0 L_q}{aB\rho}$. For ease of notation, the derivation will now proceed for horizontal only, with the understanding that the

[2]The description here is done in terms of thin lenses for simplicity, and to show the lowest-order dependences. In practice, a thick-lens computation is used.

vertical is identical save for a change in sign of the quadrupole strength.

Expressing the values of the S matrix as the product of the (fixed) R values and the variable quad matrix Q, we see that as we scan the quad strength, the measured beam size varies as follows:

$$\sigma_{11}^{(w)} = (R_{11} - K_q R_{12})^2 \sigma_{11}^{(0)} + R_{12}^2 \sigma_{22}^{(0)} + 2 (R_{11} - K_q R_{12}) R_{12}\sigma_{12}^{(0)}. \tag{3.4}$$

Gathering terms, we find that Equation 3.4 can be expressed as:

$$\begin{aligned}
\sigma_{11}^{(w)} &= K_q^2 \left(R_{12}^2\sigma_{11}^{(0)}\right) \\
&+ K_q \left(-2R_{11}R_{12}\sigma_{11}^{(0)} - 2R_{12}^2\sigma_{12}^{(0)}\right) \\
&+ \left(R_{11}^2\sigma_{11}^{(0)} + R_{12}^2\sigma_{22}^{(0)} + 2R_{11}R_{12}\sigma_{12}^{(0)}\right).
\end{aligned} \tag{3.5}$$

Note that Equation 3.5 can easily be rewritten in the form of a *parabola*:

$$\begin{aligned}
\sigma_{11}^{(w)} &= A (K_q - B)^2 + C, \text{ where} \\
A &= R_{12}^2\sigma_{11}^{(0)}, \\
B &= \left(\frac{R_{11}}{R_{12}} + \frac{\sigma_{12}^{(0)}}{\sigma_{11}^{(0)}}\right), \\
C &= R_{12}^2 \left(\sigma_{22}^{(0)} + \frac{\sigma_{12}^{(0)2}}{\sigma_{11}^{(0)}}\right).
\end{aligned} \tag{3.6}$$

By identifying the sigma-matrix elements with the projected Twiss parameters and emittances (i.e., $\sigma_{11} = \epsilon_x\beta_x$, $\sigma_{22} = \epsilon_x\gamma_x$, $\sigma_{12} = -\epsilon_x\alpha_x$, and similarly for the vertical terms) we can extract the values of the projected parameters at the entrance of the scanned quadrupole directly from the fitted parameters of the parabola. This is the computation used by the online "Thin Lens" computation. The online "Thick Lens" computation explicitly computes the transport matrix for the quadrupole at each setting and uses the scanner data rather than the parabola coefficients to determine ϵ, β, α.

The FFTB beam measurement procedure uses the first normal quadrupole in the line, Q5, to measure both horizontal and vertical emittances. When QSM1 is set to

	ϵ (SLC units)	β, m	α
x, thin lens	3.18 ± 0.25	764 ± 80	-6.04 ± 0.63
x, thick lens	3.28 ± 0.10	750 ± 58	-7.25 ± 0.57
y, thin lens	0.154 ± 0.0011	1060 ± 20	-8.23 ± 0.17
x, thick lens	0.150 ± 0.0035	1070 ± 33	-9.41 ± 0.30

Table 3.1: Twiss parameters at entrance to Q5 computed by thin and thick lens models. "SLC Units" for emittance are 1×10^{-10}m $\cdot$ rad in the laboratory frame of reference. The data shown is from the September 14, 1994 beam measurement data.

zero, the matrix from the 50B1 "treaty point" to the Q5 entrance is a drift space, and therefore the back-propagation of the Twiss parameters is trivially accomplished. The back-propagated Twiss parameters become useful during the beam matching and match-verification phases of tuning, due to the internal architecture of the SLC Control Program's online model.

During early experiments with the online quad emit package, both Thin Lens and Thick Lens calculations were performed. In the sigma measurement configuration, Q5 has an integrated strength of 145 kilogauss, corresponding to a focal length of 10.7 meters. The quad's physical length is 46.1 centimeters, and thus the thin lens calculation was expected to be adequate. As Table 3.1 shows, the discrepancy in the thin and thick calculations is not trivial, particularly for α.

What is the likely effect of such miscalculations on the final spot size? In order to determine this, note first from Figure 2.10 that there is a preliminary focus in the BSY, some 120 meters upstream of Q5. This means that, when the FFTB is properly matched from the preliminary focus to the FP, the entire line has the properties of a pure demagnification. Let the betatron function at the preliminary waist be denoted by β_w, and the distance from this to the entrance of Q5 by L_w. It can be shown that a miscalculation of β_w will result in an incorrect value of β^*, and that a miscalculation of L_w will result in an incorrect value of L^*. Let us denote the error in the preliminary drift distance and beta function with ΔL_w, $\Delta \beta_w$, and in the FP drift distance and

	β_w, m	L_w, m
x, thin lens	20.4 ± 2.1	-123 ± 20
x, thick lens	14.0 ± 1.1	-101 ± 12
y, thin lens	15.4 ± 0.3	-127 ± 4
x, thick lens	11.9 ± 0.4	-112 ± 5

Table 3.2: Betatron function at preliminary BSY focus, and distance to same from Q5 entrance, in thin and thick lens calculations. The data from September 14, 1994 is used.

beta function with ΔL^*, $\Delta \beta^*$. It can be shown that:

$$\frac{\beta_w + \Delta \beta_w}{\beta_w} = \frac{\beta^*}{\beta^* + \Delta \beta^*}, \text{ and}$$
$$\frac{\Delta L_w}{\beta_w} = \frac{\Delta L^*}{\beta^*}. \tag{3.7}$$

From the Twiss parameters in Table 3.1, it is possible to compute the values of β_w and L_w, and these are shown in Table 3.2. It should be noted that for this computation, the errors on β and α have been assumed to be uncorrelated.

As we see in Table 3.2, the most significant error implicit in the thin lens computation (assuming the thick lens values to be essentially accurate) is the betatron function at the BSY focus, which is miscalculated by as much as 40% in the horizontal and nearly 30% in the vertical. The waist shifts are at the level of 1 β_w, which would produce shifts at the FP of 1 β^*. This level is acceptable.

Early FFTB experimental runs found the normalized vertical emittance in Sector 28 to be around 5×10^{-6}m $\cdot$ rad, while the same quantity at the entrance to Q5 was closer to 1×10^{-5}m $\cdot$ rad. While even the linac emittance was far larger than the FFTB design value, the amount of emittance dilution indicated between the end of the linac and the beginning of the FFTB is considerable. At that time, the damping rings and linac were using a simple modification of the SLC timing, which damps the electron emittances for only 8.3 milliseconds. While this number is logical for SLC, which injects at 120 Hz, the FFTB's 30 Hz permits 16.7 milliseconds of damping time. Once the timing was adjusted to increase the "store time" in this fashion, the linac
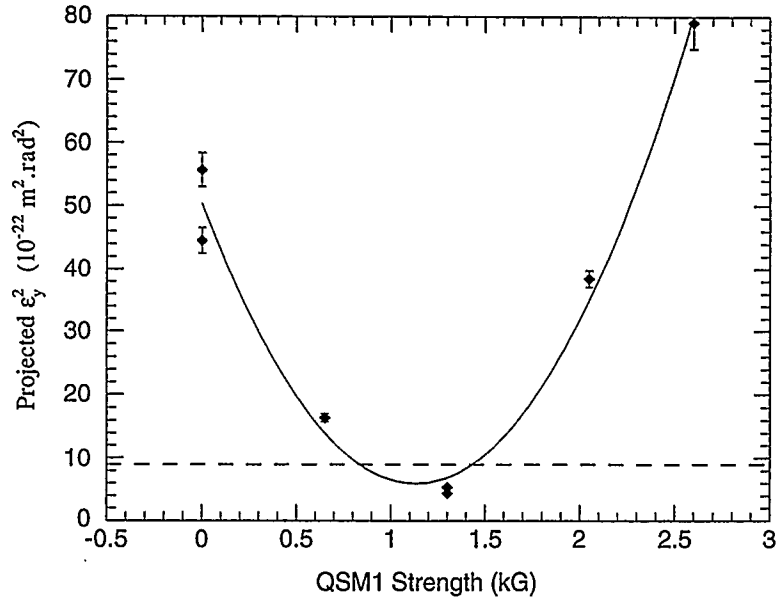
Figure 3.2: Projected vertical emittance vs QSM1 strength. The dashed line indicates the FFTB design. This measurement was performed on January 3, 1995.

emittances were reduced to values as low as 1.5×10^{-6}m · rad, with typical numbers closer to 2.0×10^{-6}m · rad. The large emittance growth was hypothesized to be a result of xy coupling. To test this theory, the quad emit scans were repeated with skew quadrupole QSM1 at different strengths.

For small values of QSM1 strength, we can estimate its effect as follows: since QSM1 gives the electrons a change in their y' values which is proportional to x, we expect that small values of QSM1 will not change the vertical divergence, but can change the correlation between y and y'. Recalling the normalized correlation coefficients, r_{ij}, we can rewrite the projected emittance in the following form:

$$\epsilon_y^2 = \sigma_y^2 \sigma_{y'}^2 \left(1 - r_{34}^2\right). \tag{3.8}$$

QSM1 = 0 kG   QSM1 = 1.3 kG
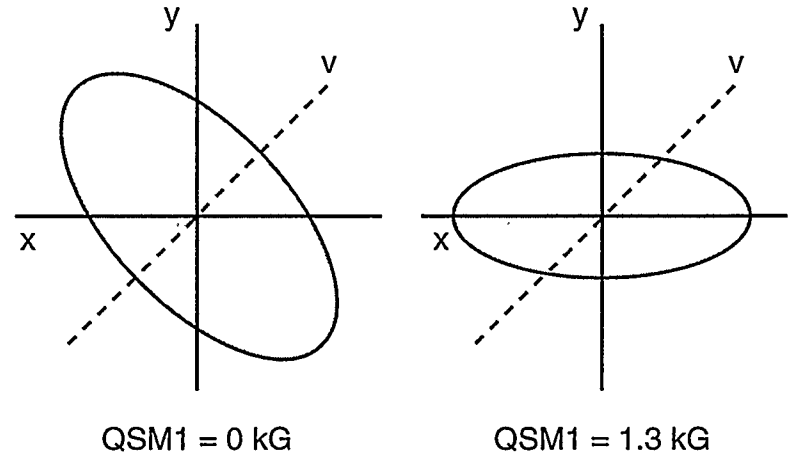
Figure 3.3: Reconstructions of the beam's xy ellipse at WS1 in beam-measurement configuration, for two different values of QSM1. With QSM1 set to zero, the beam is rolled 47° relative to the xz plane, and the RMS size along the minor axis is 14 μm. Setting QSM1 to minimize the projected emittance unrolls the beam and also reduces the minor axis RMS to 9.9 μm. Plots shown are from the January 3, 1995 data.

If r_{34} varies linearly with QSM1 strength, we can expect to see a quadratic dependence between QSM1 strength and ϵ_y^2.

Figure 3.2 shows the measured values of ϵ_y^2 as a function of QSM1, with the expected quadratic dependence. Note that the minimum value occurs at QSM1 $\sim$ 1.2kG, indicating the presence of significant coupling. The actual value of QSM1 which minimized the vertical emittance varied from 0.7 to 1.5 kilogauss between FFTB runs, and minimizing the emittance in this fashion typically resulted in vertical normalized emittances at or below the Sector 28 value at the time.

Additional evidence for xy coupling of the incoming beam can be seen in the reconstruction of the full xy ellipse at WS1, which requires measurement of the beam size by the v-wire (see Section 2.5.2). The beam measured with QSM1 set to zero typically is rolled by up to 50° relative to the xz plane, and its minor axis is also

larger than the minor axis with QSM1 set to minimize emittance. Figure 3.3 shows an example of this measurement.

Note that QSM1 gives only one degree of freedom for adjusting the incoming coupling. While the effectiveness of QSM1 in reducing the emittance is fortunate, it does not guarantee that the incoming coupling has been truly eliminated. It is therefore necessary to reconstruct the full incoming beam phase space, using a fully-coupled extension of the quad emit formalism.

3.2.2 Coupled Beam Measurements – General

We saw in the previous Section that a single normal quad scan on a wire scanner with x and y wires is just sufficient to reconstruct the in-plane terms of the sigma matrix ($\sigma_{11,12,22}$ and $\sigma_{33,34,44}$). This algorithm does not yield any insight into the cross-plane terms of the sigma matrix ($\sigma_{13,14,23,24}$). In order to determine these, we must measure the beam size at an angle to the x and y axes, and this is the purpose of wires which measure the u or v axes.

Consider the beam matrix, σ_{ij}, at the location of a wire scanner, and let the scanner's direction of measurement be rotated clockwise by an angle θ from the x axis. We can determine the beam size seen by the wire by applying the matrix for a *passive* rotation (i.e., a rotation of the coordinate system) to the beam matrix. The rotation transforms the x axis into the wire axis, and the transformed σ_{11} term is the mean- squared size reported by the wire scanner. This size shall henceforth be denoted σ_{ww}. Formally:

$$\sigma_{ww} = \left[R\sigma R^t \right]_{11} = R_{1i}R_{1j}\sigma_{ij}. \tag{3.9}$$

The R matrix for a passive rotation is given by:

$$R(\theta)_{ij} = \begin{pmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{3.10}$$

and the square of the beam size reported by the wire is therefore:

$$\sigma_{ww} = \sigma_{11}\cos^2\theta + \sigma_{33}\sin^2\theta + 2\sigma_{13}\sin\theta\cos\theta. \tag{3.11}$$

The FFTB's WS1 has a v wire in addition to the x and y wires. The v wire angle relative to the x axis is 135°, resulting in a beam size given by:

$$\sigma_v^2 = \frac{1}{2}\left(\sigma_{11} + \sigma_{33} - 2\sigma_{13}\right). \tag{3.12}$$

Note that, if all three wires (x,y,v) are scanned at each setting of the scanned magnet, then σ_{11} and σ_{33} are known from the x and y measurements, and by subtraction we may determine the value of σ_{13} from the v scan.

Let us consider the value of σ_{13} at the wire scanner under the conditions of a quad scan, as we considered the in-plane sizes in the previous Section. Using that notation,

$$\begin{aligned} \sigma_{13}^{(w)} &= S_{1i}S_{3j}\sigma_{ij}^{(0)} \\ &= S_{11}S_{33}\sigma_{13}^{(0)} + S_{11}S_{34}\sigma_{14}^{(0)} + S_{12}S_{33}\sigma_{23}^{(0)} + S_{12}S_{34}\sigma_{24}^{(0)}. \end{aligned} \tag{3.13}$$

Recalling that $S \equiv RQ$ and using the thin-quad approximation allows us to rewrite Equation 3.13:

$$\begin{aligned} \sigma_{13}^{(w)} &= \left(R_{11} - K_q R_{12}\right)\left(R_{33} + K_q R_{34}\right)\sigma_{13}^{(0)} + R_{12}\left(R_{33} + K_q R_{34}\right)\sigma_{23}^{(0)} \\ &\quad + R_{34}\left(R_{11} - K_q R_{12}\right)\sigma_{14}^{(0)} + R_{12}R_{34}\sigma_{24}. \end{aligned} \tag{3.14}$$

Gathering terms by powers of K_q yields:

$$\begin{aligned} \sigma_{13}^{(w)} &= -K_q^2\left(R_{12}R_{34}\right)\sigma_{13}^{(0)} \\ &\quad + +K_q\left[\left(R_{11}R_{34} - R_{12}R_{33}\right)\sigma_{13}^{(0)} + R_{12}R_{34}\left(\sigma_{23}^{(0)} - \sigma_{14}^{(0)}\right)\right] \\ &\quad + R_{11}R_{33}\sigma_{13}^{(0)} + R_{12}R_{33}\sigma_{23}^{(0)} + R_{11}R_{34}\sigma_{14}^{(0)} + R_{12}R_{34}\sigma_{24}^{(0)}. \end{aligned} \tag{3.15}$$

We can see immediately that, from a parabolic fit yielding coefficients A, B, C, we can derive coefficients a_2, a_1, a_0 such that $\sigma_{13}^{(w)} = \sum_{i=0}^{2} a_i K_q^i$, which can be identified with coefficients in Equation 3.15. We also see that, with three determined coefficients and four parameters, the system is underdetermined. Three parabolas yield only nine coefficients, which is insufficient to determine ten parameters of the sigma matrix. Additional information must be sought to solve the system, and we can gain insight into the information needed by using a_2 and a_1 to eliminate $\sigma_{13}^{(0)}$ and $\sigma_{23}^{(0)}$ from the equation for a_0:

$$a_0 = a_2 \left[\frac{R_{11}R_{34} - R_{12}R_{33}}{R_{12}^2 R_{34}^2} - \frac{R_{11}R_{33}}{R_{12}R_{34}} \right] + a_1 \frac{R_{33}}{R_{34}}$$
$$+ \ \sigma_{14}^{(0)} \left(R_{11}R_{34} + R_{33}R_{12} \right) + \sigma_{24}^{(0)} \left(R_{12}R_{34} \right). \tag{3.16}$$

For notational simplicity, let us define $F(R, a_i)$ such that:

$$F(R, a_i) \equiv a_0 - a_2 \left[\frac{R_{11}R_{34} - R_{12}R_{33}}{R_{12}^2 R_{34}^2} - \frac{R_{11}R_{33}}{R_{12}R_{34}} \right] - a_1 \frac{R_{33}}{R_{34}}, \tag{3.17}$$

and therefore state that:

$$F(R, a_i) = \sigma_{14}^{(0)} \left(R_{11}R_{34} + R_{33}R_{12} \right) + \sigma_{24}^{(0)} \left(R_{12}R_{34} \right). \tag{3.18}$$

Given a known transport matrix R, and measured coefficients a_i from a quad scan, $F(R, a_i)$ is a linear combination of the remaining off-diagonal terms of the beam matrix, $\sigma_{14}^{(0)}$ and $\sigma_{24}^{(0)}$.

3.2.3 4-D Quad Emit Algorithm

The most direct approach to extracting the ten sigma matrix parameters is to perform the quad scan vs σ_x, σ_y, σ_v as described in the previous Section, and then to perform an additional quad scan vs σ_v, but to do the last scan in such a manner that $F(R, a_i)$ is independent in the two v-wire scans. We see from Equation 3.18 that the ratio of $(R_{11}R_{34} + R_{33}R_{12})$ to $(R_{12}R_{34})$ must be changed from one v scan to the next to ensure this independence. This can be done by performing the first set of x,y,v scans in the

Parameter (units)	Fitted Value
$\gamma\epsilon_1$ (SLC units)	2.50
β_1 (m)	845
α_1	-7.18
$\gamma\epsilon_2$ (SLC units)	0.167
β_2 (m)	998
α_2	-7.95
B_{11}	-0.047
B_{12}	0.054
B_{21}	-0.084
B_{22}	0.078

Table 3.3: Beam parameters fitted by 4-D Quad Emit program, based on data taken on September 4, 1994 with QSM1 set to 1.3 kilogauss. The projected vertical emittance of the beam as shown is $\gamma\epsilon_y = 0.188$ in SLC units; this constitutes an emittance dilution due to coupling of 11%.

optics used for 2-D Quad Emit, changing the beta match optics, and performing a scan of Q5 vs WS1 v-wire again. This algorithm is the *4-D Quad Emit* procedure [3].

The 4-D Quad Emit procedure uses the coefficients of the fitted parabolas and a thin-lens approximation to represent Q5. Because four quad scans are used, a total of 12 parameters are used to fit 10 beam parameters, yielding an overdetermined system. The algorithm returns the beam sigma matrix, the normal-mode and projected emittances and Twiss parameters, and the coupling parameters in Spence notation (see Appendix A). The algorithm also returns normalized fit residuals to the parabola coefficients (i.e., the difference between the measured coefficients and those expected from the fitted beam matrix, divided by the measured values). The beam matrix parameters are fitted at the entrance to Q5, which is upstream of QSM1; therefore, 4-D Quad Emit can be used to measure the residual coupling after QSM1 has been set to minimize the projected emittance.

The standard 2-D quad emit optics has $R_{11} = -2.534$, $R_{12} = 25.05$ m, $R_{33} = 3.732$, $R_{34} = 43.14$ m; while the optics for the additional v scan has $R_{11} = -2.849$,

---

[3]The 4-D Quad Emit measurement was developed by W.L. Spence.

Scan	A	B	C
x	5.29×10^{-2}	-1.28×10^{-3}	-1.01×10^{-2}
y	-5.82×10^{-3}	1.52×10^{-4}	-2.09×10^{-4}
v1	3.35×10^{-2}	-8.25×10^{-4}	-2.77×10^{-3}
v2	9.33×10^{-2}	2.22×10^{-3}	-9.86×10^{-3}

Table 3.4: Fractional residuals of parabola coefficients after 4-D Quad Emit fit.

$R_{12} = 32.66$ m, $R_{33} = 1.448$, $R_{34} = 22.08$ m. For the quad emit optics this yields a ratio of $(R_{11}R_{34} + R_{33}R_{12})$ to $(R_{12}R_{34})$ of -0.01465, while the optics for the additional v scan has a ratio of -0.02164, so the necessary condition is satisfied.

Table 3.3 shows the values of the normal-mode Twiss parameters, normal-mode emittances, and the Spence parameters determined by the 4-D Emit computation after optimizing the projected emittance with QSM1. Note that the projected and normal-mode emittances are (in SLC units) 0.188 and 0.167, respectively, which indicates that the remaining dilution of the vertical emittance is only 11%. If no additional attempts are made to remove the incoming coupling, the expected FP linear beam size is 45 nm, while a fully decoupled beam would yield a linear beam size of 43 nm. While the algorithm does not return uncertainties as such, Table 3.4 shows the normalized fit residuals of the 12 coefficients which are used as input by the package. The worst fractional residual is just under 10%, which argues that the beam sigma matrix returned by 4-D Quad Emit probably has 10% precision at worst.

In Section 3.2.1, we saw that the use of a purely thin-lens computation led to noticable discrepancies in the incoming beam reconstruction, when these results were compared to a thick-lens computation. Presumably such is also the case for this algorithm, and an understanding of this issue is helpful in verifying the proper functionality of the algorithm. An additional "ground-truth check" would compare the results of the algorithm for different values of QSM1, and verify that the expected transformations in the beam are taking place. Each of these issues is addressed below.

Scan	A ($\mu m^2/kG^2$)	B (kG)	C (μm^2)	χ^2/ν
x	57.3 ± 3.4	-144.2 ± 0.08	204.8 ± 39.8	2.15
y	16.0 ± 0.75	-147.0 ± 0.06	38.27 ± 4.6	5.7
v1	28.1 ± 1.5	-144.1 ± 0.08	148.3 ± 2.8	2.1
v2	52.1 ± 4.3	-122.1 ± 0.14	353 ± 0.0	2.7

Table 3.5: Parabola coefficients for September 3, 1994 4-D Quad Emit measurement. Measurement "v2" is made using a different optics, thus the altered value of "B" relative to all other scans. Parameter "C" of scan "v2" returned an uncertainty of 0.0, and is not used in the thin-lens vs thick-lens analysis described in the text.

Thin Lens effects in 4-D Quad Emit

The most direct test of the 4-D Quad Emit algorithm's vulnerability to thin lens approximation errors is to propagate the fitted incoming beam from the entrance of Q5 to WS1, in the appropriate optics, and observe whether the beam predicted at WS1 conforms to that observed in the measurement. In this case, the fitted parabolas and the errors on the parabola coefficients are the available data, and it is therefore convenient to compare the parabola coefficients and not the beam sizes themselves. Table 3.5 contains the fitted parabola coefficients and their associated errors, along with the χ^2/ν for each parabola. Note that the returned error for the minimum spot in the second v-scan is identically zero; this term is therefore not used in the analysis below.

The propagation of the incoming beam to WS1 was accomplished using the SLAC *Final Focus Flight Simulator (FFFS)*, which uses the standard beam-optics program TRANSPORT as its mathematical kernel. The FFFS uses the SLC control system's online fitting algorithms to fit parabolas to beam scans, and these are therefore performed identically to the fits for real 4-D Quad Emit data. Table 3.6 shows the fitted coefficients obtained by tracking the input beam in Table 3.3 to WS1. Using the experimental errors from Table 3.5, it is also possible to calculate a χ^2 contribution for each coefficient, and these are also in Table 3.6. Many of the χ^2 contributions in Ta-

Scan	A ($\mu m^2/kG^2$)	B (kG)	C (μm^2)
x	54.1 (0.9)	-146.7 (952)	176.7 (0.5)
y	14.8 (3.0)	-145.4 (688)	36.2 (0.2)
v1	26.1 (1.8)	-146.2 (725)	108.6 (199)
v2	58.4 (2.6)	-123.3 (74)	340

Table 3.6: Parabola coefficients from using the 4-D Quad Emit beam parameters in a thick-lens quad scan simulation. Numbers shown in parentheses are χ^2 contributions (square of deviation between simulation value and measured value, divided by square of uncertainty in measured value).

ble 3.6 are enormous: values of almost 1000 are seen, indicating that the parameters obtained via simulation and experiment diverge by up to 30 times the experimental uncertainty.

Table 3.6 shows that the "A" and "C" parameters generally have acceptable χ^2 contributions, but that the measured and predicted values for "B", the Q5 strength which minimizes the beam size on the wire scanner, disagree unacceptably. It is now necessary to verify that the noted divergences arise from the thin lens approximation alone.

One procedure which can correct the lowest-order distortions of the thin lens approximation is shown diagramatically in Figure 3.4. The top half of the figure shows a beam line in which a thin lens is the first element; the lower half is the same, except that the thin lens has been replaced by a thick lens. The transport matrix from the downstream face of the quad to the target wire scanner, R, is the same in both cases. If the beam matrix at the wire scanner in the first system is identical to the matrix at the wire scanner in the second system, then the beam matrix at the exit of the thin lens and that at the exit of the thick lens must be identical to one another also. Therefore, a beam reconstructed with a thin-lens formalism which is propagated forward through the thin lens, and then backwards through the thick lens, should yield a beam matrix which has been corrected for the thin-lens approximation. This method suffers from one crucial defect in this case: the "v2" scan was performed
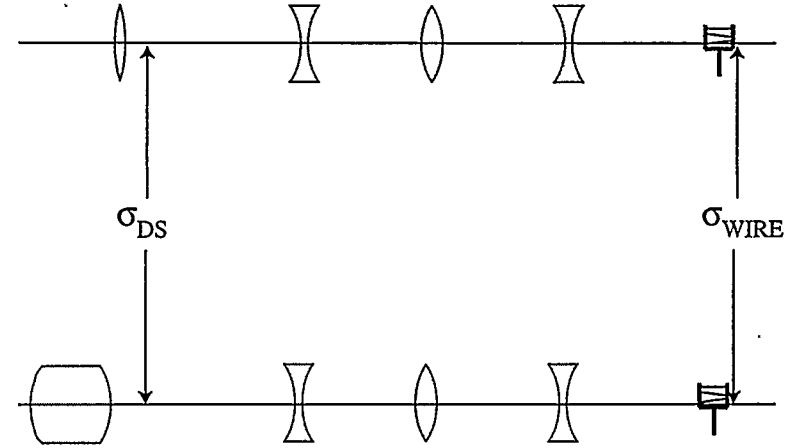
Figure 3.4: Thin versus thick lens correction for quad emit results. The two systems are identical from the downstream face of the scanned quad to the wire scanner. Identical beam matrices at the scanner (σ_{WIRE}) will therefore yield identical beam matrices at the exit of the scanned quad (σ_{DS}).

with Q5 at a different nominal strength. We therefore anticipate that application of this correction will yield less improvement in the coefficient for the "v2" scan than for the other three.

Table 3.7 shows the beam parameters measured in the thin-lens case, and those corrected by the method described above to a thick-lens case. What is most immediately obvious about this correction is that only the values of α_1 and α_2 appear to change dramatically; the emittances, betatron functions, and Spence parameters are essentially unchanged by this transformation.

Table 3.8 shows the parabola coefficients obtained by using the corrected beam in Table 3.7 to simulate quad scans from Q5 on WS1. Note that the convergence of "B" parameters is, in general, improved, but that the "A" and "C" parameters of scan "v2" are further from the experimental value, as expected. This supports the theory that deviations between 4-D Quad Emit results and simulation arise primarily from applying the thin-lens approximation to Q5; furthermore, Table 3.7 shows that, while

Parameter (units)	Thin Lens Value	Thick Lens Value
$\gamma\epsilon_1$ (SLC units)	2.50	2.51
β_1 (m)	845	799
α_1	-7.18	-8.23
$\gamma\epsilon_2$ (SLC units)	0.167	0.166
β_2 (m)	998	1041
α_2	-7.95	-9.34
B_{11}	-0.047	-0.049
B_{12}	0.054	0.053
B_{21}	-0.084	-0.084
B_{22}	0.078	0.078

Table 3.7: September 4, 1994 4-D Quad Emit results, and the parameters after correction for thin-lens effects as described in the text.

the uncoupled matching is affected by the approximation, the coupling parameters are still extracted correctly by this algorithm.

Scan	A (μm^2/kG^2)	B (kG)	C (μm^2)
x	51.4 (3.1)	-144.0 (9.5)	186.8 (0.5)
y	15.3 (1.0)	-147.0 (0.0)	34.5 (0.7)
v1	25.1 (4.2)	-144.0 (2.8)	139.9 (8.9)
v2	40.2 (7.9)	-121.6 (13.3)	278

Table 3.8: Parabola coefficients from simulation of a quad scan with the corrected beam parameters.

Non-Optimal QSM1 Values

The September 3, 1994 beam reconstruction included a measurement of the full beam matrix with QSM1 set to zero. Because the effects of QSM1 are, in principle, easily simulated, it is possible to use the beam measured with QSM1 optimized (BDES = 1.3 kG), and determine what the beam matrix would be with QSM1 zeroed. Thus, the two measurements can be compared directly. Table 3.9 shows the 4-D Quad Emit measurement at QSM1=0. Also shown are two projections, based upon the measurements made at QSM1=1.3 kG, of what the expected beam parameters would

Parameter (units)	QSM1=0 Fit	Thin QSM1=1.3 kG	Thick QSM1=1.3 kG
$\gamma\epsilon_1$ (SLC units)	2.83	2.50	2.51
β_1 (m)	764	812	768
α_1	-6.50	-6.84	-7.86
$\gamma\epsilon_2$ (SLC units)	5.1×10^{-4}	0.167	0.166
β_2 (m)	2.2×10^5	959	1001
α_2	-1.6×10^3	-7.61	-8.95
B_{11}	0.748	-0.047	-0.049
B_{12}	0.0030	0.055	0.054
B_{21}	9.93	0.663	0.657
B_{22}	0.0068	0.080	0.079

Table 3.9: Beam parameters fitted by 4-D Quad Emit algorithm on September 3, 1994, with QSM1=0. For comparison, the beam parameters measured at QSM1=1.3 kG are back-propagated to 50B1, and forward-propagated to Q5 with QSM1=0, which should result in the same beam matrix.

be at QSM1=0: both the beam parameters from the thin-lens fit, and the parameters corrected for the lowest-order thick lens effects, are included.

The most striking feature of the QSM1=0 fit is that the normal mode emittance and Twiss functions ϵ_2, β_2, α_2 are drastically different from the parameters expected by adjusting the skew-optimized beam parameters. This is because the vertical projected emittance is dominated by coupling from the horizontal plane. The vertical projected emittance given by the parameters in the second and third columns of Table 3.9 can be computed using the emittance relations in Equation A.78, and the relations in Section A.3.5: the value thus determined is 0.48×10^{-10}m $\cdot$ rad, while the normal-mode emittance is 0.17×10^{-10}m $\cdot$ rad. In Section 3.2.1, we saw that the projected emittance is the sum in quadrature of the appropriate normal-mode emittance and the coupling contribution. The coupling contribution in this case is therefore 0.45×10^{-10}m $\cdot$ rad, which is nearly as large as the projected emittance. In such a case, the uncertainty on the fitted normal-mode emittance becomes large and the results of 4-D Quad Emit become unreliable [39].

3.2.4 Super Quad Emit

A significantly different approach to measuring the incoming beam is the online application *Super Quad Emit*[4]. This application is similar to the thick-lens 2-D Quad Emit algorithm, and is based upon techniques developed by P. Raimondi for the SLC Final Focus [40]. As in the other techniques described previously in this Chapter, the beam is focused onto a dual waist at WS1 (via the 2-D Quad Emit optics); x, y, and v beam sizes are measured as a function of quadrupole strengths. In this procedure, however, the optics are not changed but instead several quadrupoles are scanned, including (potentially) skew quadrupole QSM1.

Once the quadrupole scans are complete and all beam sizes have been measured, the data (quad strengths, beam sizes, errors on beam sizes) are passed to a fitting engine (OPTFIT), along with the fixed optics of the FFTB line, arrangement of the scanned quadrupoles and wire scanners, roll angles of all elements – in short, all the physical description of the beamline. The fitting engine loads an initial "guess" of the normal mode emittances, Twiss, and Irwin parameters, computes a sigma matrix at the beginning of the FFTB line (specifically, the 50B1 treaty point), and propagates this sigma matrix to WS1 for each value of the scanned quadrupoles. By comparing the beam sizes at WS1 to the measured values for all steps of all magnets, a χ^2 value is computed. The routine for converting beam parameters to a sigma matrix, tracking the sigma matrix to the wire scanner, and constructing the χ^2 is passed to the CERNLIB minimization engine MINUIT. MINUIT then minimizes the χ^2 by computing the second derivatives of χ^2 with respect to the parameters, inverting the matrix of second derivatives to obtain a covariance matrix, and using this information to estimate the location of the minimum [41]. Note that the matrix of second derivatives is computed numerically, via finite- difference techniques, and that MINUIT therefore only requires a function which will supply the value of χ^2 when passed a set of values for the parameters to be fitted.

---

[4]Super Quad Emit was developed by P. Raimondi, L. Yasukawa, and the author

Super Quad Emit's main advantages for reconstruction of the incoming beam are: it uses a thick-lens model and the wire scanner data and errors; and it runs on the SLC control system, requiring no off-line computations or changes of optics. However, Super Quad Emit can only determine the beam parameters at the treaty point which begins a section of the SLAC accelerator complex (in this case, 50B1). Therefore, there is no way to use Super Quad Emit to measure the beam parameters downstream of QSM1, and no way to use it to measure the residual coupling after QSM1's correction.

Theoretical Considerations

Let us consider once again a thin-lens approximation of Super Quad Emit, in order to understand the lowest-order behavior of the algorithm. Let us assume a matrix from the treaty point to the scanned quadrupole, $R^{(1)}$, and a matrix from the scanned quadrupole to the wire scanner, $R^{(2)}$, and let us require that $R^{(1)}$ and $R^{(2)}$ be uncoupled. The matrix from the treaty point to the wire scanner is given by S, defined here as follows:

$$S \equiv R^{(2)} Q R^{(1)}, \tag{3.19}$$

where Q is the thin-lens quad matrix as defined in Section 3.2.1. Finally, let us define a matrix, $\mathcal{R}$:

$$\mathcal{R} \equiv R^{(2)} R^{(1)}. \tag{3.20}$$

As defined above, $\mathcal{R}$ represents the matrix from the treaty point to the wire scanner when the scanned quad is at zero[5]. With these definitions, a certain amount of algebra yields a solution for $\sigma_{11}^{(w)}$ as a function of $\sigma_{ij}^{(0)}$ (here $\sigma^{(0)}$ represents the beam matrix at the treaty point and not at the scanned quad):

$$\sigma_{11}^{(w)} = \left(\mathcal{R}_{11}^2 - 2K_q \mathcal{R}_{11} R_{11}^{(1)} R_{12}^{(2)} + K_q^2 (R_{11}^{(1)} R_{12}^{(2)})^2 \right) \sigma_{11}^{(0)}$$

---

[5]Although Q5's nominal value for the measurement optics is nonzero, we can think of "splitting" the nominal Q5 into two halves, one of which is included in the representation of $R^{(1)}$, the other in $R^{(2)}$, and consider Q a "deviation matrix" which represents the difference in Q5 from its nominal value.

$$+ \quad 2\left[\mathcal{R}_{11}\mathcal{R}_{12} - K_q R_{12}^{(2)}\left(\mathcal{R}_{12}R_{11}^{(1)} + \mathcal{R}_{11}R_{12}^{(1)}\right) + K_q^2 R_{11}^{(1)}R_{12}^{(1)}(R_{12}^{(2)})^2\right]\sigma_{12}^{(0)}$$

$$+ \quad \left(\mathcal{R}_{12}^2 - 2K_q\mathcal{R}_{12}R_{12}^{(1)}R_{12}^{(2)} + K_q^2(R_{12}^{(1)}R_{12}^{(2)})^2\right)\sigma_{22}^{(0)}. \tag{3.21}$$

Although the functional dependencies of $\sigma_{11}^{(w)}$ are more complex, the scan of Q5 versus WS1 x still provides all information needed to determine the horizontal in-plane beam matrix terms, $\sigma_{11}^{(0)}$, $\sigma_{12}^{(0)}$, and $\sigma_{22}^{(0)}$. Similarly the scan of Q5 versus WS1 y provides the corresponding vertical terms.

The dependencies of $\sigma_{13}^{(w)}$ on K_q are quite cumbersome and complex, and not represented here. However, the principle remains the same as that shown in Section 3.2.2: the three parabola coefficients are coupled to four cross-plane correlation terms, allowing two of the terms to be eliminated. This provides a single equation in two unknowns:

$$F(\mathcal{R}, R^{(1)}, R^{(2)}, a_i) \equiv f_1(\mathcal{R}, R^{(1)}, R^{(2)})\sigma_{23}^{(0)} + f_2(\mathcal{R}, R^{(1)}, R^{(2)})\sigma_{24}^{(0)}, \tag{3.22}$$

and a selection of quadrupoles must be made for which the ratio of f_1 and f_2 is different for the various quads in order to determine all ten sigma matrix terms.

For Q5, the ratio $\frac{f_1}{f_2} = 0.0162$, and for Q6 the ratio is 0.0425. This should be adequate for reconstructing all 10 parameters unambiguously, given the experience with 4-D Quad Emit in the previous Section. For additional resolution, however, Super Quad Emit gives the possibility of using a skew quadrupole scan, for which the matrix Q becomes a skew matrix, Q_s:

$$Q \rightarrow Q_s = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -K_q & 0 \\ 0 & 0 & 1 & 0 \\ -K_q & 0 & 0 & 1 \end{pmatrix}, \tag{3.23}$$

where we have used the FFTB and SLC damping ring convention that a skew quad of positive strength is a normal quad of positive strength which has been rotated clockwise by 45°, as seen from upstream[6]. When we explore the effects of propagating

---

[6]The SLC FF, by contrast, uses the convention that the rotation is counter-clockwise, and different modelling programs use different conventions as well.

a beam sigma matrix through the total matrix S, as before, we find:

$$\begin{aligned}
\sigma_{11}^{(w)} &= \mathcal{R}_{11}^2\sigma_{11}^{(0)} + 2\mathcal{R}_{11}\mathcal{R}_{12}\sigma_{12}^{(0)} + \mathcal{R}_{12}^2\sigma_{22}^{(0)} \\
&+ K_q^2\left[\left(R_{33}^{(1)}R_{12}^{(2)}\right)^2\sigma_{33}^{(0)} + \left(R_{34}^{(1)}R_{12}^{(2)}\right)^2\sigma_{44}^{(0)} + 2R_{33}^{(1)}R_{34}^{(1)}(R_{12}^{(2)})^2\sigma_{34}^{(0)}\right] \\
&- 2K_q(\mathcal{R}_{11}R_{33}^{(1)}R_{12}^{(2)}\sigma_{13}^{(0)} + \mathcal{R}_{11}R_{34}^{(1)}R_{12}^{(2)}\sigma_{14}^{(0)} + \\
&\qquad \mathcal{R}_{12}R_{33}^{(1)}R_{12}^{(2)}\sigma_{23}^{(0)} + \mathcal{R}_{12}R_{34}^{(1)}R_{12}^{(2)}\sigma_{24}^{(0)}).
\end{aligned} \tag{3.24}$$

Note that the coefficient of K_q^0 is simply the in-plane, un-scanned beam size; the coefficient of K_q^2 is dependent only on the out-of plane beam matrix terms; and that the coefficient of K_q contains all the cross-plane beam matrix terms. The QSM1 versus WS1 x scan will generally be dominated by the K_q^0 term, because the horizontal emittance is the larger of the two emittances. If we go to the WS1 y scan, however, we find that the coefficient of the linear K_q term is:

$$\begin{aligned}
a_{1,y} &= -2(\mathcal{R}_{33}R_{11}^{(1)}R_{34}^{(2)}\sigma_{13}^{(0)} + \mathcal{R}_{34}R_{11}^{(1)}R_{34}^{(2)}\sigma_{14}^{(0)} + \\
&\qquad \mathcal{R}_{33}R_{12}^{(1)}R_{34}^{(2)}\sigma_{23}^{(0)} + \mathcal{R}_{34}R_{12}^{(1)}R_{34}^{(2)}\sigma_{24}^{(0)}).
\end{aligned} \tag{3.25}$$

The power of this term to resolve $\sigma_{23}^{(0)}$ from $\sigma_{24}^{(0)}$ is given by the ratio of $\mathcal{R}_{33}$ to $\mathcal{R}_{34}$, which for the emittance measurement optics is -0.661. Therefore, including the QSM1 versus WS1 y scan is expected to significantly improve resolution of the coupling terms of the sigma matrix.

Experimental Usage

The Super Quad Emit algorithm was tested during the January 3-5, 1995 FFTB experiment; 4-D Quad Emit and 2-D Quad Emit were also used during this experiment, allowing for "side-by-side" comparisons of the results.

Table 3.10 shows the fit results for Super Quad Emit when all measurements are used. Three quadrupoles (Q5, Q6, QSM1) are scanned; each scan yields five measurements on each of three WS1 axes, for a total of 45 measurements. The quoted errors are normalized by χ^2/ν of the fit; the diagonal elements of the covariance matrix

Parameter (units)	Fitted Value
ϵ_1 (SLC units)	4.15 ± 0.19
β_1 (m)	69.7 ± 3.2
α_1	2.40 ± 0.13
ϵ_2 (SLC units)	0.15 ± 0.11
β_2 (m)	35.1 ± 27.5
α_2	0.53 ± 0.32
a	-0.098 ± 0.22
b	0.41 ± 0.16
c	-0.25 ± 0.31
d	-0.23 ± 0.11
χ^2/ν	6.44

Table 3.10: Results of Super Quad Emittance using all data from all quad scans. Values are quoted at 50B1 treaty point.

are multiplied by χ^2/ν, and the square root of these quantities are the appropriate error terms.

While the normal-mode parameters associated with the larger emittance (ϵ_1, β_1, α_1) are well determined, Table 3.10 shows that the small-emittance normal-mode parameters and the cross-plane coupling are poorly determined, and the fit quality indicated by the large χ^2/ν does not inspire confidence.

One significant clue to the difficulty can be discerned at once from Figure 3.5, in which the measured values of σ_x^2 at WS1 are plotted versus Q5 integrated strength for the Q5 scan. Note that the error-bars on the outlying points are large, and the points are systematically low relative to the fitted parabola. At these points, the beam is so much larger than the wire used in the wire scan (7 μm diameter) that the wire never intercepts a significant fraction of the beam, resulting in low signal levels and general degradation of the fit. Super Quad Emit has no built-in mechanisms for recognizing bad data, and in any event has too little redundancy to safely eliminate many points. "Bad" data points must be recognized and eliminated by the user, by hand. Once these points have been eliminated from the January data set (a total of 6 points, leaving 39 for the fit), the fit results are as shown in Table 3.11. Note that,
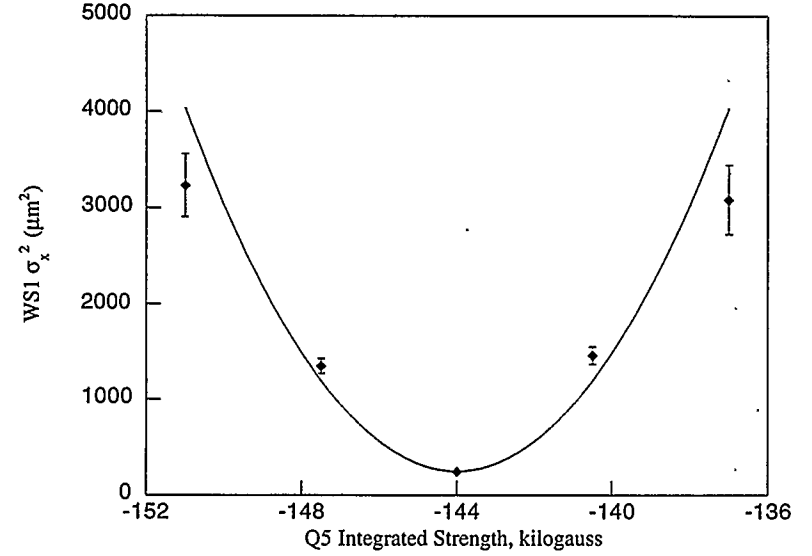
Figure 3.5: Scan of σ_x^2 as a function of Q5 from Super Quad Emit measurement.

while χ^2/ν is significantly improved, it is still larger than expected. The parabolic fits of the beam σ^2 versus quadrupole strength, once bad data are eliminated, typically yield χ^2/ν between 0.9 and 3.5, and we expect Super Quad Emit to converge at least as well.

In the theoretical discussion of Super Quad Emit, it was argued that the bare minimum amount of data needed to resolve all 10 parameters of the incoming beam was the Q5 scan and the Q6 scan, while the QSM1 scan could improve resolution of the coupling. Table 3.12 shows the fit results for Q5 and Q6 (x,y,v) data, for Q5 and Q6 (x,y,v) plus QSM1 (y), and for Q5 and QSM1 (x,y,v) data. We see right away that, while formally converging upon a solution, the Q5+Q6 data is quite poorly determined, while inclusion of QSM1 y data results in a fit only marginally better than a fit with all data including clearly bad points.

Why should the fitted values of the normal mode parameters in the small- emit-

Parameter (units)	Fitted Value
ϵ_1 (SLC units)	4.16 ± 0.17
β_1 (m)	69.3 ± 2.8
α_1	2.42 ± 0.11
ϵ_2 (SLC units)	0.189 ± 0.047
β_2 (m)	24.5 ± 5.9
α_2	0.83 ± 0.44
a	-0.298 ± 0.23
b	0.490 ± 0.057
c	-0.055 ± 0.21
d	-0.288 ± 0.52
χ^2/ν	4.50

Table 3.11: Results of Super Quad Emittance after removing fliers from fit data. Values are quoted at 50B1 treaty point.

Parameter (units)	Q5+Q6	Q5+Q6+QSM1 y	Q5+QSM1
ϵ_1 (SLC units)	4.27 ± 0.16	4.31 ± 0.27	4.29 ± 0.17
β_1 (m)	67.2 ± 4.1	73.6 ± 4.2	66.3 ± 2.8
α_1	2.29 ± 0.15	2.50 ± 0.15	2.35 ± 0.11
ϵ_2 (SLC units)	0.044 ± 0.036	0.120 ± 0.103	0.176 ± 0.043
β_2 (m)	302 ± 229	35.1 ± 31.1	25.4 ± 6.7
α_2	9.94 ± 7.47	1.23 ± 0.86	0.77 ± 0.36
a	-1.13 ± 0.46	-0.41 ± 0.37	-0.26 ± 0.21
b	0.126 ± 0.050	0.39 ± 0.17	0.485 ± 0.063
c	-0.59 ± 0.27	-0.096 ± 0.502	-0.108 ± 0.21
d	-0.057 ± 0.027	-0.27 ± 0.14	-0.283 ± 0.053
χ^2/ν	5.19	5.80	3.18

Table 3.12: Fitted beam parameters from Super Quad Emittance using different subsets of the dataset.

tance mode go haywire when QSM1 is not included? While OPTFIT uses a MINUIT call which inserts values of the Twiss and Irwin parameters into the beamline, what is ultimately tracked to WS1 is a beam matrix: the parameters from MINUIT are converted to a sigma matrix, which is then propagated to the wire scanner for each value of the quadrupoles. In the Q5 and Q6 case, the sigma matrix is not determined sufficiently well to differentiate between the projected vertical emittance which arises from the normal mode emittance, and the projected vertical emittance which is coupled over from the horizontal. This is similar to the problem of using 4-D Quad Emit to reconstruct the incoming beam without first eliminating most of the coupling, as described in Section 3.2.3.

Parameter (units)	Q5+QSM1+Q6 x	Q5+QSM1+Q6 y	Q5+QSM1+Q6 v
ϵ_1 (SLC units)	4.26 ± 0.16	4.29 ± 0.15	4.09 ± 0.22
β_1 (m)	67.2 ± 2.6	66.0 ± 2.6	70.6 ± 4.2
α_1	2.35 ± 0.10	2.34 ± 0.10	2.46 ± 0.16
ϵ_2 (SLC units)	0.181 ± 0.043	0.173 ± 0.047	0.187 ± 0.055
β_2 (m)	25.5 ± 6.4	25.9 ± 7.6	24.9 ± 7.9
α_2	0.82 ± 0.37	0.69 ± 0.28	0.79 ± 0.51
a	-0.28 ± 0.21	-0.22 ± 0.18	-0.27 ± 0.29
b	0.482 ± 0.059	0.477 ± 0.068	0.486 ± 0.075
c	-0.10 ± 0.21	-0.14 ± 0.19	-0.092 ± 0.29
d	-0.277 ± 0.051	-0.277 ± 0.054	-0.287 ± 0.069
χ^2/ν	3.66	2.81	6.64

Table 3.13: Super Quad Emit fit results using Q5 and QSM1 data, plus different subsets of the Q6 data.

Table 3.13 shows the results of Super Quad Emit when Q5 and QSM1 are used, plus a subset of Q6 data (x, y, or v scan). We see immediately that Q6 v scan dramatically reduces the quality of the fit, while Q6 y scan marginally improves it; however, the normalized error values remain the same, and the fitted values change by amounts small relative to the fitted errors. This leads to two conclusions: first, that the poor convergence of the fits in Tables 3.11 and 3.12 can be attributed to the use of Q6 v scan data; second, that the Q6 scan does not materially improve the fit

Parameter (units)	4-D Quad Emit	Super Quad Emit	Deviation (σ's)
ϵ_1 (SLC units)	3.65	4.29 ± 0.15	4.3
β_1 (m)	77.9	66.0 ± 2.6	-4.6
α_1	2.43	2.34 ± 0.10	-0.92
ϵ_2 (SLC units)	0.154	0.173 ± 0.047	-0.40
β_2 (m)	38.7	25.9 ± 7.6	-1.7
α_2	2.08	0.69 ± 0.28	-4.9
a	-0.655	−0.22 ± 0.18	2.5
b	0.450	0.477 ± 0.068	0.40
c	0.199	−0.14 ± 0.19	-1.8
d	-.260	−0.277 ± 0.054	-0.31

Table 3.14: Side-by-side comparison of Super Quad Emit and 4-D Quad Emit results obtained on same day. Deviation column represents difference between the two results, divided by error on Super Quad Emit result.

over the use of Q5 and QSM1 only.

Finally, note from Tables 3.12 and 3.13 that parameters "a" and "c" are not well-determined relative to parameters "b" and "d". This indicates that the scans of Q5 and QSM1 are not optimized to determine these parameters. Since the Q6 scan does not provide useful information, it is possible that another quadrupole, or perhaps some combination of quadrupoles in the form of a multi-knob, can be used to assist in recovering these parameters.

In order to directly compare the results of Super Quad Emit to 4-D Quad Emit, it is necessary to adjust the 4-D Quad Emit results to compensate for thin-lens effects, and back-propagate them to the 50B1 treaty point from which OPTFIT bases all its computations; in addition, the Spence parameters must be converted into the appropriate Irwin parameters (see Appendix A). The results of this conversion are shown in Table 3.14, along with the Q5+QSM1+Q6 y Super Quad Emit results. As the last column shows, the agreement between the two methods is poor, with 6 parameters disagreeing by more than 1 standard deviation. There are several possible causes to this discrepancy: the errors could be correlated in such a way that the two sets of parameters, though differing significantly when considered one at a time, actually

Parabola Coefficient	Data Value	4-D Quad Emit Beam Value	Super Quad Emit Beam Value
x scan A	80.7 ± 8.3	69.2 (1.4)	88.2 (-0.90)
x scan B	−144.0 ± 0.16	-143.7 (-1.9)	-144.1 (0.62)
x scan C	318 ± 48	290 (0.58)	323 (-0.10)
y scan A	26.6 ± 0.93	25.6 (1.1)	21.5 (5.48)
y scan B	−147.5 ± 0.049	-147.4 (-2)	-146.3 (-24)
y scan C	29.95 ± 0.0	25.6	65.3
v1 scan A	75.0 ± 3.6	67.1 (2.2)	71.8 (0.89)
v1 scan B	−145.4 ± 0.055	-145.1 (-5.5)	-144.9 (-9.1)
v1 scan C	307.2 ± 0.0	277.8	303.0
v2 scan A	87.45 ± 7.7	73.4 (1.8)	89.2 (-0.23)
v2 scan B	−121.0 ± 0.10	-120.9 (-1.0)	-121.2 (2.0)
v2 scan C	231.1 ± 0.0	195.4	116.6

Table 3.15: Parabola coefficients used in 4-D Quad Emit measurement. The first column represents the actual data, with errors scaled by $(\chi^2/\nu)^{1/2}$; the second column represents the parabola coefficients from simulating the original quad scans with the incoming beam from 4-D Quad Emit fit; the third column is simulation of the original scans with the incoming beam from Super Quad Emit. Numbers in parentheses are residuals divided by the errors from the parabola fits.

describe nearly- identical beam matrices when considered as a group; one approach could be demonstrably incorrect in its computation; each of the two fits could adequately describe the dataset from which it is derived, and the two datasets could be divergent; or the methodological differences could give rise to different systematic errors.

In the first case, we would expect that either solution for the beam parameters could equally well describe either data set; in the second case, we would expect that one of the two solutions could describe either data set well, and the other solution describes both data sets poorly. Neither of these is the case. Using the 4-D Quad Emit solution, the χ^2/ν for the Super Quad Emit dataset is 23, as opposed to 2.8 for the Super Quad Emit solution; and, as shown in Table 3.15, the Super Quad Emit solution does a poor job of reproducing the parabolas from the 4-D Quad Emit solution, particularly the "y" scan.

It is possible that the incoming beam changed between the 4-D Quad Emit and Super Quad Emit measurements; however, this possibility is rendered unlikely by the exact sequence of measurements on January 3-4, 1995. First, the incoming beam was reconstructed by the 2-D Quad Emit package, with QSM1 = 0. Then the QSM1 setting was optimized by scanning QSM1 vs WS1 vertical, yielding an optimum strength of 1.313 kG, which was applied (note that this procedure does not yield the precise value of QSM1 which minimizes the projected emittance, as shown in Figure 3.2). At this point the x and y projected beam parameters were measured by 2-D Quad Emit, and the two scans of Q5 versus WS1 v were performed; the parabolas from 2-D Quad Emit and the v scans were used by 4-D Quad Emit. QSM1 was then returned to zero, and the Super Quad Emit data was taken. Finally, the projected emittance was measured, via 2-D Quad Emit, as a function of QSM1 strength, with the final measurement at the optimized value of 1.3 kG. As a result of this sequence of events, the beam projected emittances were measured at the optimal value of QSM1 before and after the Super Quad Emit experiment, and were found to be nearly equal (2.1×10^{-11}m · rad versus 2.3×10^{-11}m · rad); the projected emittance of the Super Quad Emit beam with QSM1=1.3 kG is 3.1×10^{-11}m·rad. It seems unlikely that the incoming beam would change its parameters so dramatically, and then change them back, during the two hours of emittance measurements.

The remaining likely culprit is differing systematic errors in the two techinques. The most significant source of such differences is the different usages of QSM1. Prior to the 4-D Quad Emit measurement, QSM1 is empirically set to a value which minimizes the vertical beam size at WS1; during 4-D Quad Emit, only Q5 is scanned. Consequently, the 4-D Quad Emit procedure does not depend upon knowing the absolute value of QSM1 at any point – it only requires that the value be optimized and constant. Super Quad Emit, on the other hand, requires that the value of QSM1 be known accurately at several different values. Furthermore, Q5 and Q6 have integrated strengths of -145.7 and 174.7 kilogauss, respectively, in the beam measurement optics,
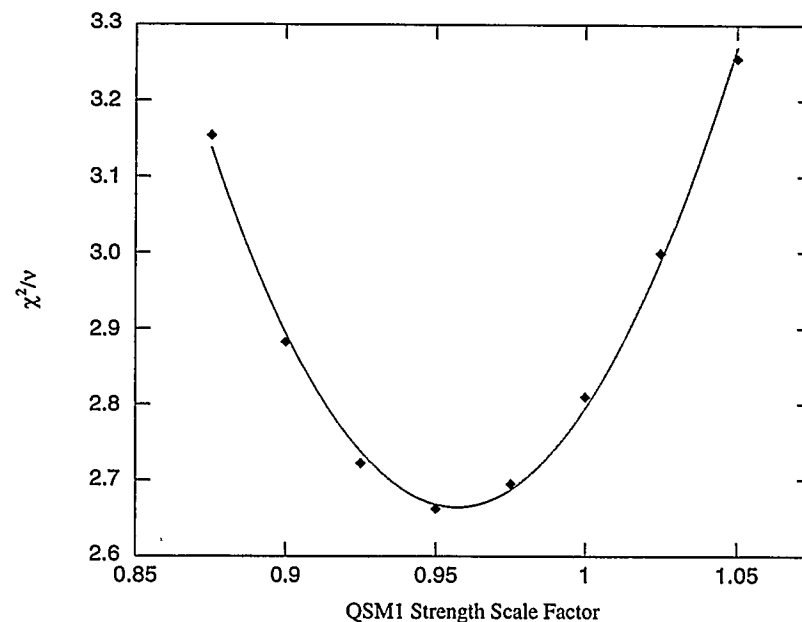
Figure 3.6: Super Quad Emit χ^2/ν as a function of QSM1 strength scale factor. The minimum of 2.67 occurs at 0.957; the "width" of the minimum, as defined in the text, is ±0.043.

corresponding to pole-tip fields of -3.64 and 4.36 kilogauss; while QSM1 is scanned from 0 kG to 3 kG, which corresponds to a pole-tip field of approximately 170 gauss. This is an extremely weak field, given that QSM1 was fabricated by cutting an FFTB Standard Quadrupole in half, and therefore QSM1 contains approximately half as much iron as the standard quads.

One plausible source of systematic error, therefore, is a scale-factor on the strength of QSM1. Figure 3.6 shows the χ^2/ν of the Super Quad Emit fit as a function of the QSM1 scale factor. Note that χ^2/ν is minimized at a QSM1 scale factor of 0.957. The width of the scale factor measurement (defined by the change in scale factor required to make χ^2 increase by the amount of χ^2/ν) is 0.043. Thus, while a scale-factor of

1.0 is within the width of this measurement, a scale-factor of 0.914 is also possible.

While a scale factor in QSM1 is believable, can such an error explain the divergence between 4-D Quad Emit and Super Quad Emit? This question can be explored by defining a "convergence coefficient" for the two datasets: for a set of n beam parameters, we can define:

$$C_{converge} \equiv \sum_{i=1}^{n} \frac{(P_{i,\text{4D}} - P_{i,\text{SQE}})^2}{\sigma_{i,\text{SQE}}^2}, \qquad (3.26)$$

where $P_{i,\text{4D}}$ and $P_{i,\text{SQE}}$ is the ith parameter in each of 4-D Quad Emit and Super Quad Emit, respectively; and $\sigma_{i,\text{SQE}}$ is the measurement error in the ith parameter from Super Quad Emit. Thus the convergence coefficient functions as a form of χ^2 test for the two parameter sets.

Because 4-D Quad Emit computes its findings downstream of QSM1, it is necessary to back-propagate the 4-D Quad Emit paramters through the scaled QSM1 for each value of the scaling factor to be considered. Also, for values of the scaling parameter which are different from the optimal value of 0.957, the errors on the Super Quad Emit quantities enlarge rapidly, and this can potentially "smear out" the test for convergence. One way to see whether this is happening is to compute the convergence coefficient with a fixed error set; for this study, the error values from the 0.95 scale factor case are used.

Figure 3.7 shows the value of the convergence coefficient for the two methods of computing it which are described above. Note that the two curves cease to track one another for values of the scale factor which are outside the acceptable range. Note also that both curves have a minimum at approximately 0.90, which is slightly outside the allowed range. Table 3.16 shows the values of the Super Quad Emit fit parameters and the propagated 4-D Quad Emit parameters at a QSM1 scale factor of 0.90. With the adjustment to QSM1, we see that all of the parameters save ϵ_1 and β_1 are within 1 standard deviation between the two sets of beam parameters. In Table 3.17, however, we see that the parabola scans used by 4-D Quad Emit are still not particularly well reproduced by the Super Quad Emit beam. This situation
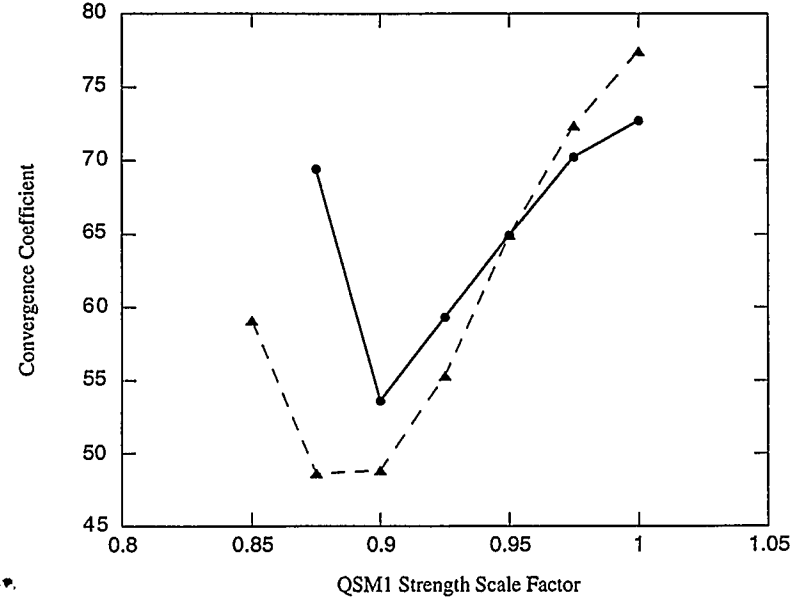
Figure 3.7: Convergence coefficient as a function of QSM1 strength scale factor. The coefficient computed with fit errors returned from Super Quad Emit for each value of QSM1 (triangles) and computed using errors from 0.95 scale factor (circles) are shown.

argues that the uncertainties on the 4-D Quad Emit measurement are smaller than those on the Super Quad Emit measurement: while the error-bars on the Super Quad Emit fit allow it to overlap the 4-D Quad Emit fit, nonetheless the actual beam parameters are sufficiently far from the fitted values to cause poor convergence with the 4-D Quad Emit data. Because QSM1 corrects the majority of the incoming coupling upstream of the quadrupole scanned by 4-D Quad Emit, the results from this method will probably always be better than from Super Quad Emit, which must reconstruct a beam whose vertical projected emittance is dominated by coupling from the larger normal- mode emittance. The remaining discrepancy, specifically the value of the horizontal normal-mode emittance and its corresponding betatron function,

Parameter (units)	4-D Quad Emit	Super Quad Emit	Deviation (σ's)
ϵ_1 (SLC units)	3.65	4.47 ± 0.16	5.2
β_1 (m)	76.3	65.3 ± 2.5	-4.4
α_1	2.43	2.43 ± 0.10	0.0
ϵ_2 (SLC units)	0.154	0.142 ± 0.048	-0.25
β_2 (m)	39.0	33.1 ± 9.9	-0.6
α_2	2.08	1.68 ± 1.00	-0.8
a	-0.545	-0.589 ± 0.30	-0.15
b	0.387	0.404 ± 0.062	0.27
c	0.144	0.096 ± 0.22	-0.21
d	-0.230	-0.240 ± 0.035	-0.29

Table 3.16: Side-by-side comparison of Super Quad Emit and 4-D Quad Emit results obtained on same day, assuming a QSM1 scale factor of 0.90.

remain unexplained. The 2-D Quad Emit measurement of the horizontal emittance recorded a value of 4.1 ± 0.1, in SLC units. No further measurements were made in the horizontal using 2-D Quad Emit; consequently, slow variations in the value at the level of 10% cannot be ruled out.

3.3 Beta Matching and Verification

As discussed in section 3.2.3, the action of QSM1 is sufficient to reduce the projected vertical emittance below the design value for the FFTB. Therefore, no effort is made to further reduce the total incoming coupling, and the beam is henceforth treated as though it were, in fact, totally decoupled. In particular, the five normal quadrupoles of the Beta Matching region are set to produce the desired beam conditions at the FP, based upon the projected emittance and Twiss parameters in each plane. This is perfectly legitimate, as the projected parameters describe the in-plane terms of the beam matrix (σ_{11}, σ_{12}, σ_{22}, and the corresponding vertical terms). As long as the transport matrix from the exit of QSM1 to the FP contains no coupling terms, the cross-plane terms of the sigma matrix will not contribute to the in-plane terms at the FP. In other words, the transport of the in-plane beam matrix terms is mathematically

Parabola Coefficient	Data Value	4-D Quad Emit Beam Value	Super Quad Emit Beam Value
x scan A	80.7 ± 8.3	69.2 (1.4)	103.2 (-2.7)
x scan B	-144.0 ± 0.16	-143.7 (-1.9)	-144.1 (-0.62)
x scan C	318 ± 48	290 (0.58)	298 (0.42)
y scan A	26.6 ± 0.93	25.6 (1.1)	21.5 (5.5)
y scan B	-147.5 ± 0.049	-147.4 (-2)	-147.0 (-10)
y scan C	29.95 ± 0.0	25.6	37.9
v1 scan A	75.0 ± 3.6	67.1 (2.2)	72.9 (0.58)
v1 scan B	-145.4 ± 0.055	-145.1 (-5.5)	-145.0 (-7.3)
v1 scan C	307.2 ± 0.0	277.8	309.7
v2 scan A	87.45 ± 7.7	73.4 (1.8)	97.8 (-1.34)
v2 scan B	-121.0 ± 0.10	-120.9 (-1.0)	-121.5 (5.0)
v2 scan C	231.1 ± 0.0	195.4	135.4

Table 3.17: Comparison of 4-D Quad Emit parabola coefficients with coefficients generated in simulated scans with 4-D Quad Emit beam and Super Quad Emit beam with QSM1 scale factor=0.90.

identical to the transport of the projected parameters, as long as the transport matrix is uncoupled.

The computation of quadrupole strengths needed to produce the correct beam parameters at the FP is straightforward, and can in principle be performed by any of a number of beam-optics programs (SAD, COMFORT, and DIMAD have all been used successfully for this operation). Once the match is implemented, however, the dual image of the FP at WS1 disappears, and further downstream images are required to ensure that the beam sent to the FP is properly matched. For this purpose, WS2 and WS3 are used. Sections 2.5.2 and 2.1.4 describe the design and placement, respectively, of these wire scanners. By placing the two scanners at the horizontal and vertical images of the FP, and separating these images by 2.85 meters, it is possible to verify both the positioning of the two waists, and the betatron functions at the waists.

In order to verify the positioning of the waists, it is necessary to scan the waist position and seek a minimum beam size on the wire scanner. This can be done in

several ways. The simplest is simply to scan a quadrupole in the beta matching region. This approach moves both waists, however, and also has effects on the betatron functions at the minimum, and is therefore unsuitable. The next simplest technique is to develop a cluster of multiknobs which move several quads by a fixed amount per knob turn, and which move only one of the waists (x or y), leaving both β_{min} valued unchanged. Because the optics of the beta match region are not constant, such a knob would have to be calculated for each matching exercise, and the knobs would only be strictly linear and orthogonal for small changes from the optics used in the calculation; once the knob has moved a large distance, the downstream quadrupoles in the beta match region cause the upstream quads to "see" a different R-matrix to the wire scanner, and change their effects.

Recently the SLC control system has added the capacity for "Irwin Knobs". These knobs take advantage of as many quadrupoles as are available to produce orthogonal transformations, and produce transformations which are linear and orthogonal over larger portions of the parameter space. This is done by dynamically recalculating the coefficients needed for the transformations over the range of the scan. Consider, for example, a waist scan over a range of $\pm W$ from the current location. The range is divided into smaller computational intervals, of length w. Over each of the small intervals, the correct changes in the scanned quads are computed to produce the pure waist shift desired, and the computation is repeated for the new values of the quads and the next interval of w. With a sufficiently small value of w, the computation approximates a continuous, orthogonal transformation.

The mathematics and concepts required for construction of the Irwin knobs are described in [42]. The concept, briefly, is this: if the existing transport matrix from the beginning of a beamline to the target wire scanner is given by R_0, then the action of an Irwin knob is to transform that matrix to $R_k R_0$, i.e., to transform it to the same matrix followed by an "effective knob matrix". The matrix R_k can be one of three

possible forms. For a waist shift, the form of R_k is a drift matrix:

$$R_k = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}. \tag{3.27}$$

For a change in β_{min}, the form of R_k is a pure magnification:

$$R_k = \begin{pmatrix} m & 0 \\ 0 & \frac{1}{m} \end{pmatrix}. \tag{3.28}$$

A third transformation is a "thin lens" transformation:

$$R_k = \begin{pmatrix} 1 & 0 \\ K_q & 1 \end{pmatrix}. \tag{3.29}$$

Note that, in a given plane of the beamline, these three transformations span the space of all possible matrices which have unit determinant. In order to freely adjust the waist and β_{min} in each plane, and require that the "thin lens" be constrained, six normal quadrupoles are needed.

The FFTB beta matching region contains only five normal quadrupoles. Therefore, a modification of the scheme described above is used. All Irwin knobs targeted at WS2 or WS3 constrain or scan horizontal and vertical waist, and horizontal and vertical β_{min}. Knobs targeted at WS2 also constrain the vertical term of the "thin lens", and those targeted at WS3 constrain the horizontal term of "thin lens". Unless the match has been quite poorly performed, the horizontal waist will be much closer to WS2 than the vertical waist; consequently, inserting a thin lens at WS2 will have relatively little effect in the horizontal, while it will have dramatic effects in the vertical. Therefore, the optics of the system will generally be more tolerant of a horizontal thin lens at WS2 and a vertical thin lens at WS3.

The utility of the Irwin knobs also depends upon the optics through the system being well known; specifically, the optics of the quadrupoles, and their excitations as a function of power supply current, must be known with better absolute accuracy than the incoming beam functions at all settings used during knob scans. Magnet-mover based lattice diagnostics were performed on the quadrupoles in the CCSX, BX,

CCSY, and FT regions, and these demonstrated that properly-standardized FFTB standard quads are set with an absolute accuracy of 0.1%, over a variety of setpoints. These diagnostics are described in Chapter 4. While the beta matching quads were never subject to the diagnostic in question, these quadrupoles are identical to those in the measured regions, and are therefore expected to be similarly well-behaved. Furthermore, no direct diagnostics of the quadrupole strengths have been performed after repeated scanning (which can potentially remove the quadrupole from the measured hysteresis curve). However, the beta matching quad power supplies are configured to perform "mini-standardize," which ensures that the magnets always approach new set-points from the correct direction. Additional description of the "mini-standardize" algorithm, as well as *indirect* evidence that it maintains the quadrupoles on their measured hysteresis curves, is also discussed in Chapter 4, and the measurement of the curves is described in Appendix B. Consequently, the beta match quads are expected to have an absolute accuracy of a few parts per thousand, which is considerably better than the knowledge of the incoming beam parameters. A mismatch at WS2 and WS3 is therefore more likely to be due to uncertainty of the incoming beam than due to uncertainty in the quad strengths, and the "Irwin Knob" scans become an appropriate tool to use in approaching convergence.

Figure 3.8 shows a scan, via the Irwin knobs, of the vertical waist position versus σ_y^2 on WS3. Two phenomena are noteworthy here. First, the position of the minimum is 12.4 centimeters downstream of WS3, indicated by the fact that the waist must be moved 12.4 cm upstream from the initial match to put the minimum spot on WS3. This is considerable mismatch, since the vertical betatron function is only 2.5 centimeters in the design optics at WS3. However, in this case the match was calculated using Twiss parameters from the Thin Lens fit, which has been seen to produce waist errors in Section 3.2.1. The second noteworthy factor is that the minimum spot size is 6.8 microns, while with the measured projected emittance (approximately 0.22 SLC units), the expected spot size is 740 nanometers.
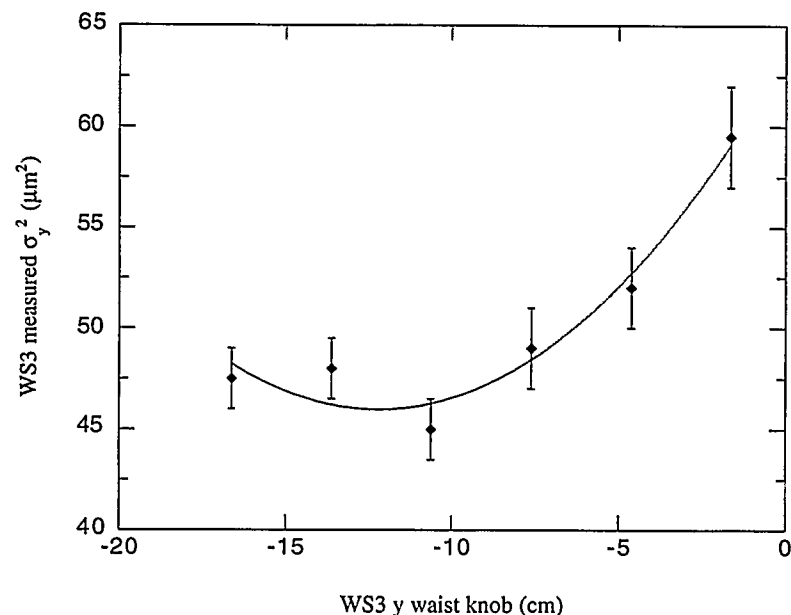
Figure 3.8: WS3 measured σ_y^2 as a function of the vertical waist Irwin knob. The location of the minimum indicates that the waist was originally located 12.4 cm downstream of the wire scanner. The beam size at the minimum is 6.8 microns. This measurement was made during the January, 1995 FFTB run.

As mentioned in section 2.5.2, by using the full set of wires on the WS3 yoke it is possible to reconstruct the full xy-plane projection of the beam at the wire scanner location, inlcuding tilt angle. When this was done, the culprit became clear: the beam was tilted by $1.33°$ relative to the wire scanner, with a minor-axis RMS size of $1.5\mu m$. Figure 3.9 shows WS3 σ_y^2 as a function of QSM1 strength. Note that, during beta matching, the projected vertical emittance was roughly minimized by setting QSM1 to 1.3 kilogauss; according to WS3, however, a strength of -0.9 kilogauss minimizes the WS3 vertical spot. Finally, Figure 3.10 shows a waist scan performed with the new value of QSM1. Note that the waist must be shifted an additional 4 cm upstream;
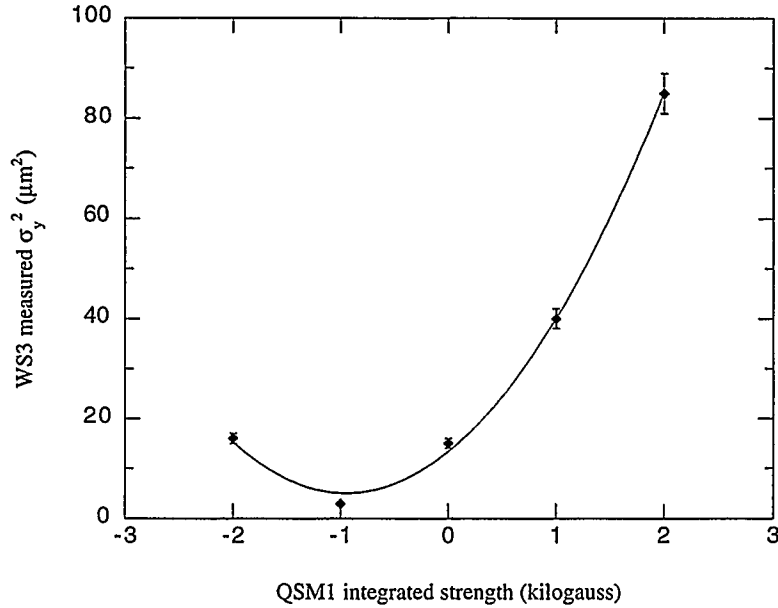
Figure 3.9: WS3 measured σ_y^2 as a function of QSM1 integerated strength, from the January 1995 FFTB run. The minimum occurs at -0.9 kilogauss, rather than the 1.3 kilogauss determined by WS1 optimization.

however, the flatness of the parabolas in Figures 3.8 and 3.10, as well as the absence of one "wing" of each parabola, makes this shift equivocal. Note also that the minimum of Figure 3.10 is extremely broad, which is an artifact of the $4\mu m$ wire size. The value of "A" in the two parabolas, which is a measure of the angular divergence, is quite consistent between the two scans (0.1100 ± 0.0030, versus 0.1127 ± 0.006). Since the divergence measurement is dominated by the scans on the wings of the parabola, where the scanner response is more linear, these values are expected to be well-measured by this technique.

The disagreement between WS1 and WS3 on the correct value for QSM1 is present in all beta-matching episodes between September 1994 and March 1995. While the
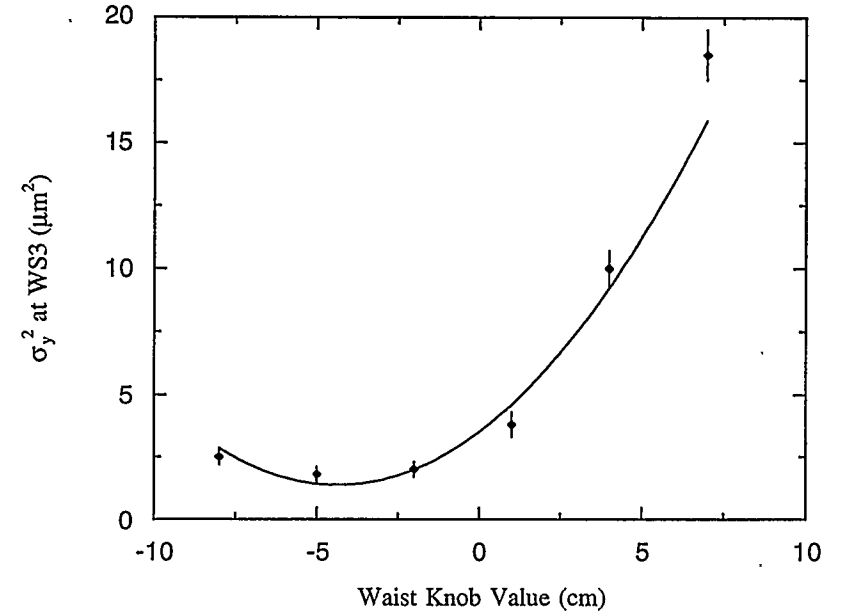
Figure 3.10: WS3 measured σ_y^2 as a function of the vertical waist Irwin knob, after setting QSM1 to -0.9 kilogauss. Note that the divergence is virtually the same, as expected, but the minimum is now closer to 1 micron.

optimal value from WS1 shifts between 0.7 and 1.3 kilogauss, the optimal value from WS3 is between 0 and -1.0 kilogauss, with no clear correlation between the two. One suspect in this dilemma was a WS3 rotation angle: since the roll angle reported by WS3 is always around +1°, it stands to reason that the scanner may be rotated in the xy plane by a similar amount. An experiment was performed to test this hypothesis: using the magnet mover of QT1, a horizontal bump was introduced at WS3, which was then closed downstream by QT3 and QT4 movers; at each value of the bump, the BPMs on either side of WS3 were read out for 100 pulses, in order to ensure that the bump was truly horizontal, and that no inadvertant vertical motion was being introduced; and finally, each wire of WS3 was scanned through the beam. The wire scanner software returns the value of the scanner stepper motor at which
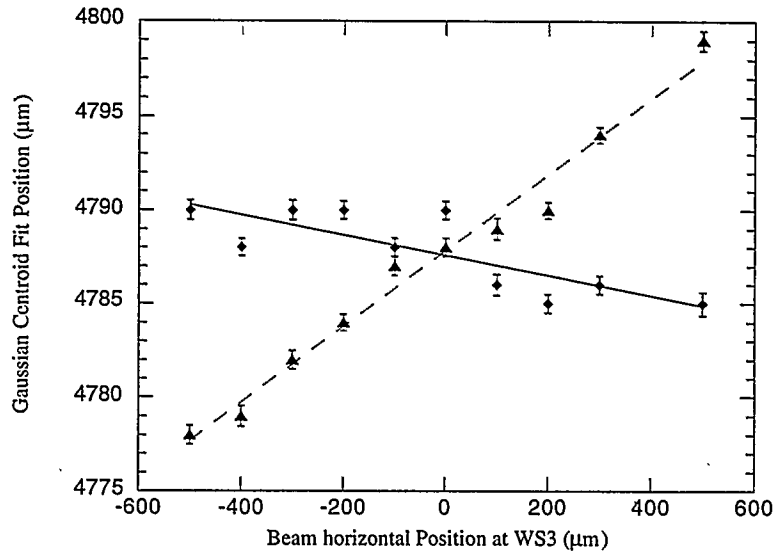
Figure 3.11: Location of the beam centroid at WS3, as determined by fitting an asymmetric gaussian to the beam, as a function of the beam horizontal position. The unrotated wire # 1 (diamonds) and 1.4° wire # 6 (triangles) are shown.

the centroid of the beam was found, allowing a correlation to be developed between the beam vertical centroid location and the amplitude of the horizontal bump. In this way, all 6 wire angles could be measured. The four odd-angle wires allowed a further test of the method in that these wires were expected to have non-zero rotation angles. Since the wire scanner software returns two centroid numbers (center of an asymmetric gaussian fit and center of the actual distribution), both sets of numbers were analyzed in parallel.

Figure 3.11 shows the variation in centroid location as a function of bump amplitude for wire 1 (nominal zero roll) and wire 6 (nominal 1.4° roll). The nominally flat wire does show a small roll, with the South end of the wire higher than the North

WS3 xpos knob value (µm)
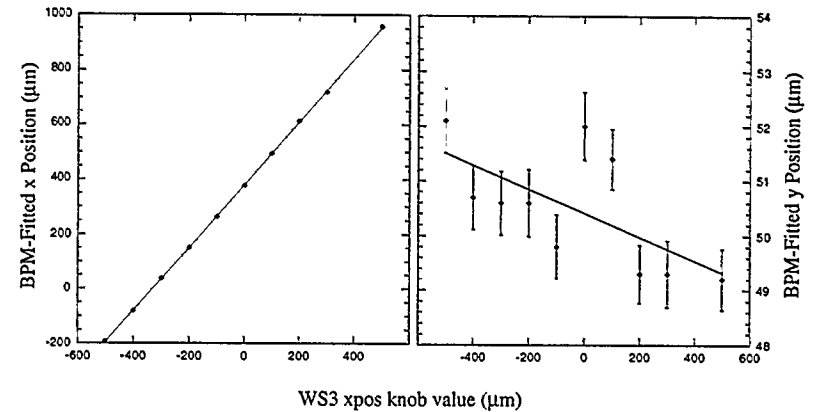
Figure 3.12: Movements of the beam at WS3 as a function of the QT1 horizontal mover knob. The positions shown are reconstructed using data from adjacent BPMs.

end by approximately 5 milliradians. The 1.4° wire is higher on the North by approximately 20 milliradians, which is slightly different from the 24 milliradian design rotation. However, the pre-installation testing of the scanner revealed that the yoke was systematically rolled by 3 ± 1 milliradians, with South end higher[43]. Table 3.18 shows the expected wire angles with this correction, the angles measured by the gaussian center, and those measured by the distribution center. Note that within errors, the three angles agree for all wires.

Because the two BPMs read out for this exercise are separated by drifts only from WS3, it is possible to form a linear combination of the readings which represents the beam position at WS3. This position, in horizontal and vertical, is shown in Figure 3.12 as a function of knob position. The slope of the horizontal correlation is 1.152 ± 0.004, rather than the expected 1.0; however, the BPMs scale factor is measured by the mover lattice diagnostic described in Chapter (), and the BPMs are found to systematically over-report by 12.4%. Furthermore, the BPM scale factors are found to vary by $\pm 5\%$, based upon bench tests [44], and the free-standing BPMs near

Wire	θ_{theor}, mrad	θ_{gauss}, mrad	θ_{moment}, mrad
1	-3	-5.3 ± 1.7	-4.1 ± 0.9
2	-15	-16.1 ± 1.5	-16.8 ± 1.2
3	+9	$+10.3 \pm 1.6$	$+8.9 \pm 1.1$
4	-3	-3.6 ± 0.9	-3.1 ± 0.4
5	-27	-29.0 ± 1.4	-28.3 ± 0.6
6	+21	$+20.0 \pm 1.2$	$+21.6 \pm 1.0$

Table 3.18: WS3 wire rotation angles measured by the QT1 horizontal mover bump technique described in the text. Shown are the expected wire angles from the bench test (θ_{theor}), the angles measured by fitting an asymmetric gaussian to the distribution (θ_{gauss}), and the angles measured by computing the center of the distribution directly (θ_{moment}). Positive angle indicates the North end of the wire is higher from the ground than the South end.

WS3 were never bench-tested for this factor. It is therefore possible that the WS3 BPMs differ from the ensemble average by an amount to explain the remaining 3% discrepancy. The vertical position data shows a correlation with the horizontal motion of -0.0022 ± 0.0009; this could be due to rolled BPMs or an imperfect horizontal knob; the correlation is far from convincing. In any event, the WS3 correlations indicate that the wire scanner has no significant roll at the level expected ($1° = 17.5$ mrad), and the BPMs indicate that the bump is performing nearly perfectly. This is a mixed blessing: while the experiment argues that wire scanners with extremely tight roll tolerances such as WS3 can be constructed and installed reliably, it also exonerates the prime suspect in the QSM1 dilemma.

What other errors in the beamline could cause such a discrepancy in coupling measurements? One possibility is that an upstream quad is rolled or shorted turn-to-turn. If all other upstream elements were held constant over the FFTB lifetime, a static quad error would give rise to reproducible QSM1 settings. However, the optics from QSM1 to WS3 are changed from run to run because of the different incoming betatron conditions, and the incoming coupling is also different, indicated by changes in the optimal value of QSM1. Consequently, even a static magnet roll could give

rise to varying values of QSM1 which optimize the WS3 spot. Magnet rotations are considered further in Chapter 4.

Returning to Figure 3.10, we see that the vertical beta match verification has a particular difficulty, specifically that the matched beam is smaller than the wire's resolution. Furthermore, the beam at the intermediate waist is potentially diluted by negative chromaticity from the CCSX sextupoles, or other aberrations. This makes verification of the betatron function itself, β_{min}, difficult. It is therefore necessary to measure the *divergence* of the beam, and to determine the matched betatron function from the divergence and the measured emittance, since $\theta_{waist} = \sqrt{\frac{\epsilon}{\beta_{waist}}}$. The divergence can be measured in two ways. The most straightforward is the vertical beam size on WS2, measured by scanning all six WS2 wires and reconstructing the xy ellipse. Assuming that the vertical minimum is located at WS3, the vertical size at WS2 is given by:

$$\sigma_y^{(WS2)} = \sqrt{\epsilon_y \beta_y^{(WS3)} + L_{WS2 \to WS3}^2 \theta_y^2} = \sqrt{\epsilon_y \beta_y^{(WS3)} + L_{WS2 \to WS3}^2 \frac{\epsilon_y}{\beta_Y^{(WS3)}}}. \quad (3.30)$$

Because the term due to beam size at WS3 is small compared to the divergence term, the approximation can be made

$$\sigma_y^{(WS2)} \approx \sqrt{L_{WS2 \to WS3}^2 \frac{\epsilon_y}{\beta_Y^{(WS3)}}}. \quad (3.31)$$

The second method is to measure the beam size on WS3 as a function of the waist position. If we replace $L_{WS2 \to WS3}^2$ with the waist knob (which is in essence a variable drift), we see that the square of the beam size depends quadratically upon the waist knob setting. Thus, the parabola parameter "A" represents the square of the divergence in the appropriate units. The units of σ_y^2 are μm^2, while the units of the Irwin waist knobs are centimeters; therefore, parabola parameter "A" has units of μm^2 per cm^2, and $100\sqrt{A}$ has units of μrad.

Figure 3.13 shows the parabola from a WS3 waist scan and the beam ellipse at WS2 side-by-side, from a measurement made in March of 1995. The fitted value of "A" is $(2.72 \pm 0.48) \times 10^{-2}$, which yields a divergence of $16.5 \pm 1.5 \mu$rad. The vertical
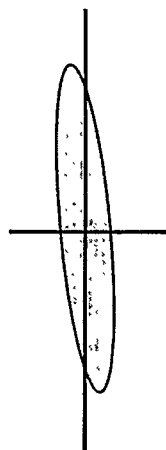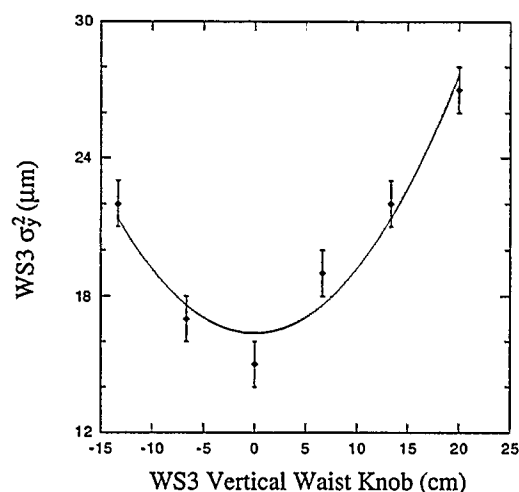
Figure 3.13: WS3 divergence measured in two ways: on the left is a scan of the Irwin waist knob, on the right a measurement of the xy-projection of the beam ellipse at WS2. These scans were performed during the March 1995 FFTB run.

beam size at WS2 is $40.8 \pm 2.5\mu$m, which yields a divergence of $14.3 \pm 0.9\mu$rad, which is acceptable agreement. The measured emittance at the end of the linac yields a projection to the FFTB entrance of 3.0×10^{-11} meter-radians, which can be combined with the divergence measurement to indicate a betatron function at WS3 of approximately 13 centimeters. At the time, the optics were adjusted to obtain a β_y^* of 1 mm, which would indicate a β_{min} at WS3 of 25 centimeters. This result demonstrates the utility of the intermediate-waist divergence measurements in optimizing the beta match. In principle a single adjustment of the magnification knob would have been capable of correcting the optics to those desired; however, the FFTB was operating in a "signal-seeking" mode, and the discrepancy in beam size at the FP was not crucial.

A similar set of measurements may be made using WS2's horizontal size and waist knob, and WS3's fitted horizontal size. In this case, however, the beam size on WS2 at

the waist is expected to be closer to 5 microns, and therefore the measured minimum size at WS2 is more meaningful in the design optics than the measured minimum size at WS3. Because the optical conditions are less exacting in the horizontal, beam matching in this plane almost always converges with little more than a waist scan.

The techniques described in this chapter are sufficient to ensure that the incoming beam is properly matched to the desired conditions at the FP. In principle, the global tuning described in Chapter 5 can then be applied, using the beam size monitors. Prior to such an operation, however, it is necessary to verify that the capture conditions outlined in Chapter 2 are met, particularly the quadrupole strength and magnet alignment conditions. In Chapter 4, we shall see how this is accomplished.

Chapter 4

Local Beam-Based Diagnostics

Chapter 2 argued that an important aspect of tuning the FFTB is ensuring that the capture tolerances in Table 2.3 are met before attempting to apply global tuning procedures. In addition, local correction of beamline aberrations and errors is always preferable to global corrections, since a cascade of local errors can add up to a level which is too large to correct with global knobs.

An additional feature of local diagnostics is that they can be designed around the BPM system, rather than the wire scanners and beam size monitors. The BPMs are fast, non-invasive, and can be read out synchronously on each pulse. In addition, the BPMs in use in the FFTB have a dynamic range of about 3 orders of magnitude (microns to millimeters), which is larger than either the wire or BSM systems. This makes the local diagnositcs less dependent on proper matching of the incoming beam, and eliminates the "bootstrapping" frequently associated with such devices (i.e., tuning the system well enough to get a signal in the first place).

The FFTB has four local, BPM-based diagnostics: quadrupole alignment, dispersion measurement, quadrupole strength measurement, and CCS sextupole alignment. Each of these shall be discussed in turn.

4.1 Preparation for Local Diagnostics

The first step in preparing the FFTB for local diagnostics is to set a beta-match optics which reduces the FP divergences significantly. This reduces the beam size in almost all magnet apertures, which consequently reduces both pulse-to-pulse jitter and "spray" on the BPMs. Additionally, reduction of the beam size in the magnets increases the range through which a magnet strength may be scanned before a downstream aperture limit is encountered, which is an advantage for quadrupole alignment. The optics selected for this set of diagnostics gives $\beta_x^* \times \beta_y^* = 3\,\mathrm{cm} \times 1\,\mathrm{cm}$ with the nominal incoming beam. Note that no attempt is made to actually match the optics to the incoming beam, as the actual sigma matrix at any given point is of relatively little importance, so long as it is close enough to the expected value that the quadrupole scan range selected for quad alignment is still acceptable from an aperture standpoint.

After setting the beta-match optics as desired, the FFTB line is "standardized": the magnets are ramped from a low current to maximum current and back down to the low current 5 times, then ramped from low current to the desired set point. This ensures that the magnets are on their measured hysteresis curves, and therefore that they are actually at their desired setpoints.

Finally, a sample of BPM data is taken to ensure that linac collimation is adjusted properly for elimination of BPM backgrounds. Early experiences with the FFTB BPM system indicated that poor adjustment of Sector 28-30 collimators resulted in degradation of BPM precision due to particles impacting the striplines. The resolution is measured by acquiring data on two or more BPMs which have relatively little phase advance between them (such as BPMs 50 and 1010, which are separated by 1 meter in the Beta Match section, or BPMs 1100 and 2020, which are in the quads bordering SF1A). The data from such BPMs should correlate linearly, with an RMS deviation equal to $\sqrt{2}$ times the BPM intrinsic resolution. Such a correlation is shown in Figure 4.1. The RMS deviation is 1.45 μm, indicating a BPM resolution of 1 μm. The bunch
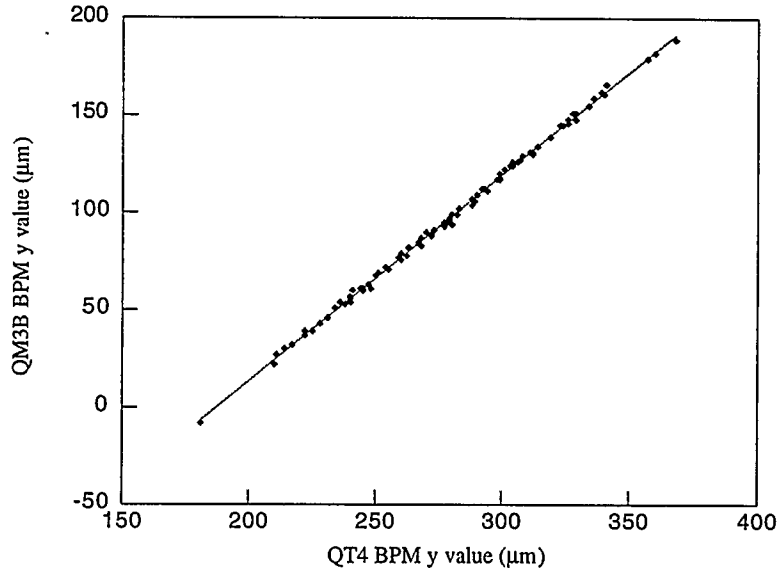
Figure 4.1: Correlation of beam vertical positions measured at the BPMs in QT4 (horizontal axis) and QM3B (vertical axis). These BPMs border sextupole SD1A, and are expected to track perfectly within the resolution of the BPMs. The RMS deviation of the data from the fitted line is 1.45 μm. This correlation is fairly representative of points in the machine with such BPM pairs.

charge at the time of this measurement was approximately 8×10^9; therefore, at the design charge of 1×10^{10}, the expected resolution is better than the design specification of 1 μm.

4.2 Quadrupole Alignment

The beam-based alignment of quadrupoles in the FFTB uses the conventional "shunt" technique familiar to accelerator physicists. Consider a situation as depicted in Figure 4.2, in which the electron beam passes off-axis through a quad. In this case the beam
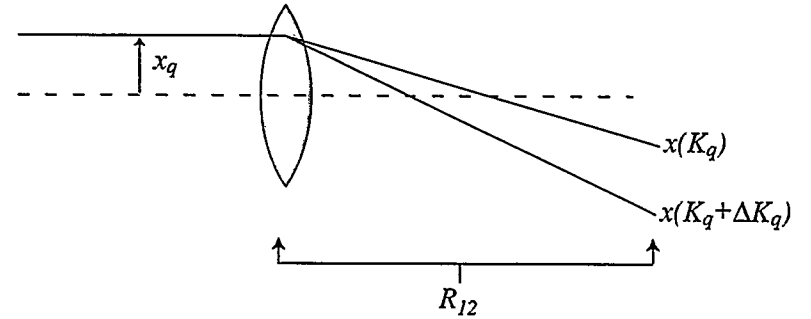
Figure 4.2: Schematic of quadrupole alignment technique. Changing the strength of the quad changes the kick seen by the beam centroid, which in turn changes the position at a downstream BPM location.

centroid sees a kick, whose amplitude is given by:

$$x' = -K_q x_q, \tag{4.1}$$

where K_q is the integrated quad strength, Δx_q is the offset of the beam centroid from the quad center, and once again the thin-lens approximation is used for simplicity of description while a full, thick lens computation is performed by the algorithm. Changing the quadrupole strength in this case will change the kick angle:

$$\Delta x' = -\Delta K_q x_q, \tag{4.2}$$

which translates to a change in position at a downstream BPM:

$$\Delta x_{BPM} = R_{12}^{q-BPM} \Delta x' = -R_{12}^{q-BPM} \Delta K_q x_q, \tag{4.3}$$

and thus the initial offset of the incoming beam can be determined from scanning the quadrupole strength, measuring the change in position on a downstream BPM, and dividing by the known R_{12} between the two. Because all quadrupoles in the FFTB can be moved via magnet movers from the control system, the magnet can then be moved into aligned position, and the independent power supplies for each quadrupole allow the measurement to be performed without any use of shunts, boosts, or trim windings.
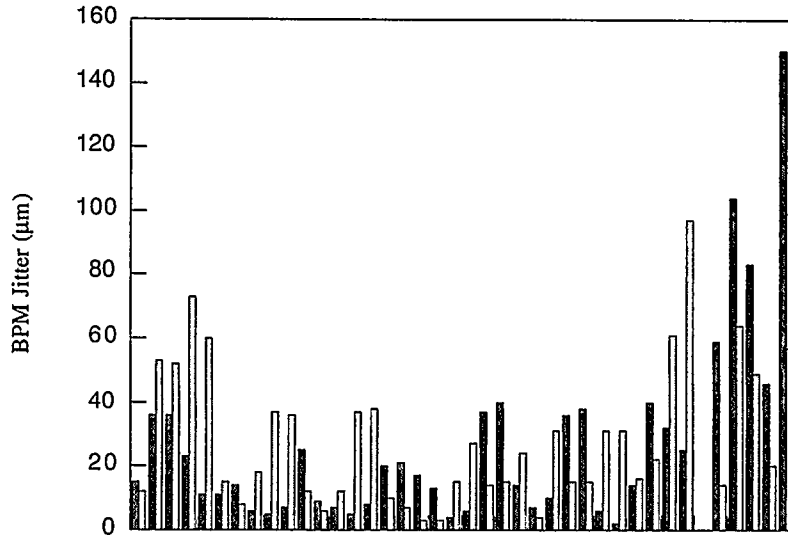
Figure 4.3: Jitter amplitude at all FFTB BPMs, from BPM 30 to the last extraction BPM, in quad-alignment optics. Both horizontal (light) and vertical (dark) are shown.

4.2.1 Initial Algorithm Design

While the alignment measurement is conceptually simple, in practice there are several complications which must be addressed. The first of these is incoming jitter: Figure 4.3 shows the horizontal and vertical RMS jitter at all FFTB BPMs during the March 1995 alignment operation. The jitter amplitude is as large as 40 μm in some locations, while the BPM resolution is close to 1 μm (see Figure 4.1). This implies that the incoming position and angle must be computed for each pulse.

The jitter also implies two further difficulties related to reference pulses. The first is this: the alignment computation in Equation 4.3 assumes that the beam pulse upstream of the quad being scanned follows the trajectory to which the magnets wish to be aligned. In other words, given a perfect system with no jitter, scanning the quadrupole in Figure 4.2 and applying Equation 4.3 to the change in a downstream

BPM will indicate a correction which moves the quad onto the line of the beam; however, if the incoming beam is jittering by some amount, then the pulse which passes through when the quad is at its nominal strength may potentially be a "flier", and not a good pulse to align to. The second difficulty is ensuring that the pulse to which the next quad downstream is aligned to is identical to the one to which the current quad is aligned to.

All of the difficulties enumerated above were anticipated and addressed in the initial design of the quadrupole alignment algorithm. The reference pulse issue was addressed by using the BPM system (rather than correlation plots) to acquire the reference orbit prior to data acquisition proper. The reference orbit thus acquired was averaged over 100 pulses and saved. The saved orbit was then compared to subsequent pulses for some time to ensure that the saved orbit was fairly representative of the incoming beam, and not systematically offset or off-energy from the majority of incoming pulses. Because the FFTB launch and energy feedback systems maintain a constant orbit (with a time scale short relative to 100 pulses), two such highly-averaged orbits which are representative of the incoming beam are quite likely to be identical. As a further correction to the reference orbit problem, a single reference orbit is used to align between 3 and 6 quadrupoles: after the reference orbit is acquired, the Correlation Plots application in the SLC Control Program (SCP) steps the first quad and acquires all BPMs at each setting of the quad; the first quad is returned to its original value, and the next quad is stepped, until all desired quads have been stepped and data acquired for all BPMs at all quad settings. This procedure ensures that the quads stepped in this way will all be aligned to the same incoming orbit, so that the alignment of the FFTB will be at least "piecewise-continuous".

In summary, the data acquisition algorithm is as follows:

- Acquire reference orbit, $N_{avg} = 100$, compare to subsequent pulses
- For each quad in the current sequence (3 to 6):
    - Step the quad through 3 strengths, acquire all BPMs at each strength

       – Return the quad to original strength

  • Submit quad strengths and BPM readings to fitting engine.

The fitting engine used for FFTB Quad Alignment is OPTFIT, a MINUIT driver which uses first-order (R-matrix) beam transport to compute χ^2 values (See section 3.2.4). OPTFIT subtracts the reference orbit from all BPM data, leaving only differences from the reference orbit. This eliminates all DC phenomena such as the Earth's magnetic field, BPM offsets, etc. The engine then calls MINUIT for minimization, using the misalignments of each quad and the incoming position vector of each pulse, $\vec{x_0} = (x, x', y, y', \delta)$, as parameters to be fitted ("fitvars," in the language of the control system), and the BPM readings as the data to be matched. The algorithm of the MINUIT call is as follows:

  • Pass new guess values for quad misalignment and position vectors to math subroutine

  • insert position vectors and quad misalignments into appropriate data slots

  • For each pulse acquired

      – Use quad strengths to compute new R_{ij}'s for stepped quads

      – Use initial position vector to track pulse through beam line, compute difference between tracked ray and data at each BPM

  • compute χ^2 as sum of squares of differences between ray tracking and BPM data

  • Iterate above with new guess values until accurate error matrix can be computed, locate minimum in parameter space.

4.2.2   Refinements to the Algorithm

Experience with the beam-based measurement process soon dictated several additional refinements to the original fitting algorithm. These included an improved approach to solving the incoming trajectory, cleanup of BPM data before the main fit call, and fitting of initial kick angles in a segment.
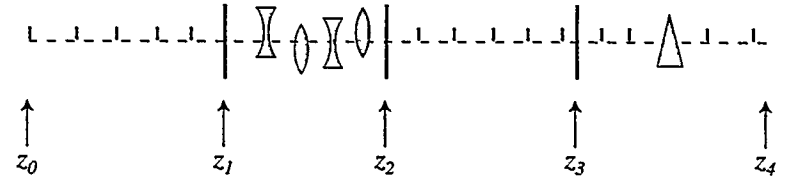
Figure 4.4: Division of the beamline into several regions. The region between z_1 and z_2 contains all the quads whose strengths are changed for the current acquisition; all other regions have constant transport matrices throughout the acquisition.

Solving the Incoming Trajectory

An examination of the fitting algorithm above shows that each pulse acquired increases the number of fitvars by 5, for computation of the incoming trajectory. Because MINUIT is attempting to compute a square, symmetric error matrix, the approximate computing time will scale as $n(n + 1)$, where n is the number of fitvars. This means that using many pulses, which is helpful in beating down statistical error, will also dramatically increase computational time. The number of pulses also quickly becomes limited by MINUIT, which will only accept 50 fitvars (and, it has been remarked, will only reliably converge for about 8 [45]). Even worse, the correlation between fitvars of interest (i.e., quad misalignments) and fitvars we don't actually care about (i.e., initial vectors of pulses) is considerable. MINUIT does not differentiate between data from upstream of the first stepped device and data downstream; thus, *even a dataset with no BPMs upstream of the first scanned quadrupole will produce a solution for the incoming ray in MINUIT.* In a case such as FFTB quad alignment, MINUIT will frequently conclude that all variation in the data is due to incoming offsets, none of the quads are misaligned, the errors on fitted quantities are enormous, and the χ^2/ν also enormous.

The incoming-ray problem can be solved by considering the beamline diagram in Figure 4.4. Upstream of the first stepped quadrupole (from z_0 to z_1) the transport matrices are constant; downstream of the last stepped quadrupole (from z_2 to z_4) the
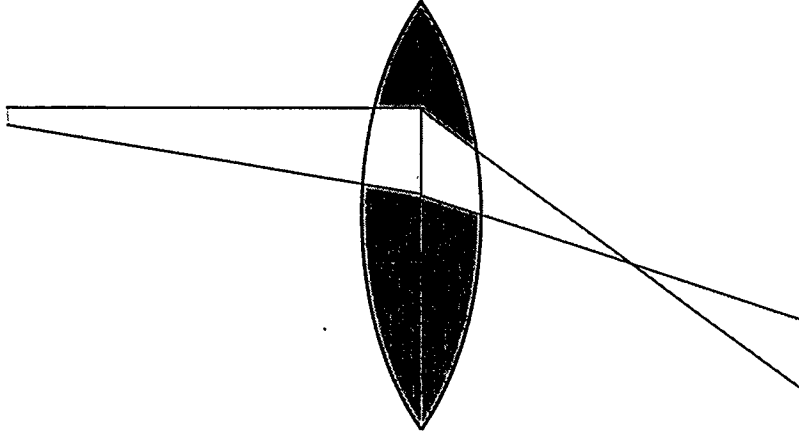
Figure 4.5: An incoming ray with some uncertainty in x and x' incident upon a quad. The lightly shaded region represents the total area in which the beam may be located, both upstream and downstream of the quad. The propagation of the incoming ray's errors must be properly folded into the OPTFIT calculation of quad alignment.

transport matrices are also constant, and in particular in the dumpline (from z_3 to z_4) only drifts and vertical bend magnets are present. Therefore, given a reference orbit and a data orbit, it is possible to compute the energy deviation of the data orbit from the BPMs in the dumpline, and the geometric deviation of the data orbit from the BPMs upstream of the first stepped quadrupole. The data used to compute these deviations can be removed from the MINUIT fit, and the deviations themselves can be used as initial conditions to the fit. This reduces the number of fitvars in the MINUIT minimization call to just those of true interest to the machine physicists.

The difficulty in removing the initial vector fitting from the MINUIT fit is that a MINUIT multi-parameter fit explicitly includes all correlations in error computation. Consider for example the situation in figure 4.5, in which a beam with an imperfectly-determined initial vector is incident on a quadrupole. The beam will pass through the quad, with an offset from the centerline which is undetermined, and this in turn will cause the exiting ray to be further undetermined. MINUIT automatically includes

this effect in its determination of the quadrupole alignment precision; any external computation of the initial vector must include a mechanism for folding these errors into the MINUIT calculation.

The method chosen to compute the incoming vectors is a linear least-squares minimization of χ^2, which automatically yields a covariance matrix for each pulse. This computation is exceedingly fast, as each pulse requires construction of a 5×5 curvature matrix from known beamline and data constants, a matrix inversion, and a matrix-vector multiplication. Note that each pulse requires its own curvature matrix because not all BPMs will return data and verify good status on each pulse. Furthermore, the energy is extracted from a 3-parameter fit (y, y', δ at the first energy feedback BPM) to decouple the fit from any upstream steering when quads are stepped. For m pulses and q quads being fitted in x and y, the scaling for a combined fit is $(5m + 2q)(5m + 2q + 1)$, whereas for the separated fit the computational scaling is $(4 \cdot 5 + 3 \cdot 4)m + 2q(2q + 1)$. For datasets with a large value of m, the separated fit wins due to its linear computation scaling; FFTB Quad Alignment datasets typically have m values between 130 and 150. In addition to the computational advantages, the separated fit is incapable of getting "lost": deviations in downstream BPMs due to changes in quad strengths never get misinterpreted as changes in the initial vector.

The inclusion of initial vector errors into the main fit can be accomplished by noting that, at each BPM, there will be two sources of deviation between the tracked ray and the data. The first is the BPM resolution, which is electronic and approximately the same for all data points, which we can define as σ_{BPM}. The second is the propagation of the error in the tracked ray, which we can define as σ_{trac}. Let us assume that for a given ray, the initial vector $\vec{x}_0$ and the associated error matrix, $\sigma_{0,ij}$, are known. The propagation law for the vector is given by:

$$x_{trac,i} = R_{ij} x_{0,j},$$

(4.4)

where $\vec{x_{trac}}$ is the position after tracking, $\vec{x}_0$ the initial position, and summation over

j is implicit. The propagation for the error matrix is given by:

$$\sigma_{trac,ij} = \frac{\partial x_{trac,i}}{\partial x_{0,k}} \frac{\partial x_{trac,j}}{\partial x_{0,l}} \sigma_{0,kl} \qquad (4.5)$$
$$= R_{ik} R_{jl} \sigma_{0,kl}.$$

Equation 4.5 can be rewritten in a familiar matrix form:

$$\sigma_{trac} = R \sigma_0 R^t, \qquad (4.6)$$

i.e., the transport law for the error matrix is identical to the transport law for the beam matrix.

Once the error matrix for each pulse is established, the transport matrix for each pulse, from the beginning of the beamline to each BPM, is computed (note that this is variable because of the scanning of intermediate quadrupoles). The transport matrices are used to transport the pulse's initial-vector error matrix to each BPM, where the intrinsic error σ_{BPM}^2 is added to the appropriate error matrix term, either $\sigma_{trac,11}$ or $\sigma_{trac,33}$; the square root of this quantity is the effective uncertainty, σ_{eff}, of the BPM data point in question. By using σ_{eff} to weight the data in the MINUIT fit, the uncertainty in the fitting of the incoming beam is taken care of.

In principle, the fitting algorithm automatically locates all BPMs between the beginning of the line and the first stepped quadrupole to use for fitting the incoming geometric ray, and uses the extraction line BPMs to fit the energy. In practice, the line upstream of the first quadrupole frequently includes bend magnets. Therefore the BPMs used to reconstruct the incoming x and x' potentially include BPMs with horizontal dispersion. The order of operations within MINUIT extracts the energy before the geometric terms of the initial vector, and the dispersive offset at each BPM is subtracted before the fit of the geometric oscillation. However, the computation of the error matrix is not quite clever enough to properly include the $x\delta$ and $x'\delta$ correlations; in order to avoid this problem, only BPMs upstream of the first bend are used to fit the incoming oscillation, and all BPMs from the first bend to the first

stepped quad are dropped from the fit. As we shall see, this does not result in an unacceptably coarse measurement of the quadrupole alignments.

Detecting Fliers in Dataset

Another area of concern is the possibility that individual pulses, individual BPMs, or individual readings may be dominated by various noise sources, which can drastically alter the fit of the quad misalignments. In the original algorithm, only the value of χ^2/ν gives any insight into this possibility; with 150 pulses, 30 BPMs, and x and y data, this is a daunting problem to solve by hand. Some form of pre-fitting examination of the data, which is capable of recognizing bad data and eliminating it from the fit, is required. In order to be fully effective, the mechanism must use information which does not directly correlate to the fitvars. Using the χ^2 contribution of the data after the fit, for example, would not be appropriate, as the fit has by that time already been tainted and it is possible that bad χ^2 contributions are due to the fit being "pulled" by the real culprit.

As mentioned previously, the region between the last stepped quad and the energy feedback BPMs is a region of constant transport properties. Using the same least-squares mathematics used to extract geometric initial vectors for the beam pulse by pulse, we can compute the geometric initial vector for each pulse starting at location z_2 in Figure 4.4. This computation uses the energy for each pulse computed by the extraction line BPMs to subtract dispersive terms, and the product $\eta\sigma_\delta$ is added in quadrature to the intrinsic BPM resolution to de-weight data appropriately (note that in this case η is the R_{16} element from z_2 to the BPM, not from 50B1 to the BPM). Once these vectors are computed, each BPM's χ^2 contribution is determined. Any BPM with a χ^2 contribution greater than 64 per data point is dropped from all further use, and the trajectories at z_2 are recomputed. This time, any BPM with a χ^2 contribution greater than 16 per data point is dropped. A final computation of the trajectories at z_2 is performed. This time, individual data points with a χ^2

contribution greater than 16 are dropped.

The first cut is made because a single BPM which is sufficiently disturbed can pull an entire fit above the threshold for the second cut. The softer cut is made in an attempt to eliminate such "monster" BPMs, such that on the second pass the remaining BPMs might converge.

The cuts described are not ideal. For one thing, a BPM which is in an optical location with sufficient lever-arm can still pull the entire fit to one side, resulting in all BPMs in the region having χ^2 contributions over 64. In this case, all the BPMs will be eliminated and no data will remain. This is prevented by only analyzing sequences of fewer than 11 consecutive BPMs, rather than performing the cuts on the entire beamline at once. In such a situation, in which a long string of BPMs have been cut from the fit, determining which ones are the real problem remains a trial-and- error problem, albeit a smaller one.

A more appropriate schedule of cuts would be as follows:

- Perform the computation of x, x', y, y' as described above, obtain the χ^2 contribution of every data point

- For every pulse which includes a data point with χ^2 contribution > 16, do the following:

  - Repeat the fit with each data point supressed and record χ^2 for each fit
  - Drop the point which yields the smallest chi^2
  - Repeat until no points have χ^2 contribution greater than 16

The scheme above does not get fooled by BPMs in crucial locations, requires only one iteration, and looks exclusively at the χ^2 contributions to a given pulse. The scheme as implemented looks first at the contributions of the BPMs across all the pulses, then at the contributions within each pulse. The disadvantage is that the former scheme is computationally immense in the presence of BPMs which are systematically misbehaving. With m pulses and n BPMs, a single bad BPM will require $m \cdot n$ executions of the code to compute geometric trajectory and χ^2 contributions; a pair of bad BPMs will require $m \cdot (2n - 1)$ executions. The scheme as implemented

always requires $3m$ executions of the code in question. Note also that the algorithm is a boolean algorithm in that data points are not arithmeticlly altered, but simply passed or not passed to the fitter; that the former scheme is almost guaranteed to remove bad data points, though the risk to "innocent bystander" data points is much higher; and that execution speed is essential, in that quadrupole alignment is intended to run on-line in the SCP. As a result, the former algorithm is acceptable for these purposes. Should a faster computer become available to the control system, reworking the cuts in OPTFIT is probably desirable. Given the highly parallel nature of the improved algorithm above, a parallel-processing computer would be ideal for the job.

Finally, it is worthwhile to note that neither algorithm can perform cuts on the data coming from BPMs inside the stepped quads. In this region there is no way to decouple the orbit changes in the BPMs from the misalignments of the quadrupoles. However, the output from OPTFIT includes a contribution to the normalized χ^2 from each BPM in each plane. This can be used to determine whether BPMs within the quads are misbehaving, and if so they can be deselected and the fit repeated.

Fitting Segment Angles

Consider a beamline with an alignment situation such as that in Figure 4.6. The quadrupoles are perfectly internally aligned, but the incoming beam does not pass through them; however, the first quad is misaligned in such a way that the outgoing beam from this quad passes through the line of the remaining quadrupoles. In this case, it would be inappropriate to simply move all the quads onto the ray of the incoming beam, since this would require a monotonically-increasing movement of subsequent magnets. Even a small angle between the magnet line and the incoming beam will quickly snowball into an uncorrectable problem. Moreover, if the beam has actually been steered to the dump before alignment begins, such a situation is likely to exist somewhere in the line.
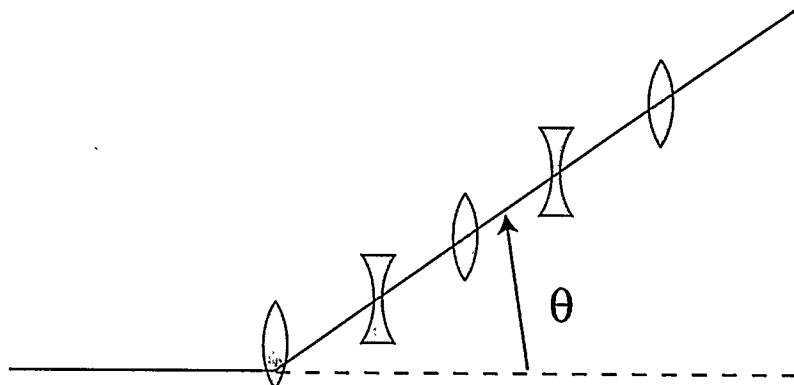
Figure 4.6: A line of quadrupoles which are perfectly aligned to one another, but with an angle θ between the incoming beam and the line of the magnets.

Existing techniques for beam-based alignment of quadrupole systems resolve this problem by defining the first and last BPMs in the region of interest as the line to which the quads will be aligned [46],[47]. In these techniques, only a single execution of the fitting algorithm is permitted, which extracts the misalignments of all quadrupoles at once. Using MINUIT to perform such a fit for the FFTB would be prohibitively expensive in computer time, but would produce a uniform solution in which all FFTB quads are fitted to a common line.

A modification of the above solution is illustrated in Figure 4.7. In this situation, the last quad in the group of stepped magnets is defined to be aligned to the desired trajectory, a "phantom" xcor and ycor are inserted into the beginning of the beamline, and their strengths are included as fitvars. These phantom elements are computational artifacts used to extract the approximate angle of the magnets relative to the incoming beam: since the last quadrupole is defined to have zero misalignment relative to the incoming beam, but the first quadrupole is producing a kick angle θ, the only consistent solution is that the phantom corrector must be producing a kick angle $-\theta$, which is being cancelled by the quadrupole kick. In this case, the
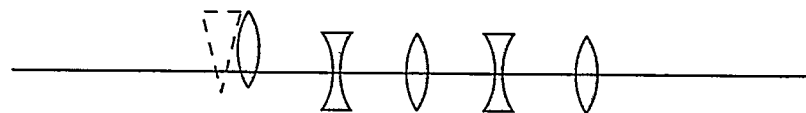
Figure 4.7: Use of "phantom" correctors in FFTB alignment. The kick imparted by the corrector is cancelled by the sum of kicks from misaligned elements, such that the last quad in the line is aligned by definition.

appropriate correction of magnet positions can be naively implemented, but only if a real corrector exists which can be set to cancel the effects of the "phantom" one (or in real terms, the steering from the first quad must be transferred to an appropriate xcor or ycor).

The requirement that each segment begin with a corrector which can be adjusted in this fashion is a significant aspect of the selection of quadrupole segments. A further consequence of not using a downstream endpoint in the extraction line is the danger of "walking the magnets into the wall:" specifically, that the incoming ray may be so poorly chosen that the magnets are driven monotonically away from the center positions of their movers. This can quickly result in magnets moved to their limits. As this problem is related to the *a priori* mechanical alignment of the beamline, and the aperture limitations in the extreme upstream end of the FFTB, the situation is considered in detail below.

4.2.3 Mechanical Alignment and the FFTB Launch Criterion

Prior to the September 1994 FFTB run, the magnets from QA1 to the dump were surveyed using mechanical techniques including laser-trackers and water levels, and placed with an expected accuracy of 100 μm in the horizontal, and 50 μm in the vertical [48]. During the May 1994 FFTB run, the incoming beam was steered such that all obstacles in the BSY, including PC90, were avoided (based on evidence from
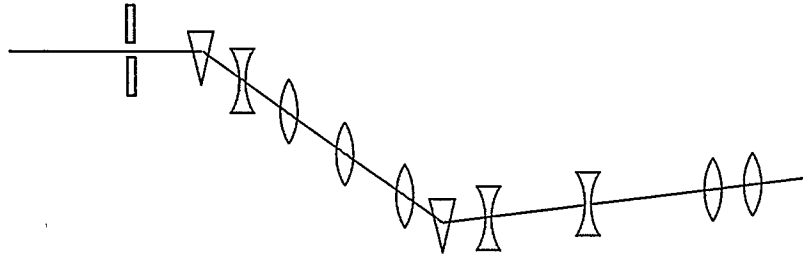
Figure 4.8: Quad Alignment "dog-leg" schematic. The beam is kinked by two correctors to steer from the line through the BSY and PC90 onto the line of the FFTB proper, and intervening quads are moved onto the angled line in between.

the loss monitors in the area). Because of the extreme radioactivity of objects in the BSY, and the fact that the area is uninhabitable during any SLAC operations, the obstacles were poorly understood and not amenable to mechanical correction. The readings on BPMs 30 and 50 (free-standing BSY BPMs) which skirted the obstacles and cleared PC90 adequately became known as the FFTB "Launch Criterion". Thus, the beam position and angle could not readily be altered upstream of XCOR and YCOR 1010 (H1D and V1D, correctors located between PC90 and QSM1).

Because of the radiation concerns, as well as the 16 meter steel wall, the first three quads could not be moved onto the FFTB line. Even if they could have been, there was no guarantee that this line would coincide with the line dictated by the launch criterion.

The problem was solved by "dog-legging" the beam in both x and y as shown in Figure 4.8. XCOR 1010 and XCOR 1070 (between B01A and B01B) in the horizontal and YCOR 1010 and YCOR 1040 (downstream of QA1) in the vertical were used to match the orbit of the incoming beam to the line of the FFTB. This was accomplished by first turning off quadrupoles Q5 through QA2, inserting ST62, and using the BPMs upstream of 1090 to perform a rough match. The beta match quads were then moved roughly to zero their BPMs in both planes, and trimmed to their nominal values for quad alignment. The orbit was corrected, ST62 removed, and further fine adjustments

8–95   7973A2
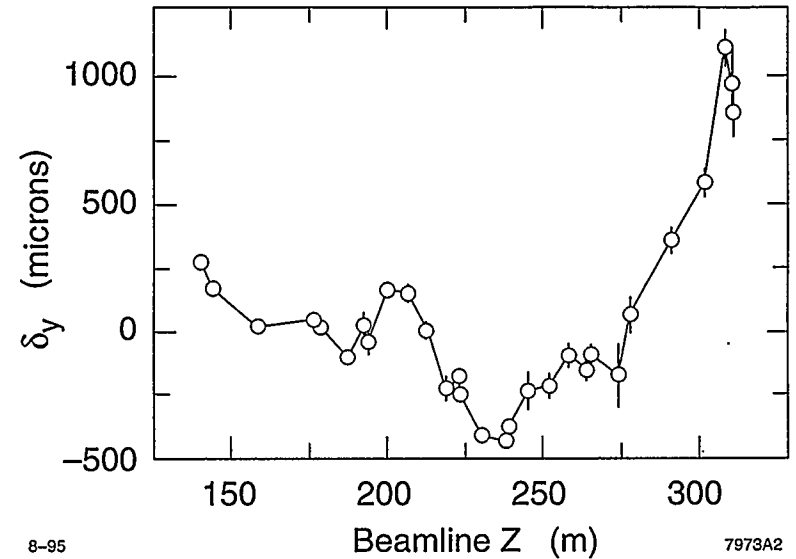
Figure 4.9: Vertical positions of FFTB quads relative to incoming beam in May 1994.

brought the beam into the FFTB and onto the line required by mechanical alignment. In this way the risk of "walking into a wall" was minimized.

4.2.4 Results of Quadrupole Alignment

The first attempt to align the FFTB from Q5 to QC1 occurred in May of 1994. During this operation, an unexpected rise and fall was encountered in the vicinity of the CCSX. This situation was handled by using the vertical corrector magnets, and also by using the quadrupole movers to produce steering in a few locations. The pre-alignment positions of the magnets are shown in Figure 4.9. In this case, the beamline was piecewise-aligned only. The monotonically-deteriorating resolution in Figure 4.9 was due to two phenomena: first, the snowballing of global errors within a string of quadrupole magnets (see discussion "Defining Quad Alignment 'Resolu-
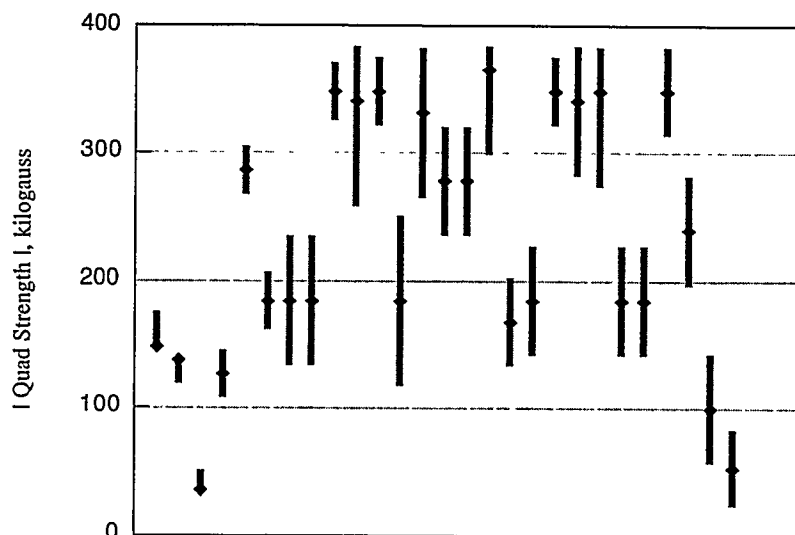
Figure 4.10: Scan ranges determined by simulation for FFTB standard quads. For simplicity of display, absolute values are shown. Diamonds represent nominal values of quads.

tion'," below); second, by monotonically-increasing values of χ^2/ν for the successive alignment segments. The value of χ^2/ν increased from a low of 0.3 (indicating that the BPM resolution, hard-coded in OPTFIT at 6 μm, was for this operation closer to 3 μm) to a high of over 20. OPTFIT reports errors normalized by $\sqrt{\chi^2/\nu}$, leading to significant blow-up of the downstream error-bars. Despite these difficulties, the background-tuning situation for the beam size monitors and wire scanners improved dramatically after this pass of alignment, and the quad emit scans were able to go to a larger range because Q5 was no longer steering as severely.

During the period from May to September of 1994, several improvements were made. The mechanical alignment of the FFTB was checked, and found to agree qualitatively with the beam-based measurements; the alignment was then corrected to the levels noted above. The optics for beam-based alignment were adjusted as
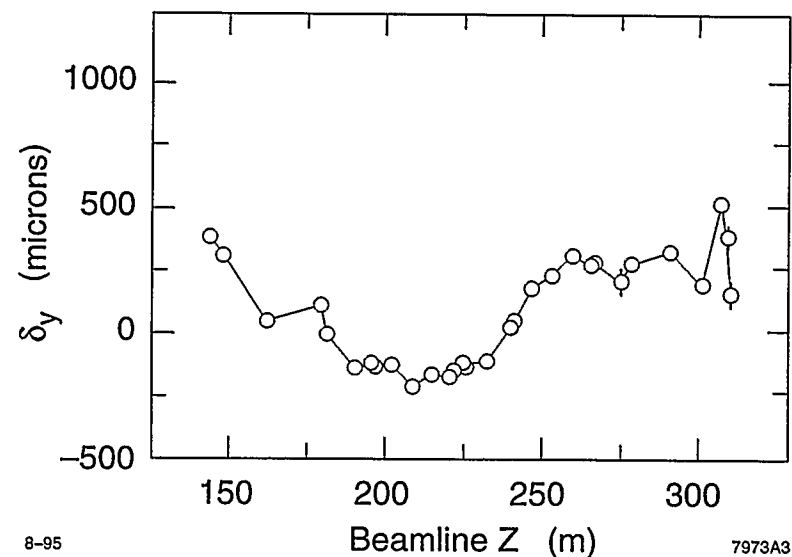
Figure 4.11: Vertical positions of FFTB quads relative to incoming beam in September of 1994.

noted in Section 4.1, and optimal scan ranges for each quad were determined in the new optics. Figure 4.10 shows the quadrupole scan ranges selected. Finally, the scheme for resolving the Launch criterion was devised.

Figure 4.11 shows the results of beam-based alignment in the vertical plane from September of 1994, including all the improvements made during the period between runs. Note that the error-bars are much smaller, and the overall amplitude of the misalignments reduced by a factor of 2 over May. In addition, the vertical misalignments are, as advertised, within about 50 μm of one another, save for a mysterious "step" in the beta exchanger of 400 μm. Furthermore, the alignment of September 1994 resulted in the FFTB operating with all vertical correctors from YCOR 1040 to the FP at zero strength for the first time.

Figure 4.12 shows the horizontal and vertical resolutions of the quad alignment procedure, as well as the χ^2/ν for each segment. Note that the beta match segment
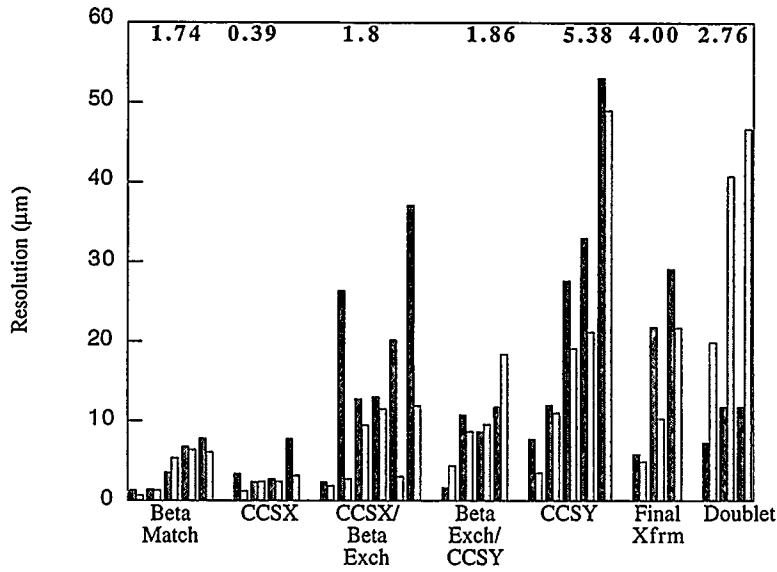
Figure 4.12: Resolutions for FFTB quad alignment experiment of September 1994. Horizontal (dark) and vertical (light) are shown. Numbers in boldface at top are χ^2/ν for each segment's fit.

Figure 4.13: Contribution to χ^2 for all FFTB BPMs, considering only pulses in which all quads were at their nominal values. Contributions are normalized by number of good data points from each BPM. The data from the upstream half of the CCSX (dark) and from the 3 quads upstream of the doublet (light) are used in this analysis, showing degradation due to loss of hysteresis.

has a somewhat larger χ^2/ν than the subsequent segment, and that once again there is a general monotonic increase in the value of χ^2/ν.

The former phenomenon is easily understood: during a repair access to the BSY, BPM 30 was damaged. The connections for BPM 30 were then swapped to BPM 20, a linac- style BPM just upstream of BPM 30. The relatively poor resolution of BPM 20 causes the incoming beam to be relatively weakly determined, since only BPM 20 and BPM 50 lie upstream of Q5. For all subsequent segments, BPM 20 is deselected and only the BPMs from BPM 50 to QA2 are used for resolution of the incoming beam.

One possible cause of the degradation of the fit quality is loss of standardization as quad alignment progresses. As Figure 4.10 shows, many of the quads are stepped through a significant portion of their total range. This may cause the upstream

transport matrices to diverge from the expected model, and this would cause fits of the incoming beam to fail. One way to examine this possibility is to use two sets of quad align data, throw away all data except when all the quads are at their nominal values, and see whether the incoming beam fits can describe the BPM data through the line, as indicated by BPM χ^2 contribution values.

Figure 4.13 shows the χ^2 contribution per data point for each BPM in the FFTB, for the CCSX1 segment (QM3A through QN2A) and the FTRANS segment (QC5 through QC3). Note that the same BPMs are used to measure the incoming beam in the two cases, and that all data with quads stepped off their nominal strengths are deleted. In effect, the two datasets should describe identical beamlines, yet the later
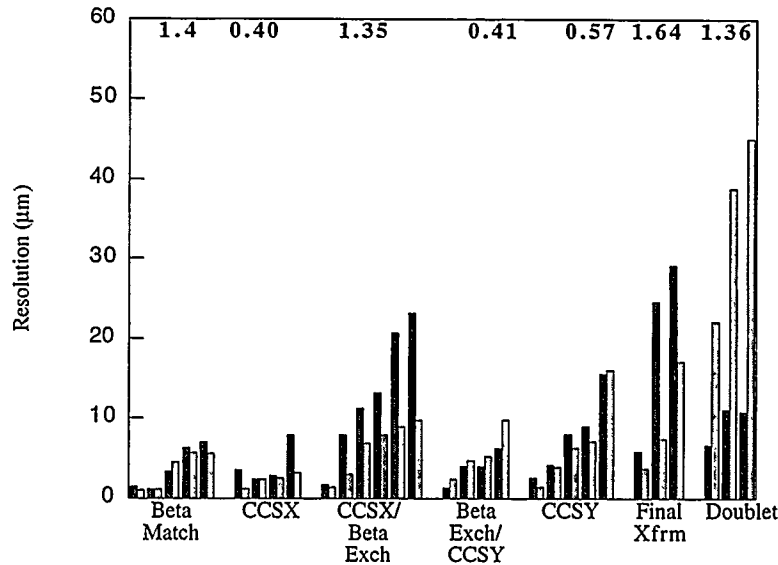
Figure 4.14: Resolutions for FFTB Quad Alignment experiment of March 1995. "Mini-Standardize" protocol was used for all quads except the doublet. Horizontal (dark) and vertical (light) are shown, along with χ^2/ν for each segment's fit (boldface).

set has relatively poor fit convergence. This is significant evidence that the loss of standardization is damaging the algorithm.

In order to correct this loss, all FFTB quads were configured for "mini-standardize" before the March 1995 run. In this mode of operation, all increases in absolute quad strength (i.e., increases in operating current) were accomplished in the normal fashion; all decreases in absolute strength were accomplished by first overshooting the new set point by 5%, then rising in current to the set point. In this way, all FFTB quadrupoles were set to new operating points from lower current to higher current. The resolutions and χ^2/ν of the quad alignment fits from March 1995 are shown in Figure 4.14. Note that χ^2/ν now achieves a high of 1.6 for the three quadrupoles preceding the doublet, and is nearly as large (1.4) for the doublet quadrupoles. Deeper

investigation indicates that the χ^2 contributions which are large in these cases are universally the horizontal, not the vertical. One possible suspect, therefore, is that the extraction line quads are not well aligned, and the values through which the last 6 quads are stepped result in particles scattering off the quad apertures onto the BPMs in the horizontal. In addition, the doublet quads were not set to mini-standardize, as the large excursion this induced changed the optics enough to cause radiation trips.

Defining Quad Alignment "Resolution"

The resolutions shown in Figures 4.12 and 4.14 are the square roots of the diagonal elements of the error matrix calculated by MINUIT, multiplied by $\sqrt{\chi^2/\nu}$ for the entire fit. Some consideration is needed, however, to fully comprehend what this number represents, or, in other words, what it is we are aligning to with such precision.

The alignment of a single quadrupole is unambiguous, as shown in Figure 4.2. The quadrupole is stepped, the beam deflection measured, and the correlation between the two can be converted into a distance between the incoming beam centroid and the magnetic center of the quadrupole. The error on this quantity may be interpreted with similar ease.

Figure 4.15 shows a situation of slightly greater complexity. Here two quads are shunted in sequence, in an attempt to align both to a common line. Note, however, that the beam is passing through the center of the second quadrupole due to the kick received from the first. In this case, shunting the second quad will not deflect the beam, even though the second quad does not lie on the line of the incoming beam. In order to determine the offset of the second quadrupole, it is first necessary to know the offset of the first quadrupole, which then determines the kick given to the beam by the first quad.

Let us define the offset of the beam from the center of the first quad by $x_{i,1}$ and that of the second quad from the beam by $x_{i,2}$; $x_{i,j}$ is simply the offset revealed by a naive correlation between quad shunt and downstream kick. Let us now define the
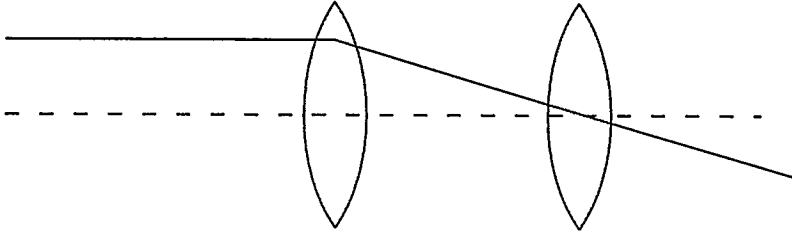
Figure 4.15: "Intrinsic" vs. "Global" resolution. Because the second quad will not deflect the beam when shunted, its intrinsic misalignment is zero; however, the first and second quad are both clearly off the line of the incoming beam, while the first quad deflects the incoming beam through the center of the second.

global offset of the quad, $x_{g,j}$, to be the true deviation of the quad centers from the incoming beam (or, the distance between the quad centers and an incoming ray of infinite energy, which is not deflected by the quads). In the thin- lens approximation, we may say that:

$$x_{g,j} = x_{i,j} + \sum_{k=1}^{j-1} -K_{q,k} x_{i,k} L_{k \to j}, \tag{4.7}$$

where $L_{k \to j}$ is the distance along the path of the beam from the kth quadrupole to the jth. This is simply a superposition of steering kicks, where the steering strength is proportional to the offset of the quad from the beam. The summation term in Equation 4.7 can be thought of as the deviation of the beam incident on the jth quadrupole from the incoming beam's vector. If the measurement of the *intrinsic* resolution, $x_{i,j}$, is not correlated between quadrupoles, then we expect that their uncertainties will add in quadrature:

$$\sigma_{g,j}^2 = \sigma_{i,j}^2 + \sum_{k=1}^{j-1} K_{q,k}^2 L_{k \to j}^2 \sigma_{i,k}^2. \tag{4.8}$$

Because OPTFIT's χ^2 engine uses tracking code rather than a closed-form solution, the results which emerge are explicitly global misalignments and resolutions. Equation 4.8 is responsible for the monotonic loss of resolution along a given segment. Figure 4.16 shows the intrinsic resolution of the technique for each quadrupole.
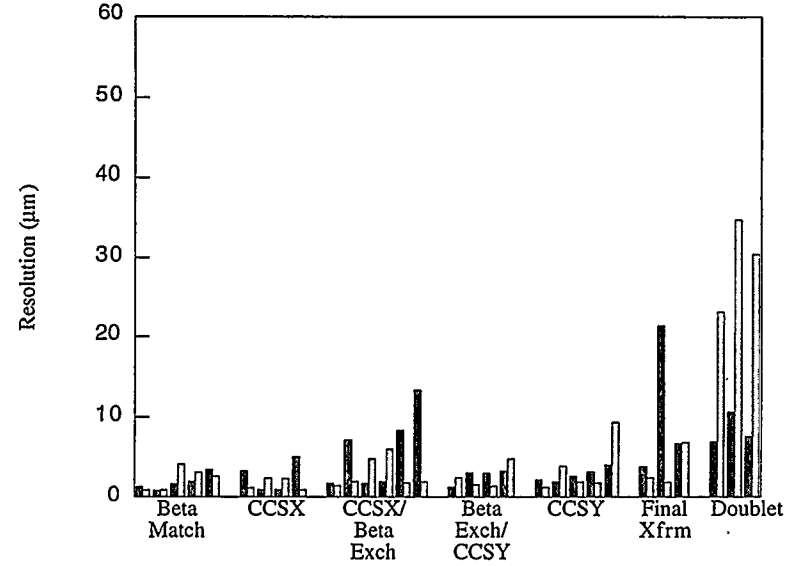
Figure 4.16: Intrinsic resolutions for all quadrupoles, using the OPT-FIT technique and the data set from March 1995.

One implication of this global alignment formalism is that the beam emerging from the last quadrupole in a segment has a deviation from the incoming given by $\sum_{k=1}^{j} -K_{q,k} x_{i,k} L_{k \to j}$, with a corresponding error $\sum_{k=1}^{j} K_{q,k}^2 L_{k \to j}^2 \sigma_{i,k}^2$. This is the error on the beam entering the first quadrupole of the next segment. It is possible, therefore, to use the intrinsic resolutions of the procedure, from Figure 4.16 above, to compute the global resolutions of the entire beamline as though all alignment was performed in a single fit to a single incoming beam. This represents the true RMS deviation of the quadrupoles from a geometrically-perfect straight line (modulo steering magnets and bends). Figure 4.17 shows the true global errors.

The global errors in Figure 4.17 were determined by first-order tracking: each quad was displaced from the beam incident on its upstream face; the displacement of the quad was determined by a gaussian-distribution random number generator,
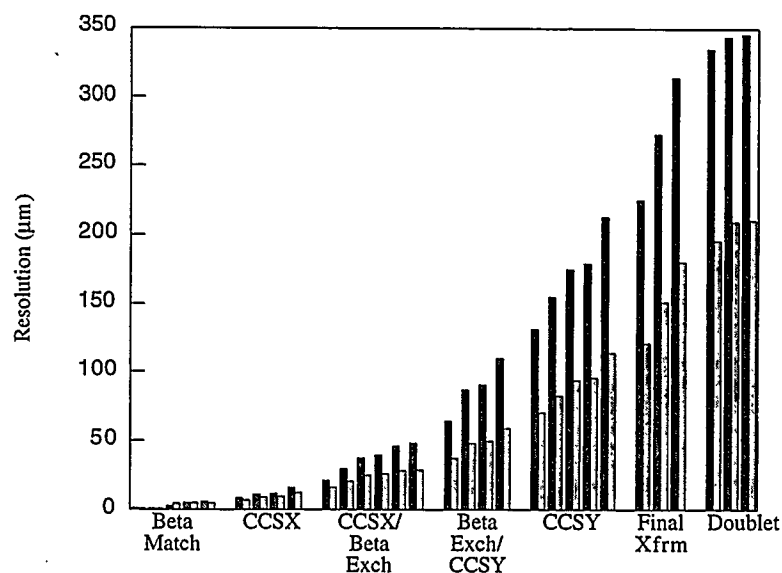
Figure 4.17: True global alignment resolution of beam-based technique described in the text. Resolutions were determined by 10,000 repetitions of first- order tracking, and using the previously-shown intrinsic resolutions.

and the RMS of the distribution determined by the intrinsic resolutions in Figure 4.16. The RMS displacement from the origin of the coordinate system after 10,000 iterations was used to generate Figure 4.17.

Note that, while Figure 4.17 represents the deviation from a straight line, Figure 4.16 represents the RMS distance between the quad centers and the beam in the quads. As has been noted elsewhere [47], beam-based alignment techniques are not sensitive to misalignment patterns with a wavelength which is long relative to the betatron wavelength, and this is graphically demonstrated by the differences in Figures 4.16 and 4.17. In the case of linac alignment, this typically results in "bowing", in which the global misalignment is largest in the center of the beamline (due to the endpoints being fixed) [46], [47]. Because the FFTB fixes only one endpoint, the

end of the beamline is "flapping in the breeze" relative to the beginning. For optical purposes, the intrinsic alignment is the crucial one, in that dispersion and higher-order aberrations are generated by the beam passing through the magnet off-center. The global resolution is important because sufficiently poor resolution can result in a solution to beam-based alignment which drives magnets outside of the range of their movers.

Global Movement Knobs

Because the FFTB line is defined by one endpoint and not two, it is conceivable that a sufficiently poor selection of initial steering and/or combination of global errors will result in a beamline which is systematically shifted or tilted, in such a way that movers at the endpoints are near their limits of travel. It is useful to be able to shift the beamline systematically to relieve these movers, or to adjust steering into the extraction line.

Because every optical element is mounted on a mover, such knobs can easily be designed by simple geometrical considerations. The knobs must properly adjust the beta match quads and the correctors used to deal with the Launch Criterion.

The FFTB has such knobs, and they were used experimentally during the FFTB run in March 1995. Results were generally good. However, steering errors crept into the system during their use. These errors are thought to be due to imperfect understanding of the launch correctors: while the beamline motions are correct, the correctors do not steer by precisely the right amount. This problem can be corrected with steering knobs which use only the launch correctors to steer position and angle at some given BPM in the FFTB. Such knobs will be in place for the next FFTB run.

4.2.5 Systematic Errors in Quad Alignment

In addition to the "snowballing" discussed in the previous section, there are several systematic errors which may appear in the measurement or correction procedures of

quad alignment.

The most straightforward systematics are scaling factors in quad strengths and/or BPMs. In Figures 4.12 and 4.14, it was argued that implementation of "Mini-Standardize" had improved the convergence of the fits; however, this argument only holds for the quads at their nominal values, not at the scanned values. A quad which has overstepped or understepped by some fraction will result in an over- or under-estimate of its misalignment. Similarly, a global error in the scale factors of the BPMs will lead to a systematic scaling of measured misalignments. Examining Equation 4.7, we see that such errors, which will scale all intrinsic misalignments uniformly, will also scale all global misalignments uniformly.

Lattice diagnostics measurements found that the FFTB BPMs systematically over-report motions by 12.4% (see Section 4.4.3). This leads immediately to a 12% over-correction of the misalignments. Because the quadrupole power supplies ramp very slowly, and a settle time is used in the control software, quadrupole scale-factors are not expected to be significant.

Measurement of horizontal misalignments is further complicated by dependence on the extracted energy offset of each pulse. A systematic miscalculation of the energy will result in BPM readings tracked around bend magnets being improperly fitted. This may also account for the greater difficulties inherent in fitting the horizontal orbits, as shown in Figure 4.13. The RMS energy jitter of the beam is approximately 1.6×10^{-4}; the maximum dispersion in the CCSX is 9.85 centimeters. Section 4.3.2 argues that the energy measurement is systematically high by roughly 6%; this argues that the error contribution at the high-η BPMs will be 1 μm RMS. As long as the jitter is roughly symmetric about the reference energy, however, the contribution from this term will primarily enlarge the χ^2, rather than systematically pulling the fit to one side. This error, though arising from a systematic source, actually is a cause of random variations.

The pulse-to-pulse jitter of the incoming beam results in possible systematic errors

in the alignment between segments. Figure 4.3 shows the RMS jitter at each BPM in the FFTB. At any given BPM, therefore, the deviation between one reference orbit averaging N pulses, and another, will be given by $\frac{\sigma_j}{\sqrt{N}}$, where σ_j is the RMS jitter amplitude. This contribution should formally be added in quadrature with the values shown in Figure 4.17; however, for 100 pulses averaging, the resultant contribution is quite small, and omitted here.

A final systematic error in the correction scheme arises from the mechanical design of the final doublet. Section 2.4.1 mentions that the three magnets of the doublet are fixed to a common vibration-isolated table. The table can be moved in six degrees of freedom: x, y, z, roll, pitch, yaw. This is sufficient to independently move the end quads (QC2 and QC1), assuming that the three quads are internally aligned with arbitrary precision. However, QX1 is found to be 233 μm to the south of the QC2-QC1 line, with an uncertainty on the order of 10 μm. Fortunately, the vertical alignment of QX1 to the line of the other two magnets is closer to 24 μm with an uncertainty of 40 μm. In the past, the doublet has been aligned by moving QC2 and QC1 as indicated by their misalignments; the discovery of the misalignment of QX1 implies that this strategy needs to be re-thought.

4.3 Lattice Dispersion Measurement

While beam-based quadrupole alignment is a test of specific errors within the FFTB line, the measurement of lattice dispersion (R_{16} and R_{36} from the beginning of the line to each BPM) is in many ways a more general test of the health and well-being of the optics. While the horizontal and vertical dispersion measurements can reveal errors in, respectively, the strength and roll angles of the CCS bend magnets (which was the original purpose), the measurement can also reveal errors in the alignment and roll of the quadrupole elements. These effects can be somewhat difficult to decouple, which causes the measurement to be less specific in its focus than the quad alignment test.

4.3.1 Measurement Technique and Algorithm

The general approach to dispersion measurement in the FFTB is similar to that used to measure the dispersion entering the SLC Final Focus from the collider arcs [49]: the beam energy from the linac is increased or decreased, and this produces an orbit shift at each dispersive BPM. The correlation between the change in orbit and the change in energy is the dispersion:

$$\eta_{x,y} = \frac{d(x,y)}{d\delta}. \tag{4.9}$$

The beam energy in the FFTB is maintained by a feedback which uses the BPMs in the vertical-dispersion region of the extraction line, described in Section 3.1.3. The beam energy is changed by changing the setpoint of the feedback. The change is virtually instantaneous: after a change in setpoint, typically only one pulse passes through the accelerator at the original energy before the feedback "catches up."

As with quad alignment, the incoming trajectory and energy must be extracted on every pulse, and the resolution of their reconstruction incorporated into the main fit. Since the computational elements needed for this operation already exist in OPTFIT, this program was chosen as the basis for the dispersion measurement as well. In the current implementation, the incoming ray fits are not correlated to the energy fits, and therefore the incoming dispersion is not determined. Even if the two were to be correlated, the launch feedback ensures that energy-dependent changes in the incoming orbit are corrected, thus eliminating the information needed for incoming dispersion measurement.

Original Algorithm

The FFTB lattice contains 14 CCS bends (including the soft bend) and 4 weak vertical correctors for correction of rolled installation of the bends: at the B01A/B location, between B01C/D and B01E/F locations (adjacent to QN1); between the B02A/B and B02C/D locations (adjacent to QM2); and at the B02E/F location. As noted in

Section 2.4.2, the bend magnets were fabricated and installed in pairs. The dispersive effects of the magnets of a pair cannot, in general, be decoupled by the BPMs in the FFTB lattice; in addition, the effect of interest is the total dispersion due to a pair, rather than the dispersion of each magnet. Also, the soft bend is too close to B02E/F to be measured separately. OPTFIT is therefore configured to fit the strengths of the four vertical correctors, and six horizontal bends (B01A, B01C, B01E, B02A, B02C, B02E), while the remaining 8 bends are constrained in the code to be at their design strengths. Note that the correctors are being used as "phantom" correctors: the locations are used as a convenient point-source of dispersion, regardless of the actual strength of the correctors (typically zero).

OPTFIT was originally written to fit a finite set of beamline optical properties using BPM and/or wire scanner data. These effects included bend magnet and corrector magnet strength, but no provision existed for directly measuring a dispersive ray by correlating position and energy data at each BPM. Instead the data was used to fit the strength of the bend magnets and vertical correctors, and from these the resultant R_{16} and R_{36} matrix elements could be computed from the beginning of the beamline. Thus the dispersive rays available to OPTFIT depended upon the determination of source terms; an unexpected source of dispersion (such as a skew quad term in the beta exchanger, for example) could not be easily fitted by this algorithm. This defect was recognized early in the development of the code, and a "work-around" devised. The algorithm was arranged as follows:

- Step energy feedback through several values, at each value acquire 10 pulses on all BPMs in FFTB

- Extract δ and σ_δ for each pulse

- Extract x, x', y, y', with errror matrix, for each pulse using BPMs 30 and 50 only

- Perform data cleanup as described in Section 4.2.2 for BPMs between bend magnets

- Using design values of bend magnets and vertical correctors, propagate error matrix onto all BPMs, add propagated error in quadrature to BPM resolution

- Perform MINUIT fit of bend strength and corrector strength

- Insert fitted strengths into beamline and compute R_{16} and R_{36} from BPM 30 location to each BPM; this constitutes the *fitted* dispersion rays

- At each BPM, correlate the fit residual of each data point with the fitted energy of the pulse

- Add the correlation above to the fitted dispersion to obtain the *measured* dispersion, and to the design dispersion to obtain the *after correction* dispersion.

Once the magnet strengths have been determined, a correction can be computed. The bend angle of the fitted bend, θ_{fit}, is added to the bend angle of its fixed partner, θ_{design}, to obtain the total bend angle of the two magnets, θ_{pair}. The difference between θ_{pair} and $2\theta_{design}$ is the amount of the correction, $\Delta\theta$. This can be accomplished via the trim windings on the pair of magnets, but it is necessary to determine how much to change the trim winding. The scheme for this correction is as follows:

- Read the bend magnet main winding and the trim winding for the pair to be corrected

- Using the hysteresis polynomial for each bend, determine the theoretical total bend angle for the pair, θ_{poly}

- Compute the trim winding setting needed to set each of the pair of magnets to $\theta_{poly} + \Delta\theta/2$

- Average the settings obtained for each of the pair of bends, this is the setting for the correction.

The correction scheme depends heavily upon the hysteresis polynomials for the bend magnets, and also upon the approximate linearity of the correction in the vicinity of the setpoint. For corrections of several percent, however, even poor agreement of the polynomial with the real magnet is adequate to converge, particularly since the dispersion measurement and fitting process takes approximately 60 seconds per iteration. Furthermore, if the bend polynomial is inaccurate by more than a few percent overall, the beam will not go through the FFTB in the first place without major trial-and-error tuning of the bend trims and main supply. Since such has not been necessary in the FFTB, the inescapable conclusion is that the process should converge.

Improvements to the Algorithm

The algorithm as constituted above is seriously flawed in several particulars. First, the computation of the measured dispersive ray is inelegant to say the least; furthermore, the dependence on the fitted bend strengths makes computation of error-bars on the measured dispersion values quite a bit more difficult than is strictly necessary; worst of all, use of the design optics to propagate the error matrix from the incoming ray means that any significant deviation of the dispersive rays from the design will introduce errors in the propagation of energy uncertainty. At a BPM with dispersion η_{BPM}, the error contribution from uncertainty in energy measurement is:

$$\sigma_{energy} = \eta_{BPM}\sigma_\delta. \tag{4.10}$$

Because the design value of η_{BPM} is zero in the vertical plane at all BPMs, this term does not contribute to vertical error at any BPM, and pulses with badly-determined incoming energy and pulses with well-determined incoming energy will be weighted equally in determining the corrector strengths. This is clearly unacceptable. A more advanced algorithm was developed which addresses these weaknesses:

- Proceed with data acquisition, energy fitting, launch fitting and data cleanup as above

- Propagate the errors from computation of x, x', y, y' only to each BPM on each pulse

- Perform tracking of each pulse, using only geometric components of launch vector (i.e., set energy offset to zero and perform tracking)

- At each BPM on each pulse, compute an η contribution in each plane by dividing the tracking residual by the δ of the pulse; compute a weight for this contribution from the BPM error, the 4-D launch error, and the energy error

- The measured dispersion at the BPM is the weighted average of the η contributions

- Using the measured η, propagate the energy fitting error onto each BPM for each pulse

- Perform MINUIT fit as above, compute R_{16} and R_{36} to obtain fitted η rays

- Measured η - Fitted η + Design η = After Correction η.

Note that a correlation exists between the fit of the dispersive rays and the fit of the magnet strengths: the data at each BPM is used to determine the dispersive ray, in correlation with the pulse-to-pulse energy variation; the dispersive ray is then used to adjust the weight of the data points for use in the MINUIT fit of magnet strengths. In order to avoid this correlation, half the dataset is used only to compute the dispersive ray and half is used only for the fitting of magnet strengths.

4.3.2 Results of Dispersion Measurements

Dispersion measurements in the FFTB are more subtle than quadrupole alignment measurements in that the horizontal plane contains nonzero design dispersion, and as a result the horizontal measurement is not a "nulling" measurement. The vertical plane measurement is a nulling measurement, due to zero design dispersion. Furthermore, any understanding of the horizontal measurement requires an understanding of the energy measurement process as well. Each of these facets will be considered separately.

Energy Measurements

The pulse-to-pulse energy variation is measured via a set of BPMs surrounding a string of permanent bend magnets in the extraction line. The value of η_y at the last BPM is 60 centimeters. A resolution of 6 μm would naively imply an energy resolution σ_δ of 1×10^{-5}. The resolution is diluted by the fact that the pulse-to-pulse y and y' variation must be computed as well, and by the fact that the last BPM also has the maximum resolution of the incoming angle. When these issues are factored in, the expected resolution is 6×10^{-5} for a set of BPMs with 6 μm resolution.

Figure 4.18 shows the reconstructed energy for each pulse during the March 18, 1995 alignment of the first half of the CCSX (QM3A to QN2A, 120 pulses total). The vertical axis shows normalized resolution ($6 \times 10^{-5} \cdot \sqrt{\chi^2/\nu}$) for each pulse. The jitter "envelope" is ± 300 ppm, with an unweighted RMS of 160 ppm. While the normalized
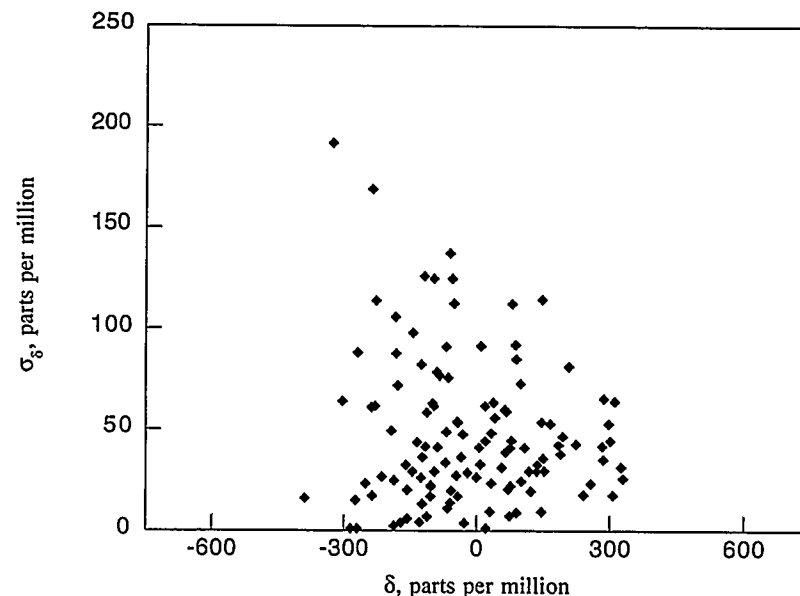
Figure 4.18: Measurement of pulse-to-pulse energy jitter made during a sequence of quadrupole alignment on March 18, 1995. Resolutions have been normalized in the fashion described in the text.

fit resolution is concentrated about 30 ppm (implying an actual BPM resolution of 3 μm in the extraction line), the resolution has a long tail going out to over 200 ppm. The cause of this behavior is not known; however, the extraction line is not as optically forgiving as the area upstream of the FP. In particular, the apertures are tighter, the quadrupoles less well-behaved, and the alignment is poor compared to the upstream components. The poor resolution is likely linked to BPM hardware issues, such as intermittent spray impacting the devices.

Figure 4.19 shows the reconstructed energy and normalized resolution from a dispersion-measurement dataset taken on March 18, 1995. In this case the energy is stepped out to ± 5000 ppm in order to measure the lattice dispersion. Note that the energy error is significantly worse for both the large excursions and the small ones
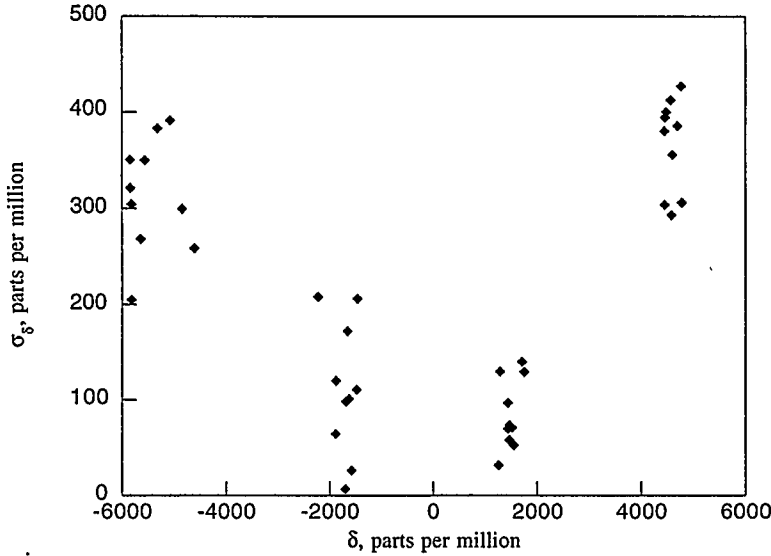
Figure 4.19: Energy values from dispersion measurement made on March 18, 1995.

(± 1800 ppm), with a wide spread of resolutions in all cases. An interesting side note is that the fractional resolution, $\frac{\sigma_\delta}{\delta}$, is centered on 6% for all four clusters. The large energy excursions were made in order to measure the dispersion with higher precision, but the larger errors on the energy fits for these pulses eliminates this improvement. For future measurements, smaller excursions of the energy are more attractive: the resultant distribution of resolutions is closer to the distribution at zero offset, and the resultant fractional resolution is as good as that for large excursions.

The normalized resolution for the energy measurement is used in all subsequent computations, including computation of the measured dispersion and propagation of launch errors onto BPMs. Normalizing the resolution of a fit in this fashion is generally an unattractive solution to unknown problems in data acquisition, fitting, etc., essentially "sweeping under the rug" everything which is unknown about the

system. Nonetheless it was kept in the fit in this case for one reason: the BPMs in the extraction line, which are used to perform the fit, are systematically different from the BPMs in the rest of the FFTB, and consequently have a resolution which is different from that in the main FFTB line but is itself unknown. This leads to an interesting systematic error in fits such as quad alignment. The total error at a BPM location is given by:

$$\begin{aligned} \sigma_{eff}^2 &= \sigma_{BPM}^2 + \sigma_{4D}^2 + \sigma_{\eta\delta}^2 \\ &= \sigma_{BPM}^2 + \sigma_{4D}^2 + \eta^2\sigma_\delta^2. \end{aligned} \qquad (4.11)$$

The value of σ_{BPM} is a function of the BPM hardware itself; the value of σ_{4D} is determined by the fit of the incoming ray and the optics from 50B1 to the BPM in question, but it is in turn a linear function of the upstream BPMs used to fit the incoming ray. Thus, a scale-factor discrepancy between the expected and actual resolutions of the standard BPM will result in a common scale factor to these two terms, which will simply result in a χ^2/ν for the fit which differs from 1 by the square of this selfsame scale factor. The term $\eta\sigma_\delta$, however, has already had any resolution scale factor removed by the χ^2 scaling described above. The most straightforward correction to this difficulty is to re- analyze quadrupole alignment data in the vertical plane (in which no energy dependence exists) to determine the best value for the standard BPM resolution, and set this value in the software.

An additional systematic difficulty is the issue of scaling-factors. Dispersion measurement requires that the BPMs measuring the energy in the extraction line and the orbit deviations in the FFTB line have the same scaling-factor: a factor of 1 would be ideal, but any value will do so long as it is the same for both systems. The scale factor for the standard BPMs was measured with a stretched wire and a pulser, and found to be 94% of the predicted value. This calibration was included in the database for each BPM upstream of the FP, and for the unit in QP1A. No such calibration was performed for the extraction-line devices, and the database values for these are equal to design. The analysis of the energy in the dumpline therefore assumes a BPM scale
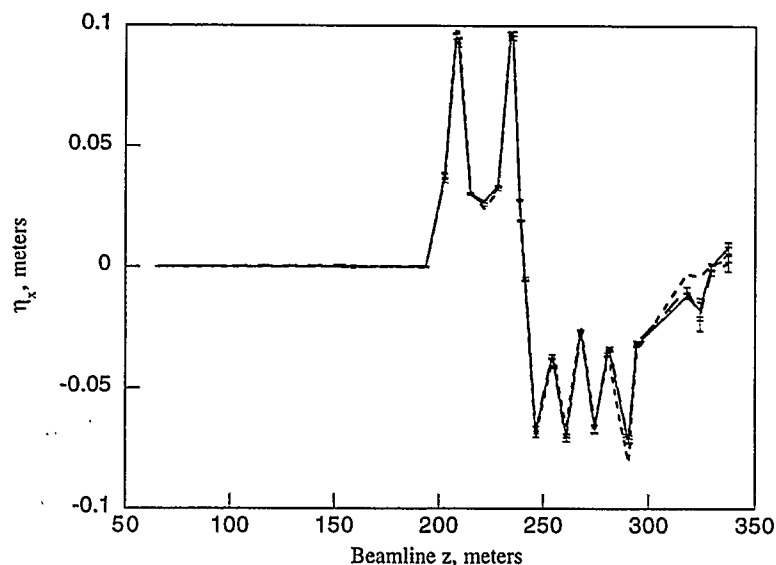
Figure 4.20: Horizontal dispersion measurement on September 3, 1994 prior to quadrupole alignment. Three datasets (solid) and the design value (dashes) of η_x are superimposed. The design and measured can only be seen separately at the downstream end of the beamline on this scale.

factor of 94% for the extraction-line BPMs.

Horizontal Dispersion

Figure 4.20 shows the measured horizontal η function on September 3, 1994, immediately prior to beam-based quadrupole alignment. The measurement was performed three times, and the figure shows the three measurements and the design dispersive ray. While the figure resembles the design dispersion well enough to declare that no gross error exists, no further conclusions can be made because the large value of the design dispersion dominates the figure. A more useful figure is Figure 4.21, in which is plotted $\eta_{design} - \eta_{meas}$. Because the third set of data from September 3 did not converge well, only datasets 1 and 2 are shown here. Note that convergence between
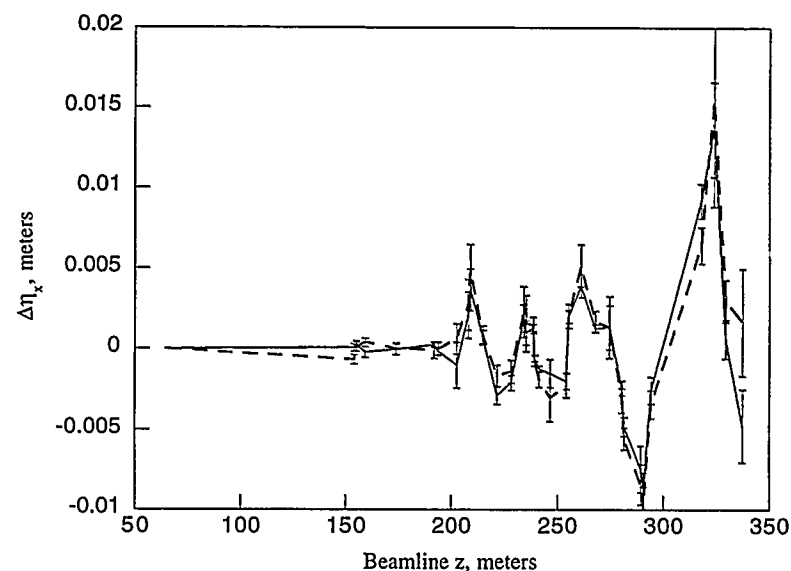
Figure 4.21: Horizontal dispersion mismatch, $\eta_{design} - \eta_{meas}$, before beam based alignment on September 3, 1994.

the two is generally good (error-bars at a given point overlap for almost all BPMs). Agreement in the last 4 BPMs (QC3, QC2, QP1A, and the first extraction-line BPM) are not as good. While most areas fall within 5 mm of their design, an excursion of over 1 cm appears at QC2, the first quadrupole of the doublet.

Two areas of interest in understanding Figure 4.21 are the process by which error-bars are computed, and the initial setup procedure for the FFTB's bend magnets and alignment. Each of these is discussed below.

At a given BPM, the contribution to the dispersive ray from a given pulse is given by:

$$\eta_{contrib} \equiv \frac{x_{BPM} - x_{4D}}{\delta}, \qquad (4.12)$$

where x_{4D} is the result of 4-D propagation of the incoming beam to the BPM. The uncertainty in $\eta_{contrib}$ will be the sum in quadrature of two terms, one of which carries

the BPM uncertainty, and one of which carries the energy uncertainty:

$$\sigma_\eta^2 = \left(\frac{\sigma_{eff}}{\delta}\right)^2 + \left(\frac{(x_{BPM} - x_{4D})\sigma_\delta}{\delta^2}\right)^2 \qquad (4.13)$$
$$= \left(\frac{\sigma_{eff}}{\delta}\right)^2 + \left(\eta_{contrib}\frac{\sigma_\delta}{\delta}\right)^2 .$$

The value of $\frac{\sigma_\delta}{\delta}$ was shown earlier to be approximately 6% for all energy pulses, and thus each pulse contains an error-term which is 6% of $\eta_{contrib}$. At the high-dispersion points this term will tend to dominate; with 20 pulses used to compute η_{meas}, the resolution of the dispersion at these points should go as 6%/$\sqrt{19}$, or 1.4%. The error-bars in Figure 4.21 are normalized by the $\sqrt{\chi^2/\nu}$ value for the given BPM dispersion fit, and these are typically 1.4% of the dispersion at the highest points (adjacent to SF1A and SF1B).

Prior to the September 1994 run-time, the quadrupoles in FFTB were re-surveyed and moved to new positions determined by the survey. The CCS bend magnet main and trim supplies were set to values determined by their excitation curves, in order to get each magnet as close as possible to its design strength of 5.624 kilogauss-meters. It is interesting to note that, despite these efforts, a centimeter-sized mismatch of the CCS dispersion appears in the Final Transformer.

During beam-based alignment, the quadrupole alignment is assumed to follow the design line of the FFTB at the half-millimeter scale. Whenever the alignment fit shows an angle between the beam and the existing quad line, the bend trims are adjusted to remove this angle, which in theory also will assist in fixing the dispersive ray. Figure 4.22 shows the post-alignment deviation of the measured dispersion from the design for two datasets. Note that the overall scale of the figure still spans 2.5 cm, and that the two datasets agree quite poorly. Error-bars are also quite large.

What has happened here? Recall that these measurements were made after quadrupole alignment, and that evidence exists (see Section 4.2.4) that the quad alignment measurement changes the optics (via loss of standardization), sufficiently that the incoming ray can no longer be tracked through the FFTB. A further change
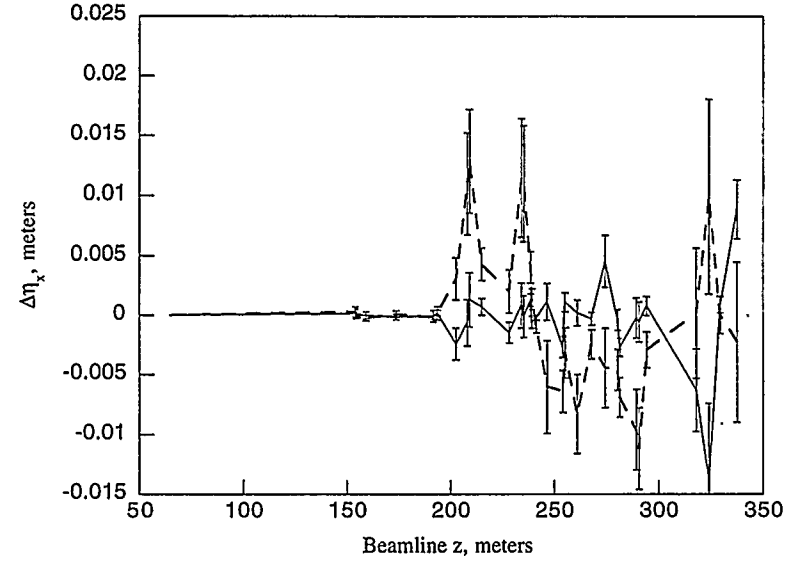
Figure 4.22: Horizontal dispersion mismatch after quad alignment, September 1994.

in the system is that the beam positions through the extraction line are allowed to change in an unconstrained manner during alignment, within the limits of beam-containment. The extraction orbit seen before and after quad alignment changes in the vertical by over 1 mm; worse, the orbit after alignment has larger values in all 4 BPMs used for energy measurement. This means that both BPM nonlinearities and possible optical effects in the B06 magnets (quadrupole and/or sextupole moments) are worse for this trajectory than the former.

Based on the September 1994 experience, several improvements were made to the setup of the beamline and the quad alignment procedure for March 1995. The quad positions and bend strengths determined during September 1994 were used as a starting point; however, the dumpline orbit was steered flat in the horizontal and vertical (all BPM absolute values under 1 mm), and re-steered to this orbit after every
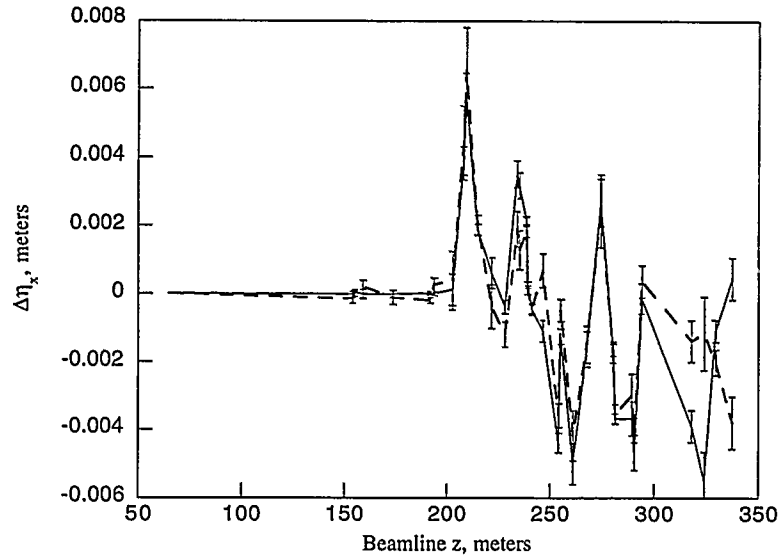
Figure 4.23: Horizontal dispersion mismatch prior to beam-based alignment on March 18, 1995.

segment of alignment. As mentioned in Section 4.2.4, the quadrupoles were set to "mini-standardize." Figure 4.23 shows the result of this preparation. The overall scale of the dispersion mismatch is half that encountered before alignment in September. Moreover, the mismatch ray through the CCSX and CCSY is qualitiatively similar to the dispersive ray itself, leading to suspicion that the "mismatch" may be dominated by scaling-factor errors between the FFTB standard BPM and the extraction line units.

Figure 4.24 shows the horizontal dispersion mismatch after beam-based alignment in March 1995. The superficial similarity between the mismatch and the design is even more pronounced, and it is highly likely that some of this is due to a scaling-factor between the two sets of BPMs. Some of the fine structure may be due to small errors in the individual BPM scale-factors. However, it is important not to be too
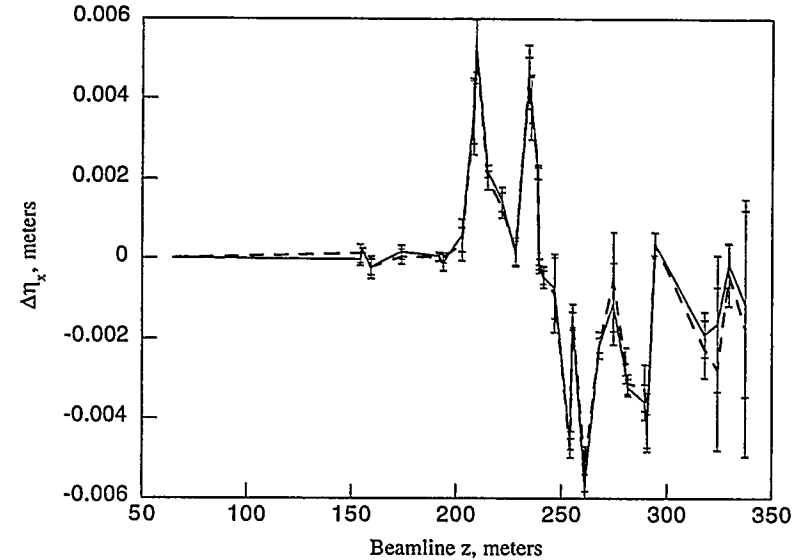
Figure 4.24: Horizontal dispersion mismatch after beam-based alignment on March 18, 1995.

easily convinced by the gross appearance of Figure 4.24. The BPM providing the first spike in the CCSY, for example, is the BPM in QT4, a quad next to SD1A. This BPM should agree with the next BPM, at the QM3B location, which makes up the first "anti-spike" in the CCSY pattern. The second "spike" in the CCSY is at the location of the BPM which, in Figure 4.24, is near zero (z = 275 meters). The BPM which is supposed to be at the location of the first "spike," the QT3 location, is near zero on this figure (upstream of QT4).

The generally poor matching of the dispersion within each CCS region can be explained by the quad alignment procedure. The bend trim windings are set by forcing the beam onto a line through the downstream quadrupoles; this will generally work if the quads downstream of the bend being set have sufficiently long lever arms to constrain the angle. The quads within the CCSX and CCSY, however, do not have

such long lever arms. Consequently, the bend angle of bends B01C/D and B02A/B are not well constrained by this process, and the dispersion within each CCS may be poorly matched. The QT3/QT4 discpepancies remain unexplained.

Note further that, while SF1A and SF1B are at equal values of η_x, within the error of the measurement, SD1A and SD1B appear to have 0.5 mm difference in their η_x values. Such a mismatch could cause generation of second-order horizontal dispersion (T_{166}); however, the magnitude of this aberration generated by such a small mismatch in η_x is approximately 30% of the FFTB's inherent T_{166}, which is a relatively small contribution.

In conclusion, while progress has been made in using the horizontal dispersion measurement, the domination of systematic errors and poor resolution of the energy still prevents this measurement from being of great value in beamline tuning.

Vertical Dispersion

The design value of the vertical dispersion function is identically zero at all points upstream of the extraction line vertical bends; the vertical dispersion measurement is therefore a nulling test, as mentioned above, and relatively insensitive to scaling effects between the energy BPMs and the main FFTB line BPMs. The vertical plane is also more sensitive to all aberrations. This combination makes the vertical dispersion measurement both more robust and more relevant to the overall tuning scheme.

Figure 4.25 shows the vertical dispersion function measured prior to beam-based alignment on September 3, 1994. The dispersion function is clearly nonzero, with a peak of +3 mm at QN3B in the CCSY, and an oscillation of ±8 mm between the CCSY sextupoles, which continues into the final transformer. The figure shows two measurements superimposed, and the agreement between the two is good.

Could such an oscillation arise solely from quadrupole misalignments? Consider the 8 millimeter dispersion at SD1A ($z \approx 250$ meters in Figure 4.25). This point is $\frac{\pi}{2}$
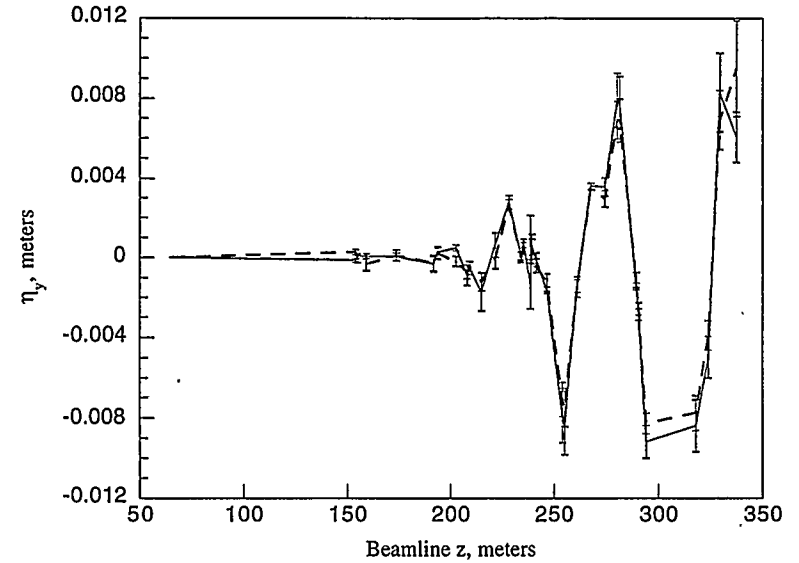
Figure 4.25: Vertical dispersion function measured prior to beam-based quadrupole alignment on September 3, 1994. Note that the function is manifestly nonzero, particularly in the vicinity of the SD1A/B sextupoles.

out of phase with quadrupole QN1, at the center of the CCSX, with an R_{12} between the two of -24.4 meters. A vertical dispersion of 8 mm at SD1A could be caused by a vertical η_y' of 328 μm at QN1. This in turn can be caused by a misalignment of QN1 of 1.5 mm in the vertical. While this is large compared to the misalignments found for individual quads, recall from Figure 4.11 that the peak-to-peak vertical misalignments of FFTB were over 500 μm.

Figure 4.26 shows, on the same scale as Figure 4.25, the post-alignment vertical dispersion from September 1994. The oscillation through the SD1A/B locations has been reduced by a factor of 2 over the pre-alignment values, and the final transformer oscillation is also reduced, albeit by a smaller factor. Interestingly, the dispersion at QM2, in the center of the CCSY, and QN2B are also reduced by a smaller factor
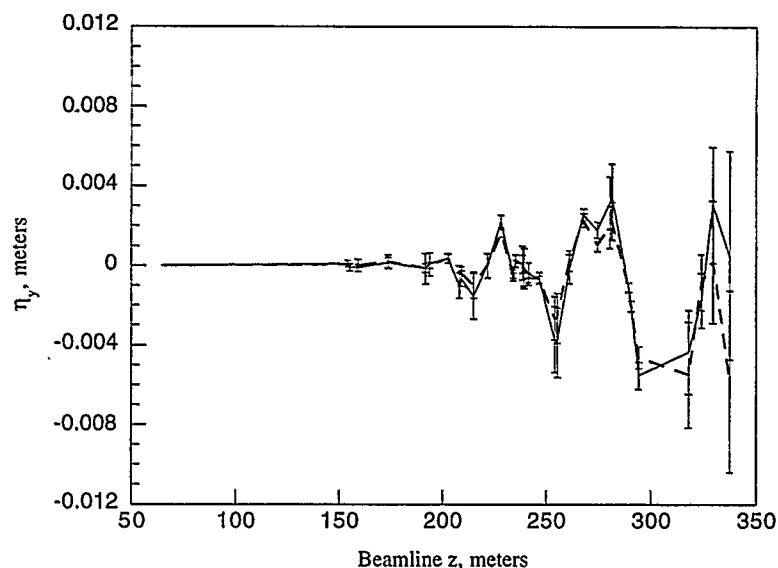
Figure 4.26: Vertical dispersion function measured after beam-based quadrupole alignment on September 3, 1994.

than the sextupole oscillation. As in other analyses of the September 1994 local correction data, the convergence of the fits, and the correspondence between fits, degrades monotonically from upstream to downstream due to magnet excitations upstream leaving their hysteresis curves.

Figure 4.27 shows the pre-alignment vertical dispersion measured on March 18, 1995; the scale in the top half is identical to the scale for Figures 4.25 and 4.26, for comparison, while the scale in the bottom half is expanded. While the overall dispersion pattern remains the same, the amplitudes are reduced still further from Figure 4.26. This may be from the beneficial effects of recovering standardization on all FFTB magnets. Figure 4.28 shows the post- alignment dispersion function, and also a common-axes comparison between the pre- and post-alignment vertical dispersion rays. The majority of the figure is unchanged: η_y reaches peaks of ± 2.5

mm at SD1A and B, and a peak of 3 mm at QM2. However, the peaks at QN2B and QC5 are now larger than before alignment.

What conclusions can be reached from the transformations of the vertical dispersive ray from September 1994 through March 1995? For one thing, the vertical dispersion is measurably reduced by beam-based alignment, assuming that the initial alignment was sufficiently problematic. Second, the FFTB contains sources of vertical dispersion which are not improved overmuch by beam-based alignment, which may even be worsened by the procedure. One such source of vertical dispersion is rolled bend magnets or quads in the dispersive region.

In order to determine possible rolled magnets and corrective strategies, a least-squares fit to the measured dispersion of Figure 4.28 was performed. This fit forms a curvature matrix from the linear beam transport properties, and inverts to obtain a solution. In the fit, BPMs were considered firmly attached to their quadrupoles, such that a quad rolled through an angle θ contained a BPM rolled through the same angle. Such BPMs give rise to spurious dispersion, since the vertical striplines now sample the horizontal dispersion.

Magnet	Case "A"	Case "B"	Case "C"	Case "D"
B01B	0.0	-9.2 ± 2.7	0.0	-15 ± 1.3
QM3A	-4.1 ± 1.1	0.0	-7.0 ± 0.6	0.0
QN3B	0.0	0.0	-3.8 ± 0.9	-2.7 ± 0.9
B01C	11.3 ± 2.8	10.0 ± 3.4	0.0	0.0
QN3C	8.9 ± 1.2	8.5 ± 1.2	9.0 ± 1.1	8.7 ± 1.1
QT4	-10.5 ± 2.7	-10.1 ± 2.7	-10.8 ± 2.7	-10.1 ± 2.7
QM3C	-8.4 ± 3.1	-8.9 ± 3.1	-7.6 ± 3.1	-8.5 ± 3.1
QM1C	-7.4 ± 2.2	-7.7 ± 2.2	-7.0 ± 2.2	-7.6 ± 2.2
χ^2/ν	10.8 / 17	12.5 / 17	14.6 / 17	15.5 / 17

Table 4.1: Results of fitting magnet rolls to vertical dispersion data, for four different cases. Roll angles are reported in milliradians.

Table 4.1 shows the fitted roll angles in four cases. Note that the solutions contain several degeneracies: the fit cannot easily distinguish the effects of quads which are close together in betatron phase (such as the quads on opposite sides of CCS
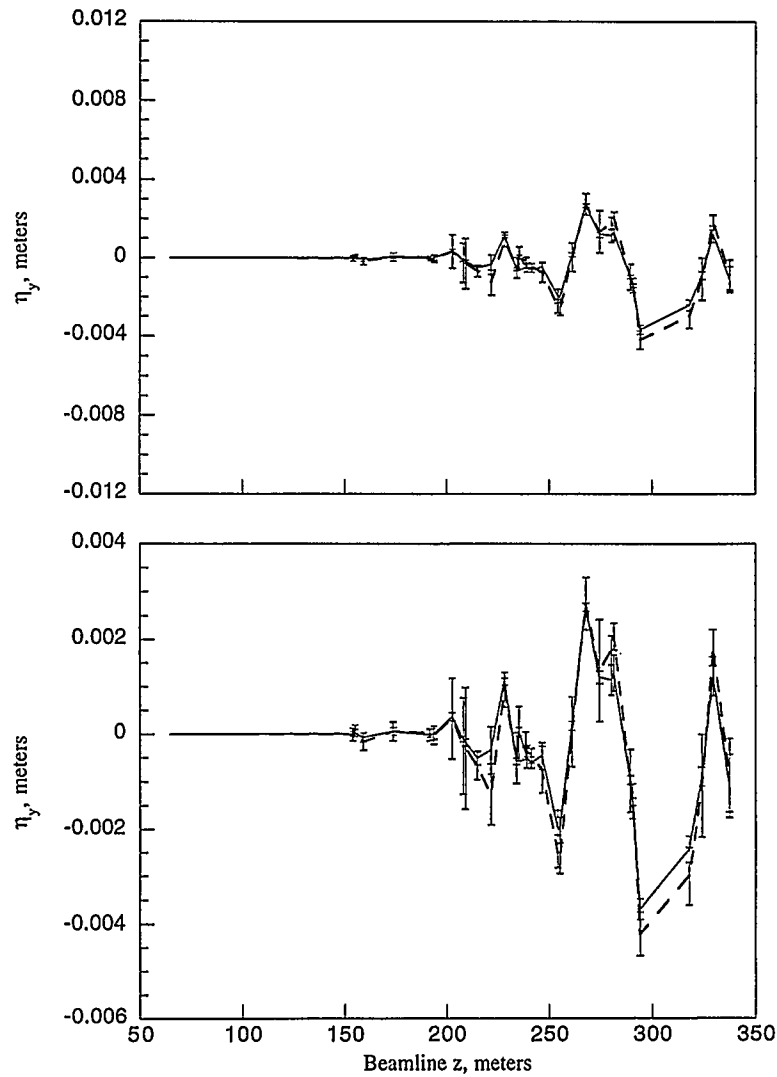
Figure 4.27: Vertical Dispersion measured on March 18, 1995, prior to beam-based alignment. Dispersion is displayed with same scale as September 1994 data (top), and expanded scale (bottom).
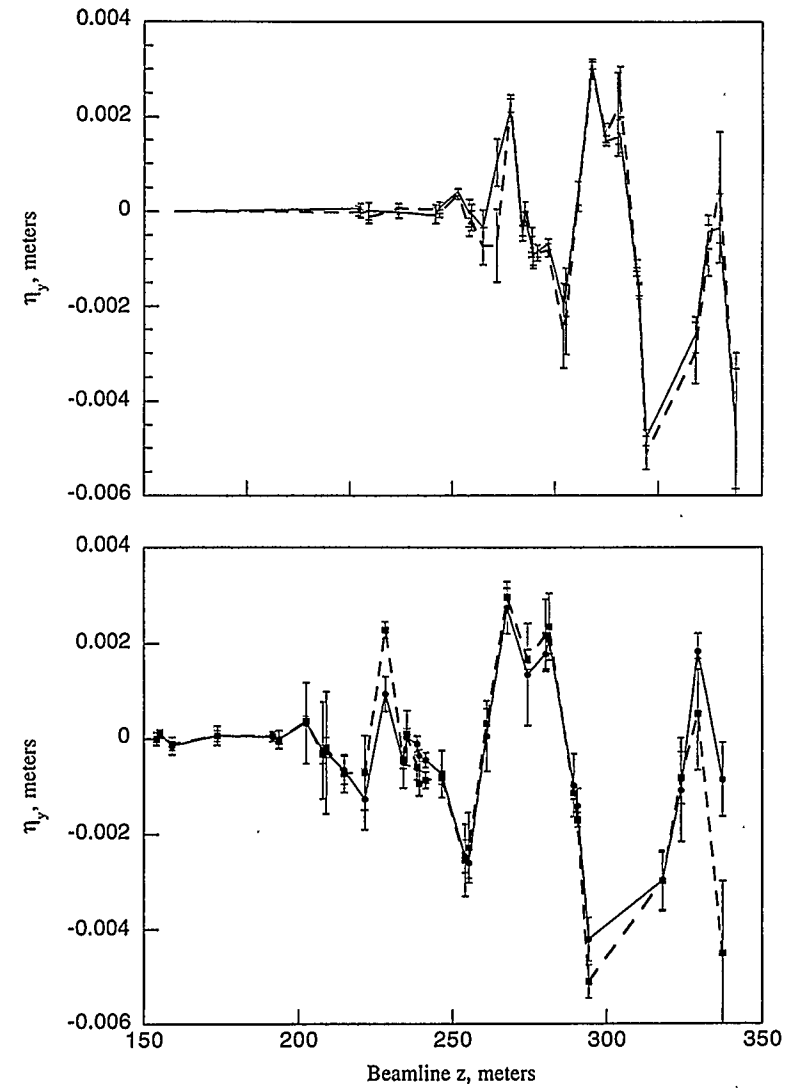
Figure 4.28: Vertical dispersion ray after beam-based alignment on March 18, 1995 (top); both before (solid) and after (dashed) on common scale.

sextupoles, or QT2A/QT2B), and so the total roll of both magnets in such a pair is represented by a single quad in Table 4.1 (i.e., the "QN3B" roll angle of -3.8 mrad in case "C" could actually be -1.9 mrad for QN3A and -1.9 mrad for QN3B). The same is true for neighboring bend magnets. Note also that a rolled B01B is hardly distinguishable from a rolled QM3A, and similarly QN3B and B01C are hard to distinguish.

While the four cases in Table 4.1 result in dispersive rays which are nearly indistinguishable, Case "C" is of the greatest interest because it arises solely from rolled quadrupoles. Because the FFTB magnets are on movers with a roll-angle degree of freedom, it is conceivable that the rolls in Case "C" could be at least partially corrected during the next FFTB run. In addition, the quad rolls in any of the cases in Table 4.1 would give rise to the 1° roll measured at WS3 (see Chapter 3). Figure 4.29 shows the measured dispersion, and the expected dispersion from the rolls in Case "C" including the effects of rolled BPMs.

The presence of uncorrected η_y in the FFTB lattice has several implications. First, the ray used in Figure 4.29 has a value of 140 μm at the FP, which would enlarge the spot size through dispersion. As we shall see in Chapter 5, the total range of the global η_y^* knobs is more than adequate to eliminate this problem. Second, the presence of vertical dispersion at the optical elements creates the possibility of higher-order chromatic aberrations which limit the achievable beam size at full energy spread. This possibility will be addressed in Chapter 6, on chromatic studies and bandwidth of the FFTB. Third, the vertical dispersion at the SD1 sextupoles will cause the sextupole mover knob for the FP (see Chapter 5) to also change the FP dispersion. We can measure this effect by noting that the FP value of η_y' is 1.3 mrad. Therefore, a 1 mm waist shift will cause 1.3 μm of η_y^*. Since the design value of β_Y^* is 100 μm, a 1 mm move of the waist is quite large, and therefore the change in dispersion of 1.3 μm is the largest which is likely to occur due to this effect.
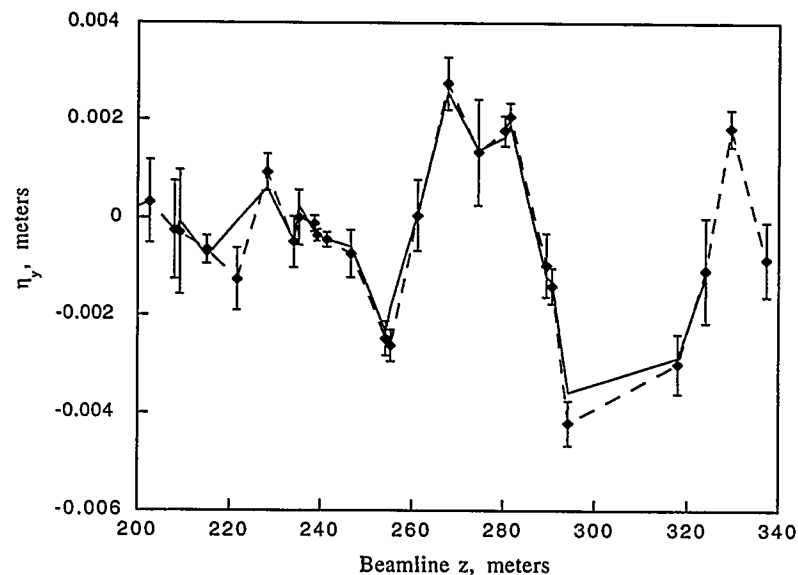
Figure 4.29: Measured dispersion ray (points and dashed line) and Case "C" fit (solid), described in Table 4.1.

4.4 Lattice Diagnostics

The process of lattice diagnostics is in many respects the opposite of quadrupole alignment: while the former changes the strength of a quadrupole in order to determine its offset from the beam, lattice diagnostics changes the trajectory of the electron beam through a region in order to determine the transport properties from one point to another. From this information, the strengths of individual quads can be inferred.

The traditional lattice-diagnostics algorithm, as performed in linacs and storage rings, is as shown in Figure 4.30[50]. A vertical or horizontal corrector is stepped through a range of values, and the displacements at a downstream BPM are measured at each value of the corrector. By correlating the position with the corrector, the R_{12} from the corrector to the BPM can be determined. This is typically done using a
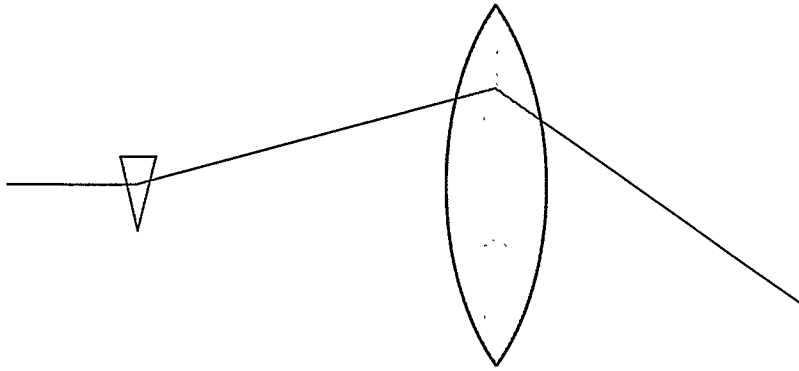
Figure 4.30: Schematic of conventional lattice diagnostics for a FODO array. The beam is offset in the magnet, and the kick thus imparted is measured downstream.

BPM and corrector separated by 90° of betatron phase, such that the angle kick at the corrector is converted into a position kick at the BPM.

Such a scheme is optimal for a matched FODO lattice, such as a storage ring, in which the set of correctors in a ring span the full range of betatron phases. Due to the unusual phase-advance properties of the FFTB, it is generally impossible to locate a corrector exactly 90° degrees away in betatron phase from a given magnet or BPM. Therefore, an alternative scheme had to be devised.

4.4.1 Alternative Lattice Diagnostics Algorithms

Figure 4.31 shows the first approach considered for FFTB lattice measurements. Two correctors are employed in such a way as to produce a pure position or angle oscillation at the location of the CCS sextupoles. In this way, an oscillation equivalent to a single corrector at the correct phase is generated. By producing both position bumps and angle bumps at the sextupole, both betatron phases are probed. The "zeroth-order" bump test is merely to ensure that the angle bump closes, and that the position bump is of equal amplitude at both ends of the $-I$ transform. Modified versions of this
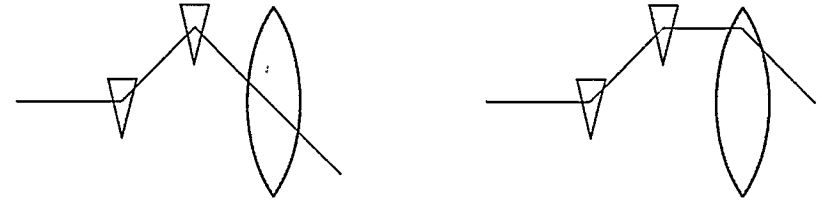
Figure 4.31: Schematic of proposed FFTB lattice diagnostics. Two correctors are used to produce oscillations in each phase, including pure angle bumps (left) and position bumps (right).

test can be employed to test a magnification module; and the deviations from perfect behavior can be used to measure individual quads in the CCS regions.

The bump method described above has several drawbacks, which eliminated it from serious consideration early on. First, creating a pure position or angle bump from two correctors requires precise knowledge of the corrector strengths as a function of current. Tests of the FFTB corrector magnets with beam indicated significant scale factors (in some cases, over 10%[51]). Second, the tests of quad strengths and $-I$ transport properties relied upon the linearity and scale factors of all BPMs in the system. Like the horizontal dispersion measurement described in Section 4.3.2, this is not a nulling test. In order to use the corrector method, both the corrector and the BPM must have scale factors and linearity known at least as well as the desired resolution of the quad strength, and the ratio of oscillation size to BPM resolution must be of the same order. In order to diagnose quadrupole errors at 0.1%, therefore, the 1 μm BPMs must be linear to within 0.1% out to 1 mm offsets, and have scale factors known to 0.1%.

A method which eliminates some, but not all, of these difficulties is shown in Figure 4.32. In this case, the magnet mover is used to create the bump, and the quad strength is determined by the ratio of the amplitude of the kick downstream to the amplitude of the motion. The magnet mover positions are externally monitored by the LVDT system, and the mover system as a whole is calibrated to μm precision and
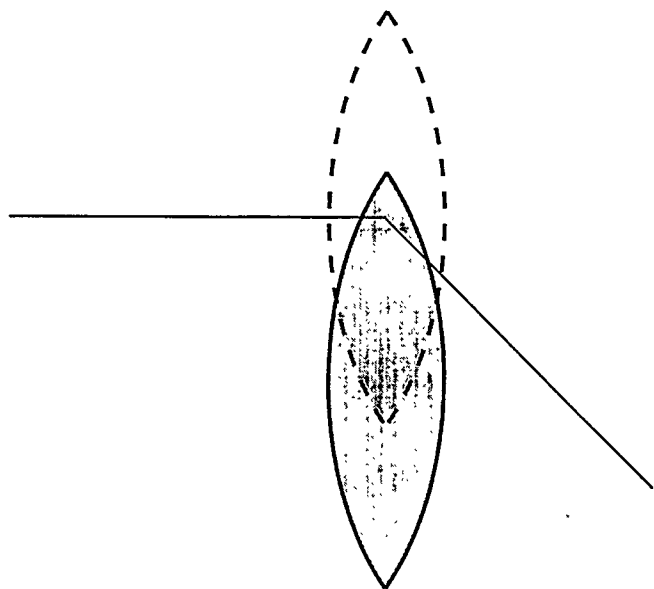
Figure 4.32: Schematic of magnet-mover based diagnostic. The quad to be measured is moved transverse to the beam, and the kick produced by the change in position is measured downstream.

0.1% accuracy over its full range of motions [31]. In this case, both the quadrupole being moved and any quads between it and the target BPM are being measured. However, this test still requires high precision and accuracy from all BPMs.

The technique which ultimately was used in FFTB is shown in Figure 4.33. In this algorithm, several quadrupoles are moved at one time, in a pattern which will produce a closed orbit bump – the BPMs downstream of the last quad will see no motion if the magnets are at their design strengths. This test is a nulling test, in that no change on the BPMs is expected if the quads are at their design strengths. The offsets are small compared to the linear range of the system. In addition, the scale-factor of the BPMs in use in the system no longer couples to the absolute strength of the quads, but to the strength errors of the quads. In other words, a 10% scale
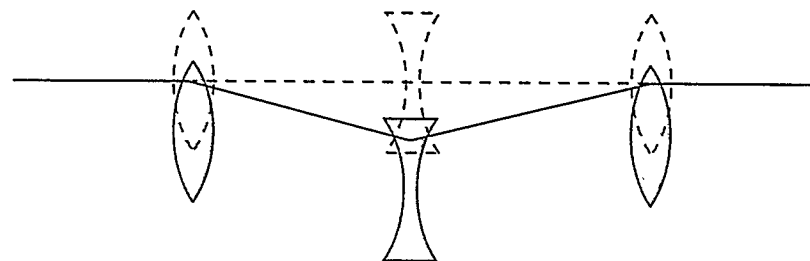
Figure 4.33: Schematic of lattice diagnostics method used in FFTB. Several quads are moved via their movers simultaneously, in patterns which will produced a closed oscillation if all quads are at design strengths.

factor error will not result in a 10% miscalculation of the quad strength, but a 10% miscalculation in the strength error (deviation between design and actual strength). In this respect the test is similar in its result characteristics to the vertical dispersion test described in Section 4.3.2.

4.4.2 Detailed Description of the Algorithm[52]

The lattice diagnostics algorithm has been developed in such a way that data acquisition is performed by the SLC Control Program, primarily using correlation plots (although the magnet configuration and online model facilities are also used), while analysis is performed offline. This is due to the large size of the data set which is fit: because a closed bump generally requires 3 or 4 quadrupoles moving synchronously, the deviation of the downstream trajectory is a function of several different quadrupole strengths. In order to separate the effects of the different quads, each quadrupole is included in several different bumps, and the full set of 100 bumps is used to fit 30 quadrupole strengths. In addition, each bump is repeatedly cycled across its full range, and many BPM samples are acquired at each step. Fitting such a large dataset with so many parameters is de facto an offline task.

Despite this difference, the lattice diagnostics algorithm bears many similarities

to the quad alignment and lattice dispersion algorithms described above, both in acquisition and analysis. The acquisition algorithm is as follows:

- Acquire reference orbit with no bumps present

- Select next bump to introduce, maximum amplitude, number of steps, and number of iterations of stepping to use

- For n = 1 to number of iterations of stepping do
  - for m = 1 to number of steps for this bump do
    * Set the bump to the required amplitude
    * Acquire 50-100 pulses of BPM data for all FFTB BPMs
  - End do (steps for this iteration of this bump)

- End do (iterations of this bump)

- Select next bump and appropriate iteration parameters, repeat above until all bumps are exercised.

Typically each bump would be set to 5-10 different amplitudes, the full pattern of amplitudes would be performed 5-10 times, and 50-100 pulses would be acquired at each step of each bump. The analysis algorithm is as follows:

- For n=1 to number of bumps do begin
  - Identify all pulses with bump amplitude=0, select first pulse to be "temporary reference orbit"
  - Subtract "temporary reference orbit" from other amplitude=0 orbits, use BPM data upstream of bump and at high η point (if necessary) to fit x, x', y, y', δ for each of these pulses
  - use theoretial (no-error) optics to subtract jitter from each amplitude=0 pulse; if residual for a pulse is large, discard the pulse; average residuals at each BPM and add to "temporary reference orbit" to form final reference orbit
  - Subtract final ref orbit from all others, fit incoming oscillation as described above; elminate contribution of incoming oscillation at all BPMs; at each bump setting compute average and RMS of each BPM, eliminate fliers
  - Fit correlation of BPM values computed above vs. bump amplitude to straight line, using BPM RMS as weighting factor; eliminate BPM values which do not fit straight line

- end do

- Slope and error of straight line fits, along with χ^2 of straight line fits, become the input data for MINUIT fit of magnet strengths.
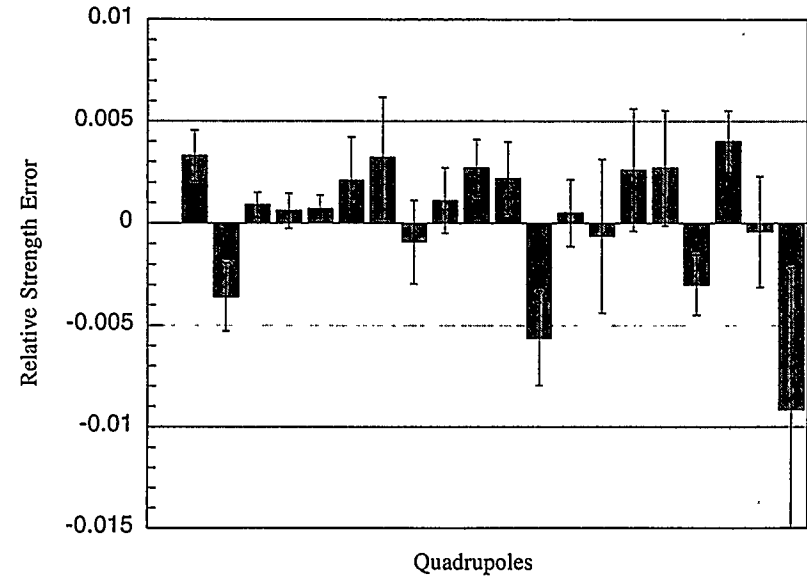
Figure 4.34: Relative strength errors, $(K_{model} - K_{meas})/K_{model}$, for FFTB quads from QN3A in CCSX to QC3 in FT. Result for QT2A is not shown.

4.4.3 Results of Lattice Diagnostics

The earliest set of data taken with quadrupole movers utilized single-magnet open oscillations. This data determined that the BPM system is operating with an overall scale factor of 13% (i.e., the BPMs are over-reporting deviations by 13%). This factor, and variations in same between BPM units, was a prime mover in the decision to use closed bumps for lattice diagnostics. The BPM scale factor is still needed to reduce the χ^2 contribution coming from BPMs inside quads which are moved in a given bump.

Early measurements using small sets of closed-bump oscillations showed persistent difficulty in reducing the χ^2 to reasonable values. It was shown that the difficulty was a mismatch between the actual energy of the incoming beam and the energy of the
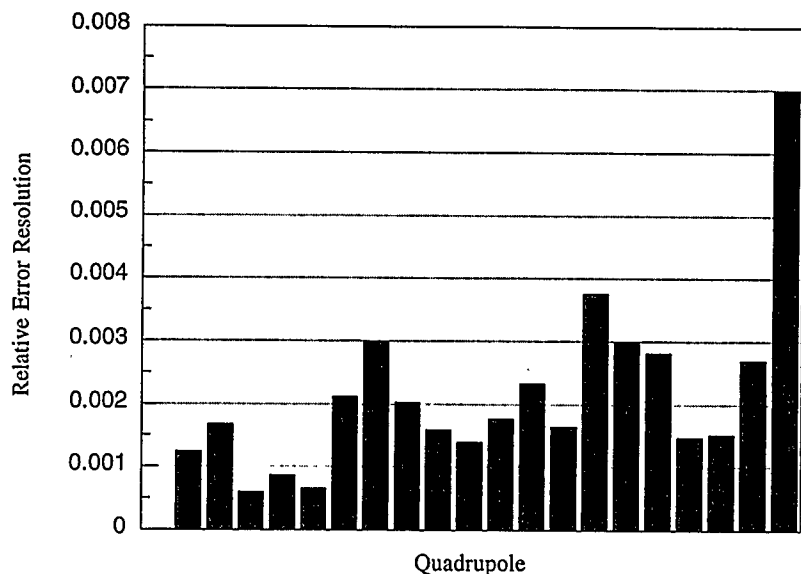
Figure 4.35: Resolution of FFTB lattice diagnostics algorithm for all quadrupole scanned from QN3A to QC3 (QT2A excepted).

lattice. There is a difference between a scaled quadrupole and a scaled energy, which is that the lattice between BPMs downstream of the stepped quad is also incorrect. In a large dataset, this manifests itself with a common strength scaling factor for all quads. The resolution of the lattice diagnostics measurement to incoming energy is 20-30 MeV (recall that the beam energy is 46.6 GeV). Bump-based lattice diagnostics are now routinely done early in an FFTB run, in order to correctly set the beam energy. The energy measured at the beginning of the run is typically high by 100-300 MeV relative to the FFTB design.

Figure 4.34 shows the quadrupole strength error, $(K_{model} - K_{real})/K_{model}$, from a simultaneous fit of over 100 bumps. Note that virtually all quads are within 2 standard deviations of the model (a quad at its model strength has a deviation of zero in Figure 4.34). Figure 4.35 shows the statistical measurement uncertainty of

the quad strengths. Figures 4.34 and 4.35 include all quads up to the doublet, save for QT2A. The relative strength error reported for QT2A is -0.103 ± 0.045. The reason for the large error is that QT2A and QT2B are too close in physical space to decouple in this measurement. Note, however, that typical resolutions in Figure 4.35 are 2×10^{-3}, with a worst-case of 7×10^{-3} for QC3.

The reason for the large QC3 error is that QC3's measurement is strongly coupled to the measurements of the strength of the doublet magnets. The doublet strengths are arguably the most crucial in the FFTB system, and yet ironically their measurement precision is poor, on the order of 1% for each magnet. The first extraction-line quad, QP1A, has a precision worse than 3%. The doublet magnets are long, powerful devices set close together. Because of their arrangement on the common table, the freedom of motion for the magnets is minimal. Furthermore, the arrangement of BPMs downstream of the doublet is far from ideal for this measurement. The most useful location, the FP, does not have a BPM of any kind (largely because the relevant real estate is crammed with beam size monitors). A doublet strength measurement would be quite useful because the doublet quads are each unique (unlike the virtually-identical FFTB standard quads), and are too strong to mini-standardize when scanned (see Section 4.2.4). In addition, the settings of the doublet which set the waist at each monitor have never reproduced well from run to run (see Section 5.3.2). However, such a measurement is essentially prohibited by the FFTB geometry.

Measuring Quadrupole Roll Angles

A feature of the FFTB latttice diagnostic procedure is the ability to diagnose rolled quadrupoles. This is done by introducing a horizontal bump and measuring the vertical kick imparted, or vice-versa. For a quadrupole with a known R_{21} and R_{43}, the out-of plane kick element, R_{41}, for a roll angle θ is given by:

$$R_{41} = \frac{1}{2} \sin 2\theta (R_{43} - R_{21}), \qquad (4.14)$$

and R_{23} is equal to R_{41} in this case. For an FFTB standard quad operating at 300 kilogauss integrated gradient in a horizontally-focusing polarity, $R_{21} = -0.19015\text{m}^{-1}$, and $R_{43} = 0.19397\text{m}^{-1}$. To lowest order, therefore, we expect that the resolution for $\sin 2\theta$ will be equal to the resolution for the normal quadrupole strength. This will yield a resolution for θ which is on the order of 1 mrad.

The difficulty with this measurement is the possibility that the mover axis is not parallel to the magnetic coordinates of the FFTB. For example, consider a magnet which is not rolled in some coordinate system, and a mover system which transports this magnet; and postulate that the mover coordinate system has an angle ψ relative to the magnet, such that a move in the "x" coordinate of the mover by an amount p causes a translation of $p\cos\psi$ in the horizontal and $p\sin\psi$ in the vertical. In this case, the "effective R_{41}, or $R_{41,eff}$, is:

$$R_{41,eff} = R_{43}\sin\psi. \tag{4.15}$$

The only difference at this level between a magnet with roll angle 2θ and a magnet mover with a roll angle ψ is that R_{43} does not exactly equal $\frac{1}{2}(R_{43} - R_{21})$. If a magnet has a mover with rolled coordinate axes, the roll angle determined by introducing horizontal bumps will therefore be different from the roll angle determined with vertical bumps; this difference permits the rolled mover and the rolled quadrupole to be distinguished. Note, however, that the difference between R_{43} and $\frac{1}{2}(R_{43} - R_{21})$ is only 1% of R_{43} itself. Unless the roll angle can be determined with 1% precision, it is not possible to tell the difference between rolled movers and rolled magnets. Furthermore, at the level of hundreds of μrad, the roll angles of the BPMs themselves become suspect. The problem of defining a coordinate system for a beamline, and rolling all devices onto that system at the milliradian level, remains for the present unsolved.

Systematic Errors

The role of BPM scale factors in systematically altering the correction applied to the magnets has been discussed above. As mentioned, a global scale factor can result in a fractional change in the applied correction, but this problem can be eliminated through repeated applications of the lattice diagnostics algorithm. Another possible systematic error is the aforementioned rolled mover system. As shown above, a mover system with a roll angle ψ will produce a move which is scaled by $\cos\psi$ in the plane of the bump. This in turn will lead to a magnet strength which is also scaled by $\cos\psi$ from the actual value. Even large values of ψ, such as 10 mrad, result in a systematic error at the 10^{-4} level, too small to see with this technique and this dataset. An individual mover scale factor will directly scale the strength of a magnet (as opposed to BPM scale factors, which scale the correction to the strength of the magnet). Such a scale factor is only possible in the case of a common scale-factor for all 3 LVDT's incorporated into a given mover. Note, furthermore, that a common scale factor for all mover LVDT's will result in all magnet motions having a common scale factor, and this will be interpreted by the fitting process as a contribution to the BPM scale factor; only the case of 3 LVDT's with a common scale factor placed on a single mover will result in that magnet's strength fit being scaled. The expected amount by which a given LVDT may be scaled relative to the ensemble is 0.1% [31], and therefore the amount by which 3 LVDT's may be collectively scaled relative to the ensemble is close to 0.05%.

4.5 Sextupole Alignment

Once the linear optics elements of the FFTB line have been measured in strength and alignment, the CCS sextupoles (SF1A/B and SD1A/B) are excited to magnetic field levels near their design values ($\int Sdl = 33,000$ kG/m, where design $\approx 32,500$ kG/m for Laser-Compton BSM operation). Because horizontally-misaligned sextupoles pro-

duce normal quadrupole effects (causing waist shifts and horizontal dispersion at the FP), and vertically-misaligned sextupoles produce skew quadrupole effects (causing coupling and vertical dispersion at the FP), it is necessary to move the sextupoles onto the line through the quadrupoles determined in Section 4.2. This, in turn, requires a beam-based alignment technique for sextupole magnets.

The technique used at the SLC Final Focus for sextupole alignment is to measure the waist location, coupling, and dispersion at the IP for two or more levels of sextupole excitation [53]. Such an approach was considered for the FFTB as well. The method has the advantage of directly sampling the effects which must be cancelled (quadrupole aberrations). The disadvantage of the technique is that it requires the ability to measure the beam at the FP as a prerequisite, which implies that some degree of tune-up be performed prior to this alignment (tune-up which becomes to varying degrees obsolete once the sextupoles are moved into position). The technique also becomes more sensitive as the FP beam size decreases and the horizontal and vertical divergences increase. Thus the FP-based technique would benefit from use of the experimental beam-size monitors, which in turn requires that all BSM-related overhead and preparation be absorbed prior to sextupole alignment.

The FFTB's arrangement of sextupoles on movers invites an alternative technique, which depends only on the BPMs and magnet movers, can be performed at low divergences and at any stage of tune-up, is relatively quick to implement, and extremely simple in design and execution.

4.5.1   FFTB Sextupole Alignment Method [54]

Consider the vertical magnetic field, $B_y(x, y)$, in a magnet with (potentially) nonzero dipole, quadrupole, and sextupole terms. The field as a function of coordinates (x_m, y_m), where x_m and y_m are the coordinates relative to an arbitrary origin (such as the origin of a magnet-mover system), is given by:

$$B_y(x_m, y_m) = B_{0y} + G_N x_m - G_S y_m + S_N(x_m^2 - y_m^2) - 2S_S x_m y_m. \qquad (4.16)$$

Here G_N and G_S refer to the normal and skew components of the quadrupole field, respectively; and S_N and S_S refer to half the normal and skew components of the sextupole field, respectively (see Appendix B for details of this expansion). We can apply a coordinate transformation to a different origin:

$$x_m = x + x_0, \qquad y_m = y + y_0, \qquad (4.17)$$

where x and y represent the distance from the new origin to the point of interest, and x_0 and y_0 the distance from the mover origin to the new origin. Substituting Equation 4.17 into Equation 4.16 allows us to expand and collect terms as follows:

$$
\begin{aligned}
B_y(x, y) &= B_{0y} + G_N x_0 - G_S y_0 + S_N(x_0^2 - y_0^2) - 2S_S x_0 y_0 \qquad (4.18) \\
&\quad + G_N x - G_S y + 2S_N(x x_0 - y y_0) - 2S_S(x y_0 + y x_0) \\
&\quad + S_N(x^2 - y^2) + 2S_S x y.
\end{aligned}
$$

A further collection of terms can be made, in which we replace all constant terms with a single term, B_{1y}, and gather terms proportional to x and to y:

$$
\begin{aligned}
B_y(x, y) &= B_{1y} \qquad (4.19) \\
&\quad + x(G_N + 2S_N x_0 - 2S_S y_0) - y(G_S + 2S_N y_0 + 2S_S x_0) \\
&\quad + S_N(x^2 - y^2) + 2S_S x y.
\end{aligned}
$$

We can now require that in the current coordinate system, the coefficient of x and the coefficient of y are identically zero. In this coordinate system there is no gradient, and the coordinates (x_0, y_0) represent the transformation from the origin of the initial coordinate system to the center of the sextupole, defined as the point at which the gradient vanishes. After some algebra, we find:

$$
\begin{aligned}
x_0 &= \frac{-G_N S_N - G_S S_S}{2(S_N^2 + S_S^2)}, \qquad (4.20) \\
y_0 &= \frac{S_S G_N - G_S S_N}{2(S_N^2 + S_S^2)}.
\end{aligned}
$$

We may also derive the values of the original G_N and G_S in Equation 4.16:

$$G_N = 2S_S y_0 - 2S_n x_0, \qquad G_S = -2S_N y_0 - 2S_S x_0. \qquad (4.21)$$

We can also rewrite Equation 4.16 in another fashion, that of a parabola in x:

$$B_y(x_m, y_m) = S_N(x_m - B)^2 + C, \qquad \text{where} \qquad (4.22)$$
$$B = \frac{2S_S y_m - G_N}{2S_N} \equiv x_E, \qquad \text{and}$$
$$C = B_{0y} - G_S y_m - S_N y_m^2 - \left(\frac{2S_S y_m - G_N}{2S_N}\right)^2.$$

Equation 4.22 indicates that if B_y is measured as a function of x_m, the correlation between the two will be parabolic, with extremum occurring at $x_m = \frac{2S_S y_m - G_N}{2S_N}$. Let (x_1, y_1) be defined as the initial position of the magnet in question, such that the initial (x_m, y_m) are given by $(x_1 + x_0, y_1 + y_0)$. In this case, the extreme value of the parabola, x_E, can be shown to be given by:

$$x_E = x_0 + y_1 \frac{S_S}{S_N}. \qquad (4.23)$$

The measurement of B_y can be practically accomplished by moving the sextupole across the incoming beam via its mover. This produces a horizontal kick which translates to horizontal offset in position downstream. By plotting the beam position at a downstream BPM versus the mover position, a parabolic correlation can be seen. The mover position which minimizes or maximizes the horizontal BPM signal is x_E from Equation 4.23. Note that in the case of $S_S \equiv 0$, scanning the magnet horizontal position will always give $x_E = x_0$. If we assume a skew sextupole contribution which is no more than 10% of the normal sextupole, and a magnitude of y_1 no larger than 1 mm, we find that x_E may deviate by 100 μm from x_0

Let us now assume that the magnet is moved to position x_E from Equation 4.23, and rename this point x_2. In analogy with Equation 4.22, we can write an equation for B_y as a function of y_m:

$$B_y(x_m, y_m) = -S_N(y_m - B)^2 + C, \qquad \text{where} \qquad (4.24)$$

$$B = \frac{-2S_S x_m - G_S}{2S_N} \equiv y_E, \qquad \text{and}$$
$$C = B_{0y} + G_N x_m + S_N x_m^2 - \left(\frac{-2S_S x_m - G_N}{2S_N}\right)^2.$$

Some manipulation reveals a relationship between y_E, the value of y_m at which the extreme value of B_y occurs, and y_0:

$$y_E = y_0 + y_1 \left(\frac{S_S^2}{S_N^2}\right). \qquad (4.25)$$

The error contribution in this case is reduced by a factor of S_S/S_N, as we have assumed that the magnet was moved to within $y_1 \frac{S_S}{S_N}$ of x_0 in the horizontal. It is evident that repeated iterations of this procedure will suffice to converge (x_E, y_E) to within any desired precision of (x_0, y_0).

During FFTB sextupole alignment, the CCS sextupoles are stepped from -1.4 mm to +1.4 mm in the mover coordinate system, in steps of 400 μm. At each mover position, the downstream BPMs and the BPMs adjacent to the sextupole are read out for 9 pulses. This allows sufficient redundancy to discard individual bad pulses, and does not significantly increase the time needed to perform sextupole alignment (as the time needed to reposition the magnet is the rate-determining step). The procedure calls for only one iteration of horizontal and one iteration of vertical alignment per magnet. Equations 4.23 and 4.25 show that such an algorithm creates the potential for systematic misalignments, in the presence of nonzero skew sextupole. This possibility is considered below.

Expected Resolutions and Systematic Errors

In order to determine the expected resolution of the parabolic fit to a sextupole mover scan, it is necessary first to consider a quadratic fit:

$$x_{BPM} = a_0 + a_1 x_s + a_2 x_s^2, \qquad (4.26)$$

where x_{BPM} is the beam position at the BPM and x_s is the sextupole mover position. We can construct a curvature matrix for the polynomial fit:

$$\alpha_s = \begin{pmatrix} \sum \frac{1}{\sigma_{BPM}^2} & \sum \frac{x_s}{\sigma_{BPM}^2} & \sum \frac{x_s^2}{\sigma_{BPM}^2} \\ \sum \frac{x_s}{\sigma_{BPM}^2} & \sum \frac{x_s^2}{\sigma_{BPM}^2} & \sum \frac{x_s^3}{\sigma_{BPM}^2} \\ \sum \frac{x_s^2}{\sigma_{BPM}^2} & \sum \frac{x_s^3}{\sigma_{BPM}^2} & \sum \frac{x_s^4}{\sigma_{BPM}^2} \end{pmatrix}. \qquad (4.27)$$

Because the pattern of x_s steps is symmetric about zero, all odd-powered sums in Equation 4.27 are identically zero. Substituting the step pattern described above for x_s and assuming 1 μm for the resolution σ_{BPM} allows us to invert α_s into an error matrix, ϵ_s. For the sake of brevity, only the relevant terms of ϵ_s are reproduced here:

$$\epsilon_s(a_1, a_1) = 1.49 \times 10^{-7} \equiv \sigma_1^2, \ \epsilon_s(a_2, a_2) = 2.33 \times 10^{-13} \mu\text{m}^{-2} \equiv \sigma_2^2, \ \epsilon_s(a_1, a_2) = 0. \qquad (4.28)$$

Recalling the conversion from a_i to parabola coefficients $y = A(x - B)^2 + C$, we find that $B = -\frac{a_1}{2a_2}$, and therefore

$$\begin{aligned} \sigma_B^2 &= \sigma_1^2 \left(\frac{\partial B}{\partial a_1}\right)^2 + \sigma_2^2 \left(\frac{\partial B}{\partial a_2}\right)^2 \\ &= \frac{1}{A^2} \left(\sigma_1^2 + B^2 \sigma_2^2\right). \end{aligned} \qquad (4.29)$$

Given the values of σ_1 and σ_2 in Equation 4.28, the error term proportional to B contributes when $|B| \approx 1$mm, and represents the familiar difficulty of fitting a parabola center when one wing of the parabola dominates. For this analysis the "one-wing" term will largely be ignored.

All that remains in determining the value of B is to determine the value of A, the parabola curvature parameter. "A" represents the response of the beam measurement at the BPM to the movement of the sextupole, and can be approximated by $A \approx K_s R_{12}^{\text{sext} \to BPM}$, where $K_s = \frac{B_0 L}{a^2 B \rho}$, the integrated sextupole strength divided by the magnetic rigidity. For FFTB sextupoles at integrated second derivative ("BDES") of 33,000 kG/m, $K_s = 10.61$ m^{-2}. Given the choice of downstream BPMs (based largely on experience, observing which BPM gives the best resolution for a given

sextupole), the values for A are: $2.23 \times 10^{-4} \mu$m^{-1} for SF1A, $9.86 \times 10^{-5} \mu$m^{-1} for SF1B, $4.24 \times 10^{-5} \mu$m^{-1} for SD1A, and $6.00 \times 10^{-5} \mu$m^{-1} for SD1B. When we substitute these values for A into Equation 4.29, we obtain the following resolutions for the sextupole alignment procedure:

$$\sigma_{\text{SF1A}} = 1.7 \mu\text{m}, \ \sigma_{\text{SF1B}} = 3.9 \mu\text{m}, \ \sigma_{\text{SD1A}} = 9.1 \mu\text{m}, \ \sigma_{\text{SD1B}} = 6.5 \mu\text{m}. \qquad (4.30)$$

The resolutions quoted above are only appropriate for BPM resolution of 1 μm; however, no attempt has been made to fit out incoming jitter from the BPM values, and 9 BPM values are read in at each mover position. The effect of this is to worsen the resolution by a factor of $\sigma_{jit}/3$, where σ_{jit} is as shown in Figure 4.3. Including this factor, the expected statistical resolution falls to:

$$\sigma_{\text{SF1A}} = 3.8 \mu\text{m}, \ \sigma_{\text{SF1B}} = 3.9 \mu\text{m}, \ \sigma_{\text{SD1A}} = 12.1 \mu\text{m}, \ \sigma_{\text{SD1B}} = 34.7 \mu\text{m}. \qquad (4.31)$$

Note that, while the sextupole alignment resolution with jitter and 9 pulses per mover value remains within the global tolerances quoted in Table 2.3, significant improvement is possible for SD1B with reconstruction of incoming trajectories.

In Equations 4.23 and 4.25, the systematic error arising from the initial position of the magnet is noted. The maximal systematic error is estimated with a skew sextupole content of 10% relative to the normal sextupole content. In Appendix B, the technique for sextupole field measurements with a stretched wire is described. The measurement includes a determination of the roll angle of the sextupole field relative to the split planes (i.e., the roll angle needed when the magnet is installed, "upright" relative to its mechanical coordinates, to make the skew sextupole field vanish). The maximum value found for this angle is 2.5 mrad. Assuming an installation error of 2.5 mrad as well, the maximum roll angle of the sextupole field with respect to the external coordinate system is 5 mrad. This gives a skew component equal to $S_N \sin(3\theta)$, or 1.5%. For a 1 mm initial offset of the magnet, this translates to 15 μm systematic error in the initial determination of x_0. With a second pass of horizontal alignment following the vertical, this error can be reduced below 1 μm.

The other potential source of systematic error is a magnet mover which moves the magnet at an angle relative to the beamline coordinate system. Consider in this case a sextupole magnet with no skew component, sitting at some position, (x_1, y_1), relative to the magnetic center:

$$B_y = S_N(x_1^2 - y_1^2) + B_0. \tag{4.32}$$

Now assume the magnet mover horizontal axis is set at an angle θ relative to the magnetic field axis. When the magnet is translated by an amount p along this axis, the translation in x is $p\cos\theta$, the translation in y is $p\sin\theta$. At this location the magnetic field is given by:

$$B_y = B_0 + S_N\left[(x_1 + p\cos\theta)^2 - (y_1 + p\sin\theta)^2\right]: \tag{4.33}$$

A certain amount of manipulation gives a magnetic field equation in terms of p:

$$B_y = B_0' + S_N(p - p_0)^2, \text{ where } p_0 = y_1\sin\theta - x_1\cos\theta. \tag{4.34}$$

If the magnet mover is now positioned at the location p_0 along its direction of travel, the resultant coordinates in the magnetic coordinate system are:

$$x = x_1(1 - \cos^2\theta) + y_1\sin\theta\cos\theta, \quad y = y_1(1 + \sin^2\theta) - x_1\sin\theta\cos\theta. \tag{4.35}$$

For any reasonable roll angle, the $\cos^2\theta$ contribution to the x position can be ignored. The significant contribution is that the new x position is different from the magnetic center position by $\frac{y_1}{2}\sin 2\theta$. For an initial offset of 1 mm, and a mover roll angle of 10 mrad, the accepted value of x differs from the center position by 10 μm. Based upon lattice diagnostics, a 10 mrad angle for this quantity is large, but not implausible. Note that, as in the case of skew sextupole content, this error depends upon the initial offset of the non-scan plane from the center, and can easily be eliminated by iterating the tuning procedure.
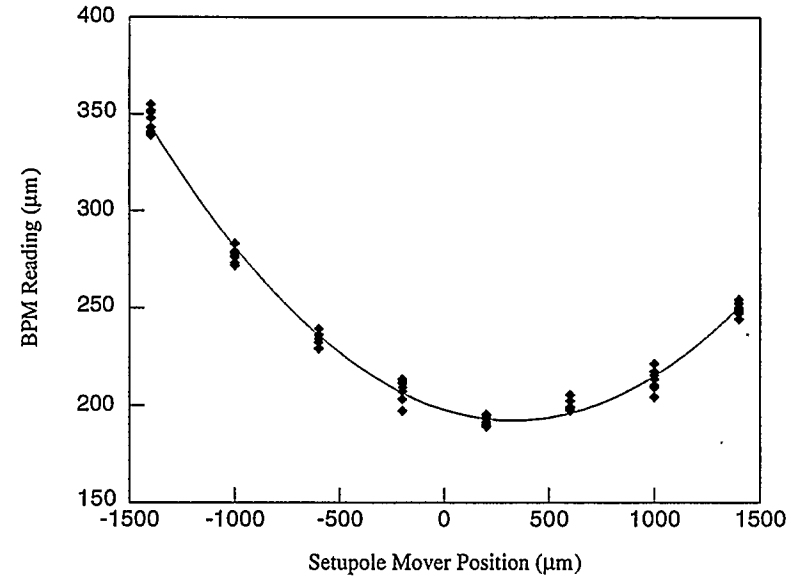
Figure 4.36: Horizontal BPM reading from QM2 BPM as a function of SD1A vertical mover position.

4.5.2 Results of Sextupole Alignment

Figure 4.36 shows a representative BPM position versus sextupole mover scan. The characteristic parabolic dependence is clearly visible, along with the pulse- to-pulse jitter discussed above. The scan in Figure 4.36 has a fitted center resolution of 8 μm.

Note that the concavity of the parabola is a function of sextupole polarity and transport properties from the sextupole to the BPM. In the FFTB, all sextupoles have the same polarity; therefore, for BPMs separated by π or less in betatron phase from the sextupole being scanned, the expected concavity is "up" for vertical mover scans and "down" for horizontal. During the first FFTB sextupole alignment tests in April of 1994, the first CCSY sextupole (SD1A) was found to have the opposite concavity pattern. Gaussmeter tests subsequently confirmed that the sextupole had
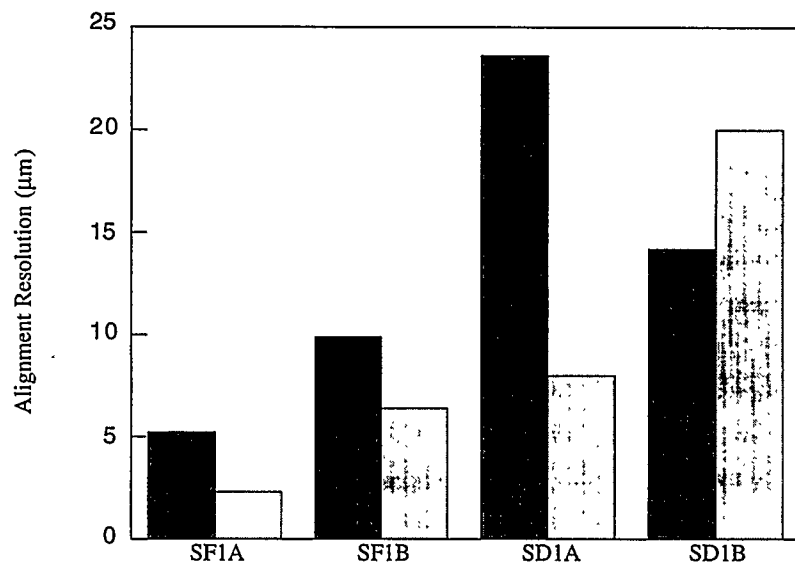
Figure 4.37: Resolution of sextupole alignment procedure from March 1995, for both horizontal (dark) and vertical (light) measurements.

been miswired.

Figure 4.37 shows the horizontal and vertical sextupole alignment resolutions from March of 1995. While most results were within a factor of 2 of the expected resolution, the horizontal resolution for SD1A was considerably worse than expected. This is because the horizontal distance between the center of the mover system and the position which centered the magnet on the beam was almost 1 mm, and the second term in Equation 4.29 became significant.

Once all local diagnostics have been performed, and corrections applied, the FFTB optics can be adjusted for large FP angular divergences, using the incoming beam matrix measured in Chapter 3 to compute settings for the beta matching quadrupoles. The match is verified on WS2 and WS3, and the final tuning is then performed using the FP beam size monitors. This tuning is the subject of Chapter 5.

Chapter 5

Global Beam-Based Diagnostics

Once the local tuning procedures of Chapter 4 have been performed, and the incoming beam has been measured and matched via the techniques described in Chapter 3, the only remaining diagnostic is the beam size at the focal point. The design spot size at the focal point of a linear collider is the smallest of the entire machine, and the only one which sees the entire beamline up to the collision; it is therefore *de facto* the only point relevant for final tuning [1].

Tuning the FP spot size implies several pre-requisites: one or more beam size monitors which may be used to measure the beam; a finite list of aberrations which are to be tuned by minimizing the measured spot; and orthogonal knobs for changing the strength of each of these aberrations.

Section 2.5.4 described the two exotic spot-size measurement devices present at the FFTB focal point, and Section 2.5.2 described WS6A and WS6B, the more conventional solid-wire measurement devices. All of these play a role in the tuning of the final spot, and their uses are described more fully in Section 5.2 below.

Table 2.2 lists the principal aberrations which are expected at the FFTB, and the knobs which are expected to cancel same. Note that Table 2.2 includes only those

---

[1] It is conceivable that a final focus system could be designed with a tuning image downstream of the IP [55], which would see the entire beamline; only the IP, however, will be sensitive to the smallest aberrations.

aberrations which are inherent to the FFTB, i.e., those which arise from errors in the strength, alignment, or construction of the magnets in the FFTB itself. Three additional errors are possible in the FFTB. The horizontal and vertical β^* can be misset, and this arises primarily from errors in the measurement of the incoming beam and/or errors in matching verification. An additional error is coupling between x and y at the FP, which can come from the incoming beam or from a roll angle in the beam size monitors, but not from errors in the FFTB itself. As discussed in Sections 5.2.2 and 5.2.3, the Laser-Compton BSM is potentially subject to errors in roll installation, while the Gas-Ion BSM is immune to same. WS6A is, of course, also subject to these errors, but since WS6A cannot be used to measure the minimum spot size this is not a critical limitation. Corrections for each of these three additional errors are possible within the limits of FFTB optics. The betatron error can be corrected with a magnification "Irwin Knob" of the type discussed in Section 3.3. The xy coupling or monitor roll is more difficult. The FFTB contains 3 skew quadrupoles, such that the incoming coupling can be tuned to eliminate xy and x'y coupling at the FP (corresponding to Irwin notation elements b^* and d^*, see Appendix A), and the third skew quad can be used to eliminate the effects of rolled quads in the FFTB. In practice, the location of QSM2 (see Section 2.1.2) is nearly degenerate with QSM1 for all reasonable strengths of Q5 through QA1. A knob for tuning the xy coupling at the Laser-Compton monitor has been devised which uses QSM2 in the beta match and QS3 in the final transformer; essentially, the two skew quads fight one another, and the QSM2 also gives a very small amount of b^*. This knob is sufficient to diagnose a roll angle in the monitor, and to correct approximately 0.5°, according to simulation. For larger roll angles, the knob begins to break down: as the determinant of the out-of-plane R sub-matrix diverges from zero, the determinant of the in-plane sub-matrices diverge from 1, and the in-plane beam begins to change. One way to think of this is that the first skew quad produces coupling between the xx' and yy' phase planes; the second skew quad also produces such coupling, but the two phase planes

are already coupled by the first skew quad. Thus the combined effect of the two skew quads is to cause a net change in the downstream xx' phase plane which depends upon the upstream xx' phase plane.

5.1 Generation of Global Knobs

Horizontal and Vertical Waist: The FFTB waist knobs use the QC2 and QC1 magnet strengths. The range available is considerable: in 1993 operations, these knobs were used to move the waists onto WS6B, 75 cm downstream of the Laser-Compton monitor. These knobs are specific to particular FT and final doublet configurations; in particular, different QC2 and QC1 coefficients are needed to move the waists for each of the three possible FP's (Laser-Compton, WS6A, Gas-Ion). While the knobs are designed to be orthogonal and linear around zero, they are non-orthogonal and noticably nonlinear outside of ±5 cm motions.

Horizontal and Vertical Dispersion: Horizontal and vertical dispersion knobs use correctors H5D and V5D (in the final transformer) to displace the beam in the final doublet. The strong chromaticity of the doublet generates η^* in the appropriate plane. Knobs are orthogonal and linear up to the corrector strength limits (corresponding to $\eta_x^* = \pm 4.9$ mm, $\eta_y^* = \pm 691$ μm).

Coupling: The x'y coupling term is controlled by QS3, which produces this aberration in nearly pure form (due to its proximity to the doublet). The knob is calibrated in units of R_2 [56], where $\sigma_y^2 = \sigma_0^2 + R_2^2 \sigma_x'^2$. In this, the R_2 knob functions similarly to an "out-of-plane waist" knob [57]. The knob range is up to $R_2 = \pm 33$ mm, which would enlarge the vertical spot by almost 6 μm.

Sextupole: The four sextupole aberration knobs are determined by first-principle calculations of the optics. A normal sextupole will produce a horizontal kick given by:

$$\Delta x' = K_s(x^2 - y^2), \qquad \Delta y' = -2K_s xy, \qquad (5.1)$$

where $K_s = \frac{B_0 L_s}{a_s^2 B \rho}$. This kick will then propagate to the FP:

$$\Delta x^* = R_{12} K_s (x^2 - y^2), \qquad \Delta y^* = -2 R_{34} K_s xy, \qquad (5.2)$$

where the x and y on the RHS of both equations is the position of the test particle at the sextupole. We can convert this into the *angle* of the test particle at the FP by noting that the transport from almost any point in the beamline to the FP is parallel-to-point, therefore:

$$x_{sext} = -R_{12} x'^*, \qquad y_{sext} = -R_{34} y'^*, \qquad (5.3)$$

where $R_{12,34}$ are the terms in the transport matrix from the sextupole to the FP. Combining this expression with Equation 5.2 yields:

$$\Delta x^* = -K_s (R_{12}^3 x'^{*2} - R_{12} R_{34}^2 y'^{*2}), \qquad \Delta y^* = +2 K_s R_{12} R_{34}^2 x'^* y'^*. \qquad (5.4)$$

We see that a normal sextupole produces two aberrations: one aberration proportional to R_{12}^3, which enlarges only the horizontal beam size; and one proportional to $R_{12} R_{34}^2$, which enlarges horizontal and vertial sizes. A suitable combination of two such sextupoles can be used to produce the two aberrations independently. Similarly, the effects of a skew sextupole on a test particle are given by:

$$\Delta x^* = +2 K_s R_{12}^2 R_{34} x'^* y'^*, \qquad \Delta y^* = -K_s (R_{12}^2 R_{34} x'^{*2} - R_{34}^3 y'^{*2}). \qquad (5.5)$$

Thus the skew sextupole gives rise to two vertical aberrations, one of which has the same dependencies as a single horizontal aberration. Two skew sextupole magnets can therefore be adjusted in combinations which excite only one of the two aberrations, and hold the other fixed.

The approximation of linearity and orthogonality for these knobs is good over their entire range. The range of adjustability (starting with all sextupoles at zero) is: NSX_X (normal sextupole, x and y aberration), ±1.3 units; NSX_YX, ±2.7 units; SSX_YX, ±1.8 units; SSX_Y, ±1.7 units. The knob scales are adjusted such that

1 "unit" of any knob will enlarge the beam size by 41% of σ_{design}^*; in other words, turning the knob by 1 unit will add 1.0 σ_{design}^* in quadrature with the design spot. With the exception of NSX_X, the knobs are scaled based on their effects on the vertical beam size.

Chromaticity: Because the FFTB sextupoles are not interleaved, the horizontal chromaticity may be adjusted by changing the SF1 strengths only, while the vertical may be adjusted by changing the SD1 strengths. As shown in Figure 2.11, the uncorrected chromaticity of the FFTB line is enormous; therefore, the range of correction is similarly enormous.

The global tuning knobs described above have several operational difficulties. The waist knobs are difficult to scan reproducibly because of magnetic hysteresis effects, and because the required step sizes are quite small. The dispersion knobs produce steering in the FT sextupoles, introducing additional normal and skew quad effects, and also produce orbit changes in the extraction line. For these reasons, the FFTB also makes use of a set of *sextupole mover knobs*[2]. These knobs make use of the normal and skew quad effects generated when a beam passes through a normal sextupole off-center. Because the FFTB CCS sextupoles are all in phase with one another, the linearity and orthogonality of these knobs holds for larger ranges – in effect, the sextupoles do not "see" the effects of upstream sextupole translations. In addition, the sextupole movers operate without hysteresis, and with micron precision, which enhances the settability and reproducibility of the knobs.

The knob coefficients can be calculated by simulating the effect of a thin quadrupole at each of the sextupole locations. The quadrupole will produce waist shifts, coupling and dispersion at the FP, and a linear correlation between quad strength and aberration derived from linear optics. The thin quad strength is then converted to sextupole offset to determine the transformation from sextupole motion to aberrations at the FP. This relationship is inverted to obtain the transformation from desired aberration

---

[2]The sextupole mover knobs were proposed by K. Oide.

to sextupole motions, M.

There are four CCS sextupoles, each of which may be moved in x or y, for a total of 8 possible knobs. In practice, five knobs are used routinely (x and y waist, x and y dispersion, x'y coupling), while a sixth quantity (second-order dispersion, T_{166}) is held fixed. While the range of each knob is quite considerable, the strengths of the sextupoles ultimately may need to be varied to minimize the chromaticity. Therefore, some understanding of the interaction of sextupole strength with the mover knobs is required before use.

Consider the horizontal motions of the sextupoles, and the normal-quadrupole aberrations arising therefrom (waists, η_x, T_{166}). The relationship between these quantities can be written as a matrix equation:

$$
\begin{pmatrix} \text{SF1Ax} \\ \text{SF1Bx} \\ \text{SD1Ax} \\ \text{SD1Bx} \end{pmatrix} = \begin{pmatrix} 2.515 & 5.055 & 93.69 & -8.425 \\ 2.857 & 2.695 & -93.69 & 8.425 \\ -1.028 & -15.79 & -357.1 & -41.62 \\ 0.903 & -28.57 & 357.1 & 41.62 \end{pmatrix} \begin{pmatrix} W_x \\ W_y \\ \eta_x \\ T_{166} \end{pmatrix} , \text{ or} \quad (5.6)
$$

$$ \vec{S} = M\vec{k}. $$

The first column of M shows the moves needed for 1 mm of horizontal waist; the second, one mm of vertical waist, etc. The knob units are all mm, and the mover units are all μm. We can invert this relationship:

$$ \vec{k} = M^{-1}\vec{S}. \quad (5.7) $$

Increasing the strength of the SF1 sextupoles by fraction f is equivalent to multiplying the first two columns of M^{-1} by a constant $1+f$; similarly, increasing the SD1 strength by fraction d will increase the third and fourth columns of M^{-1} by $(1+d)$. Let W be defined as the matrix which is formed by these transformations to M^{-1}. When the mover knobs are incremented, the driving software will use M to move the magnets, but the matrix W converts these moves into real changes in the beamline:

$$ \vec{k}_{real} = WM\vec{k}_{request}. \quad (5.8) $$

The product WM contains all information about the scaling and crosstalk of the hard-wired knobs due to sextupole strength changes:

$$
WM = \begin{pmatrix} \begin{matrix} 1+f \\ +0.004(f-d) \end{matrix} & 1.448(f-d) & 0 & 0 \\[2ex] \begin{matrix} -2.83\times10^{-3} \\ \times(f-d) \end{matrix} & \begin{matrix} 1+f \\ -0.004(f-d) \end{matrix} & 0 & 0 \\[2ex] -1.09(f-d) & \begin{matrix} 7.52\times10^{-3} \\ \times(f-d) \end{matrix} & \begin{matrix} 1+f \\ -0.403(f-d) \end{matrix} & \begin{matrix} -5.3\times10^{-2} \\ \times(f-d) \end{matrix} \\[2ex] \begin{matrix} -1.3\times10^{-2} \\ \times(f-d) \end{matrix} & \begin{matrix} -8.6\times10^{-2} \\ \times(f-d) \end{matrix} & -4.48(f-d) & \begin{matrix} 1+f \\ -0.597(f-d) \end{matrix} \end{pmatrix}.
$$
$$ (5.9) $$

Equation 5.9 shows only three cross-talk terms which are non-trivial: $(WM)_{12}$, $(WM)_{31}$, and $(WM)_{43}$; these terms relate the crosstalk *from* the y-waist *to* the x-waist, from the x-waist to η_x^*, and from η_x^* to T_{166}^*, respectively. The $(WM)_{12}$ is not significant because typical vertical waist motions are hundreds of microns, while $\beta_x^* = 10$ mm. Similarly, while $(WM)_{43}$ could potentially increase the T_{166}^* by millimeters, this still only amounts to a horizontal beam size increase of nanometers at the full energy spread. The only mixing term of concern, therefore, is $(WM)_{31}$, which increases the horizontal dispersion if the horizontal waist is not set to zero. Note further that if $f = d$, corresponding to changing both sextupole pairs by the same fraction during chromaticity scans, then Equation 5.9 reduces to scaling all the knobs to $1 + f$ of their set values.

How far may the sextupole mover knobs be moved from their zero (aligned) positions? In order to answer this question, it is necessary to determine the maximum value of f needed to constrain the chromaticity. Simulation studies have shown that,

Knob	Total Range	Maximum Set Value	RMS value from sext. alignment
x waist	±400 mm	±32 mm	2.28 mm
y waist	±35 mm	±435 μm	640 μm
η_x^*	±2.8 mm	±3.2 mm	38 μm
η_y^*	±1.6 mm	±104 μm	17 μm
R_2	±46 mm	±1.9 mm	490 μm
T_{166}^*	±24 mm	±629 mm	372 μm

Table 5.1: Range considerations for sextupole mover knobs. The total range is determined by allowing the sextupoles to be scanned up to 1 mm from their aligned positions; maximum set value is derived from the "scan with impunity" criterion in the text; the RMS value is the quantity of each aberration caused by the finite sextupole alignment resolutions in Figure 4.38.

at the design energy spread of 0.3%, a scan of ±6% of the sextupole strength is sufficient to determine the sextupole strength which nulls the chromaticity to within the tolerance, assuming a BSM with 10% resolution. Similarly, a scan range of ±4.6% is required to null the vertical chromaticity to within the tolerance. Any knob which is not set to zero, therefore, will become 6% stronger at one extreme of a horizontal chromaticity scan, and 6% weaker at the other; and similarly the vertical chromaticity scan will change the strength of the knobs by +4.6% and -4.6%.

The limits on the sextupole mover knobs can be determined by applying a "scan with impunity" tolerance: over the total range of the sextupole scan, changes in the strengths of the mover knobs do not change the FP spot size by more than 2%, and therefore the chromaticity scans can be considered unambiguous. With the expected beam conditions, this means that the horizontal waist can shift by up to 2 mm, the vertical waist can shift by up to 20 μm, and η_x^* can shift by up to 113 μm. Therefore, 2 mm represents 6% of the maximum setting of the horizontal waist knob, 20 μm represents 4.6% of the maximum setting of the vertical waist knob, etc. The full ranges of the sextupole mover knobs are shown in Table 5.1.

One other factor in the use of sextupole mover knobs is that the sextupole align-

ment procedure in Chapter 4 has a finite resolution, shown in Figure 4.37. Therefore, the residual incoherent misalignments of the sextupoles after beam-based alignmnent will give rise to waist shifts, dispersion, etc. The expected RMS values of these aberrations from residual misalignments are the last column of Table 5.1. For most of the knobs, the RMS aberration from this source is small compared to the maximum value of the knob which is allowed. This is not true for the vertical waist knob: the RMS waist shift from misalignments is 50% larger than the "scan with impunity" maximum value. This introduces ambiguity into the sextupole scans. Because of this, the chromaticity is measured via a different procedure, described in Chapter 6.

During the initial tuning of the FP spot, the value of a given sextupole mover knob which minimizes the spot size is frequently out of the allowed range shown in column 2 of Table 5.1. In this case, the desired value of the mover knob is converted to an equivalent magnet knob, which is then set to minimize the spot. Thus, while the mover knobs are always used for FP scans, only small increments of the knobs are applied.

5.2 Tuning on the Beam Size Monitors

Once a complete set of multiknobs is defined for each set of beam size monitors, the global tuning procedure can begin. This tuning takes somewhat different forms on each monitor, corresponding to their various designs.

5.2.1 Tuning on WS6A

In order to prevent breakage of the 4 μm carbon wires which make up WS6A, it is necessary to maintain beam sizes such that $\sigma_x \sigma_y \geq 3 \mu\text{m}^2$ (see Section 2.5.2). This leads to a conundrum: how can one guarantee that the beam sizes are *a priori* large enough to measure with the wires without first making such a measurement?

In early FFTB runs, no systematic attempt was made to answer this question,
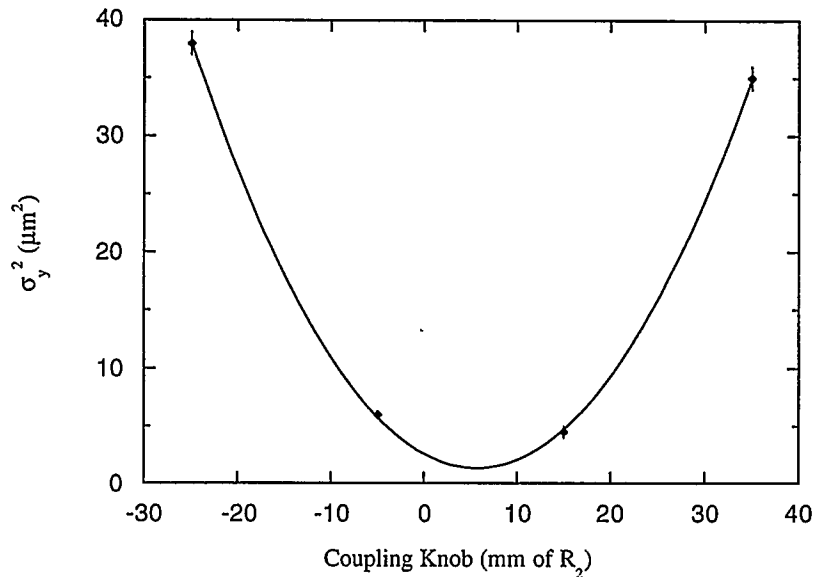
Figure 5.1: Use of WS6A to tune aberrations. In this scan, the vertical waist has been moved off the wire such that the minimum spot seen by the wire is 1.2 μm, while the focused spot is close to 300 nm.

and as a result several wires were in fact broken. During this time it was noticed that the strength of QS3 needed to optimize the beam size on any monitor never exceeded ±1 kilogauss. For later runs, a technique was developed which allows the beam to be tuned on WS6A without breaking any wires, by first setting QS3 to 10 kilogauss to ensure that the vertical beam size is large enough to prevent breakage. The aberrations (waist, dispersion, coupling) can then be tuned, so long as at least one aberration remains large enough to maintain the beam size. For example, after setting QS3 to 10 kilogauss, the value of QS3 is slowly reduced until the beam size at WS6Ay reaches 2 μm. At this time the waist knob is moved until the beam size reaches 4 μm, and the coupling is scanned. Figure 5.1 shows such a scan. The technique works because the location of the minimum of the parabola depends upon the measurements in the wings of the parabola, and not the measurement of the

minimum.

What is the advantage of tuning the beam size on WS6A? Sections 5.2.2 and 5.2.3, below, note among other things the difficulty of acquiring a signal on the more exotic beam size monitors. The Laser-Compton monitor, in particular, has a very small dynamic range – the "Small-y" mode can measure reliably only between 45 and 200 nanometers. The process of obtaining a 200 nanometer spot becomes a potentially painful exercise in bootstrapping, unless the total aberration content at the waist can be reduced before looking for a signal at the BSM. The waist can be moved to WS6A, at which point the dispersion and coupling can be minimized. The optics needed to focus the waist at WS6A and either of the two beam size monitors are sufficiently similar that this technique allows reduction of the two aforementioned aberrations, and increases the likelihood of capturing a signal in the Laser-Compton or Gas-Ion BSMs with only waist scans. While this technique has been developed too recently to be fully exploited, future FFTB runs will certainly utilize it for increased efficiency.

One difficulty encountered in using the WS6A wires is that the wires are fixed in position; WS6A tuning is always preceded by several wide scans used to find the wire clusters. This is especially problematic in the vertical, where the corrector magnet which scans the beam across the Laser-Compton BSM's interference pattern is used to scan the beam across the wires. The total range of motion is quite small, and frequently either the horizontal or vertical wires have not been located at all due to installation alignment difficulties. This problem is being addressed by a redesign of the wire scanner forks: the vertical yoke will retain 50 μm wire spacing but will be strung with 20 or 30 wires, while the horizontal yoke spacing will be increased to 100 μm [58].

5.2.2 Laser-Compton Beam Size Monitor

In order to tune the electron beam on the Laser-Compton Beam Size Monitor, it is first necessary to tune the monitor itself, most particularly the many laser pathways,
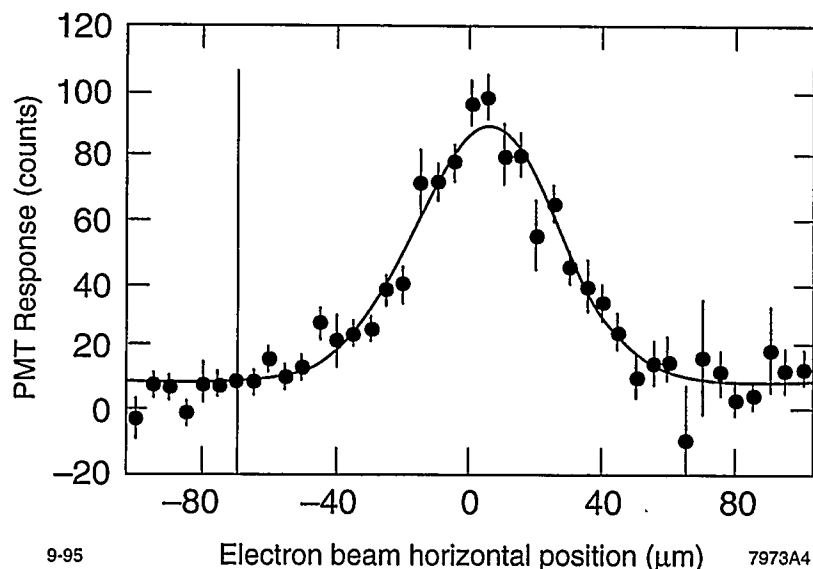
Figure 5.2: Scan of a single laser pathway with the electron beam, for transverse alignment of the laser pathway. Note that the size of the resultant Gaussian is simply the RMS size of the laser.

the Nd:YAG laser itself, and the signal-to-noise performance of the system. Each of these is discussed below.

Tuning the Laser Pathways:

In order to form the interference pattern shown in Figure 2.13, it is essential to cross the two laser pathways at a point in space; and furthermore, to ensure that the point is also a point along the path of the electron beam. This involves a *transverse* alignment, and also a *longitudinal* alignment. Each laser pathway includes at least one mirror which may be rotated in two degrees of freedom to accomplish this alignment. The total system also includes mirrors used for the alignment of the incoming laser, mirrors for feedback on the laser position, etc. These additional mirrors and their functions are beyond the scope of this discussion.

Transverse alignment of the laser pathways is done explicitly with the electron
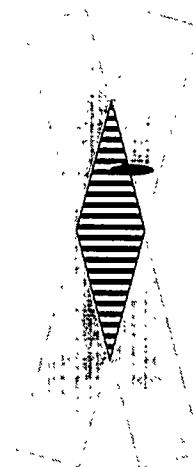
Figure 5.3: Loss of transverse alignment of the Laser-Compton BSM relative to the electron beam (oval) results in both reduced modulation depth and reduced average intensity, since one laser pathway now contributes relatively few photons in the path of the electron beam.

beam. The shutters inside the BSM are adjusted such that only one laser pathway is available (thus no interference pattern is formed); the electron beam is scanned across the laser path, and the photomultiplier tube downstream is read out. The resulting pattern of the detected photon intensity vs. electron beam position is a Gaussian distribution, with RMS width given by the width of the laser beam at the FP, and offset given by the misalignment. The laser path is moved by the amount indicated by the fit, and the scan is repeated until convergence. Figure 5.2 shows such a scan. Note that the RMS width of the laser is only 25 μm. Earlier operations with the BSM used a larger laser "target". The smaller laser size at the FP results in higher photon intensities at the FP and lower intensities at the lenses in the BSM, but also makes the alignment more critical.

All of the laser pathways may be aligned in turn. This alignment needs to be repeated periodically during the course of a day. As Figure 5.3 shows, the diagnostic signal that triggers transverse alignment is weakening of the average Compton signal

(a)   (b)

z (electron
beam
direction)

$\longrightarrow$

|←——————————————→|

100 µm slit aperture
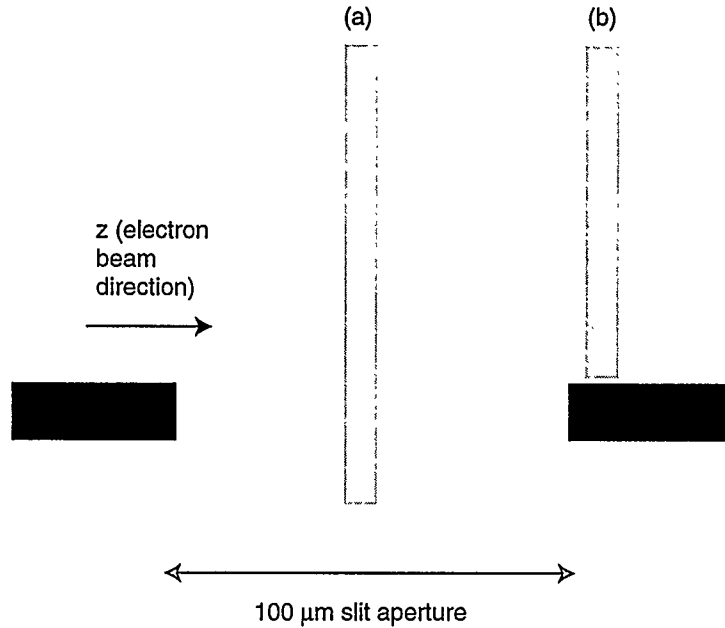
Figure 5.4: Longitudinal alignment of the Laser-Compton BSM. A slit is inserted into the path of the lasers, which prevents the laser from passing through if it is more than 50 µm from the slit center (case (b)), but permits it to pass through and strike a photodiode otherwise (case (a)). This constrains all laser pathways to meet at a common z location.

at the PMT.

Longitudinal alignment of the laser pathways is done initially without the electron beam. A slit is introduced into the center of the FP chamber (Figure 5.4); the laser longitudinal motors are stepped, and the laser intensity in a photodiode across the cavity from the mirror is read out. The slit width is 100 µm, and the photodiode intensity is nearly constant as long as the laser beam is within the aperture; the center of the flat-top distribution is determined manually. Each of the laser pathways is aligned in turn in this fashion, and the slit is removed.

Note that the longitudinal alignment described above does not guarantee that the

beams are longitudinally aligned at the location of the electron beam. A second pass of longitudinal alignment is required, in which one of the laser pathways is moved, and the beam size measured as a function of the waist knob. The minimum spot size from this procedure occurs when the longitudinal overlap of the laser pathways is maximized.

One important factor in the longitudinal alignment is that poor longitudinal overlap does not reduce the average signal seen by the PMT; rather, the sinusoidal modulation is reduced because the interference pattern is weak or missing. Therefore, such a misalignment will result in systematically enlarged spot size measurements.

Tuning the Nd:YAG Laser:

The primary features of the laser oscillator itself which require tuning are the power output, timing, and spatial and temporal coherence. These are primarily tuned in the laser shack itself, with the exception of the timing of the laser pulse with respect to the electron beam. The relative timing is done by scanning the laser trigger time and maximizing the PMT signal from the electron beam-laser interaction. While reduced laser power output results in poor signal strength, spatial and temporal coherence of the laser result in systematic errors which are discussed in Section 5.2.2.

Tuning the Backgrounds:

The strong dependence of the Laser-Compton BSM upon the maximum and minimum heights of a sinusoid implies that backgrounds, which will tend to "fill in" the minimum, are to be avoided at all costs. This is done by adjusting the collimators and the extraction orbit of the beam to minimize the signal seen on the PMT GADC (Gated Analog-to-Digital Converter) when the laser is absent. In the May 1994 FFTB run, a major improvement was made by designing a "low-noise" extraction optics, in which vertically-focusing QP1 and QP2 quads were run at full strength to reduce the size of the electron beam in the other extraction quads. In practice, it is possible to reduce the background to as few as 20 counts, although 35 is more common. The average Compton signal seen during the September 1994 FFTB run, with all systems
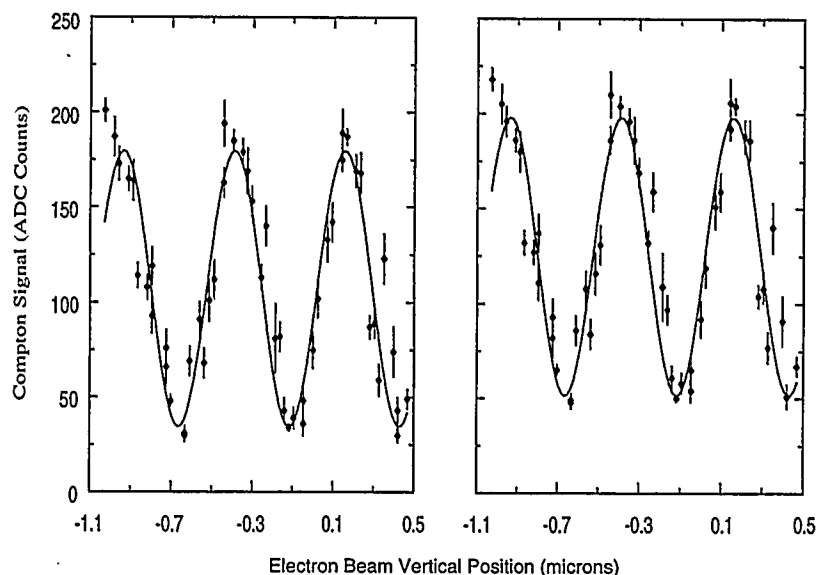
Figure 5.5: Laser-Compton beam size measurements with (left) and without (right) pulse-to-pulse background subtraction. Note that the pattern on the right is systematically higher.

tuned up and a near-maximal electron population in each bunch, was in the vicinity of 100 counts, with 200 counts possible and expected during later runs. Clearly, a background of 35 is far too large without some correction scheme.

Currently the Laser-Compton BSM operates with a background subtraction algorithm. While the electron beam operates at 30 Hz, the laser only fires at 10 Hz. Each pulse with laser present is preceded and followed by a pulse with laser absent. The average of the preceding and following GADC values is subtracted from the with-laser GADC value; 6 repetitions of this pattern (6 pulses with laser and 12 without) are averaged to comprise one point in the BSM scan. Figure 5.5 shows one Laser-Compton BSM scan, both with and without this background subtraction. The error-bars on individual points are given by the RMS of the 6 pulse-patterns averaged to form the data point, divided by $\sqrt{6}$. The unsubtracted scan has a fitted size of 88 nm, while
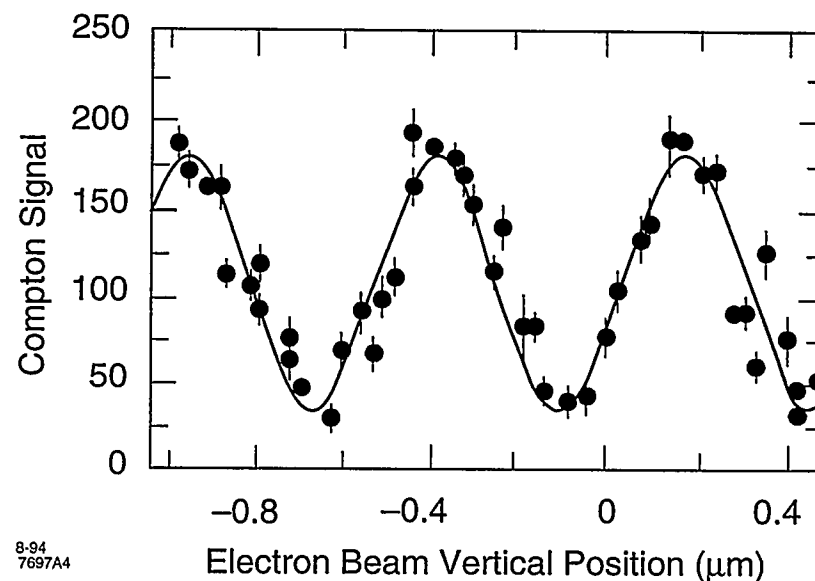
Figure 5.6: Laser-Compton beam size measurement performed in May of 1994. The measured size is 77±7 nanometers.

the background-subtracted scan has a fitted size of 75 nm.

It is worth noting that while the same PMT is used for both signal and background detection, the two inputs reach the computer through different GADC's. A systematic difference between the two is possible.

Result of Laser-Compton BSM Tuning

Figure 5.6 shows a 77 nm beam spot measured in the KEK BSM in May of 1994. Figure 5.7 shows a histogram of the size measurements made over the course of several hours during that run. The average measurement is 77 nm, with an RMS width of 7 nm. For the laser intensity available at the time, the 10% width is consistent with the measurement uncertainty of the individual measurements.

The 77 nm beam size is known to be enlarged by 10% over the actual beam size
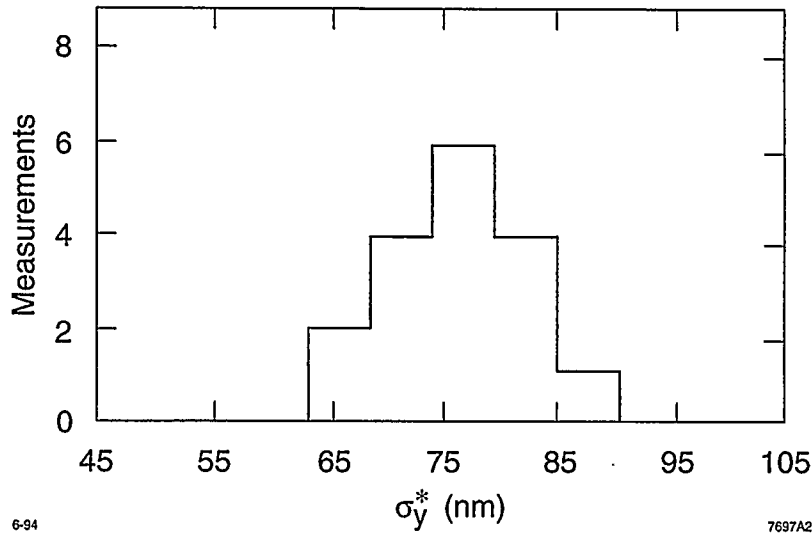
Figure 5.7: Histogram of measurements made during the last 3 hours of the May, 1994 FFTB run. Average size measured was 77 nm, with an RMS of 7 nm.

by a systematic error related to the longitudinal size of the laser pattern. The RMS size of the laser beam at the FP was known to be 50-60 μm from scans similar to Figure 5.2. Because the electron beam has a β_y^* of 100 μm, the laser interference pattern is long enough in space to sample the beam at locations where the beam is not in focus. These tend to systematically enlarge the measured size, and the 60 μm laser measurement was found to correspond to a 10% enlargement. The beam size in May of 1994 is therefore believed to have been reduced to 70 nm. This systematic was another motivating factor in retuning the laser to a smaller cross-section at the FP, since the current 25 μm pattern causes an enlargement of less than 1%.

The 70 nm vertical size was reacquired in September of 1994; however, at no time did any measurement show any sign that the beam had been reduced to a significantly smaller size. At the time, the emittance was averaging approximately 2×10^{-11} m·rad, and the RMS energy spread was 5×10^{-4}; the linear beam size $\sqrt{\epsilon\beta^*}$ expected for

$\beta_y^* = 100\mu$m was 45 nanometers. Contributions from uncancelled aberrations have not been estimated for these parameters; however, the below-design emittance and energy spread both imply that the total blow-up will be less than the 10% predicted for the design parameters; therefore, a beam size under 50 nanometers should have been possible. The most likely explanations of this discrepancy are errors in the beam optics and/or tuning, and systematic errors in the measurement process. Possible contributions to each of these are discussed below. Note that, with the exception of the background subtraction GADC issue mentioned above, all systematic errors lead to an overestimate of the beam size.

Possible Tuning Errors in the Beamline

β_y^* Tuning Errors:

The most straightforward tuning error imaginable is simply mis-setting the vertical betatron function at the FP. When ϵ_y is as small as 2×10^{-11} m·rad, the vertical beam size is too small to measure on WS3; and while WS2 can measure the WS3 divergence, this measurement requires several wires and is quite time-consuming, and thus is not often done. It is thus not inconceivable that the linear beam size is set incorrectly at the FP. Furthermore, larger values of β^* correspond to smaller FP angular divergences, which in turn correspond to smaller beam sizes in the FFTB apertures. Therefore, background tuning will tend to "favor" larger β^* values, creating an operational tendency to raise β^*, or at least to not reduce it, when backgrounds are intractable.

The FP divergence monitor, WS6B, is generally capable of yielding an aberration-free measurement of the value of β_y^*, presuming that the vertical emittance measured at WS1 and/or Sector 28 of the linac is to be trusted:

$$\sigma_y(WS6B) = \sigma_{y'}(FP)L(FP \rightarrow WS6B) = \sqrt{\frac{\epsilon_y}{\beta_y^*}}L(FP \rightarrow WS6B), \qquad (5.10)$$

where Equation 5.10 assumes that the distance from the waist to the beam size monitor is small compared to the distance from the monitor to WS6B (which in

this case is 75 cm), and that the beam size at the monitor is small enough that $\sigma_y^* R_{33}(FP \rightarrow WS6B)$ is negligible.

The Laser-Compton BSM was well-tuned and measuring reproducible 70 nm spot sizes during a period from September 14 through September 16, 1994. In this same period the Sector 28 vertical emittance $\gamma\epsilon_y$ was oscillating about an average value of 2×10^{-6} m·rad. Seven measurements of $\sigma_y(WS6B)$ made during this period give a weighted average of 362 μrad, with an RMS of 32 μrad. Assuming that the projected emittance at the FFTB FP is given approximately by scaling the Sector 28 normalized emittance by the appropriate relativistic factor, the divergence of 362 μrad corresponds to a β_y^* of 167 μm. Such a large β_y^* would result in a linear beam size of 60 nm. The WS6B measurement is also affected by systematics, principally a possible reduction of the measured beam size due to limited acceptance of the photon extraction line downstream of the FFTB; the FFTB's measured value of ϵ_y is typically smaller than the Sector 28 number would indicate, by up to 15%. Thus the 167 μm β_y^* represents an upper bound, but one which is not likely to be wrong by more than 25%. .

Additional software tools have been added to the SLC control system to allow faster and more efficient evaluation of the angular divergence measurements at WS2 and WS6B. These should reduce the risk of inadvertant mis-tuning of the betatron function.

Geometric Sextupole Errors:

Section 2.4.1 mentions that the tolerances on the allowed sextupole content of the final doublet quads are quite tight, while the tolerances on the remaining FFTB quads are also strenuous. In bench measurements, the FFTB standard quads were found to contain an RMS sextupole content approximately twice the allowed value (see Appendix B). The doublet quads were not machined to meet this tolerance [14], but were instead machined to meet the looser tolerance on octupole content; the sextupole content was then measured and eliminated using a set of trim windings

on each magnet. The measured sextupole contents of the doublet magnets are large enough to increase the FP beam size from 46 nm to almost 70 nm when the emittances and betatron functions are set as they were in September of 1994. Because the doublet sextupole measurements were equivocal [59], the decision was made to include the FT sextupoles and eliminate sextupole content globally by scanning the sextupole aberration knobs, rather than using the doublet trim winding scheme.

The FT sextupole magnets made a measurable improvement in the beam size measured at the Laser-Compton BSM in May of 1994; their effect was more equivocal in September of 1994. Shortly after this it was determined that the alignment tolerances of the FT sextupole magnets are quite tight: while the SK2/SX2 magnets have tolerances of 200 μm in x and in y (defined by the misalignments which will cause 2% spot dilution due to waist or coupling at the full sextupole strength), the SX1 horizontal and SK1 vertical tolerances are 25 μm.

An experiment in January of 1995 (see Section 5.3.1) determined that the sextupoles were misaligned, by up to 1 mm in some cases. While the gross errors were then corrected by mechanical realignment, it is not practical to attempt to align the sextupoles down to their tolerances because these magnets are not supported on remotely-controlled movers. In addition, the sextupole content of the beamline is expected to be quasi-static, and therefore the sextupoles will not need to be quickly and efficiently scanned repeatedly. In future runs a "2.5-dimensional" scan will be used: the current sextupole knob will be set to a value, and each of the waist and coupling will be optimized; this will be repeated at each setting of the knob, and the best value determined from this set of scans. Once set, the knob is unlikely to require scanning again.

Possible Systematic Errors

Spatial and Temporal Coherence of the Laser:

The phase coherence of the Nd:YAG laser source in both position and time is
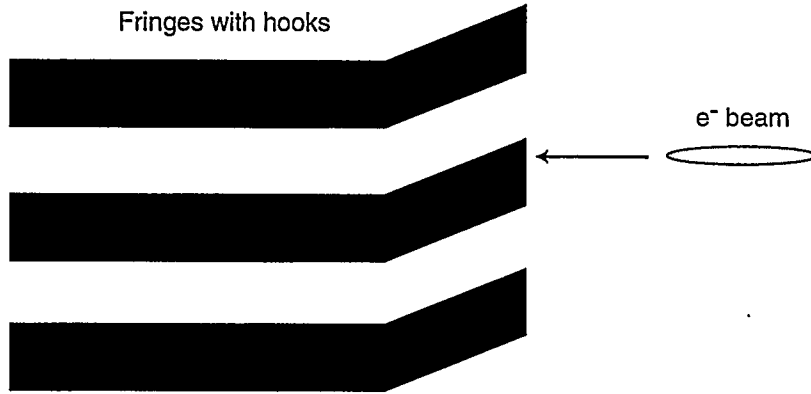
Figure 5.8: Poor spatial coherence of the Nd:YAG laser results in non-planar interference fringes, as shown. The electron beam can no longer pass through a well-defined "null", resulting in an enlarged beam size measurement.

critical to constructing and maintaining stable interference patterns. Temporal incoherence would result in the interference pattern "jittering" in the path of the electron beam. Spatial incoherence would result in fringes which are not planar: the incoming laser beam is no longer a pure plane wave, but rather a set of plane waves with phase differences across the area of the wavefront. A minor case of spatial incoherence produces fringes with "hooks" as shown in Figure 5.8. This effect fills in the minima in the fringe pattern, again resulting in larger measured spot sizes.

The spatial and temporal coherence of the laser oscillator were measured in the September 1994 FFTB run. The temporal coherence was found to be quite good, while the spatial coherence was poor [60]. This situation was corrected, resulting in instant reduction in the measured spot size. Laser coherence is now routinely measured prior to FFTB operations.

Longitudinal Size of the Laser Spot:

As mentioned above, the longitudinal extent of the interference pattern can systematically enlarge the measured spot. This is because, with a β_y^* of 100 μm, a large interference pattern will produce Compton scattering at locations where the beam is
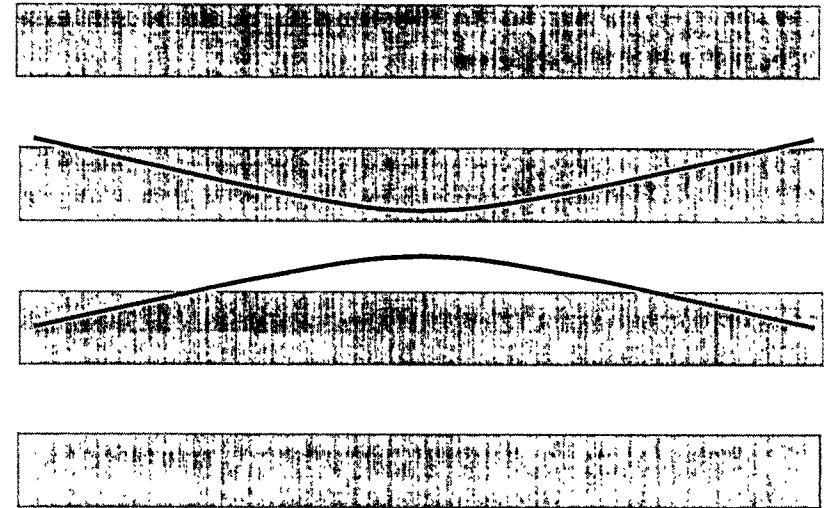
Figure 5.9: Beam envelope and interference pattern when $\beta_y^* < \sigma_z$ of the laser. Scattering at points distant from the FP will systematically enlarge the measured beam size.

larger than the focused size, as shown in Figure 5.9. Furthermore, the length of the laser pulse (in time) is enormous compared to the length of the electron beam (in time). Therefore the probability of scattering off the wings of the laser beam, where the electron beam is large, is equal to the probability of scattering off the core of the laser beam, where the electron beam is small.

During the May 1994 FFTB run, the size of the interference pattern was approximately 60 μm, resulting in a 10% increase in the measured size of the electron beam. During the September 1994 run, the pattern size was reduced to 25 μm, which reduced this effect to a 1% correction. Note that there is a trade-off between a large laser spot (which produces a systematic dilution of the measurement), and a small laser spot (which requires more frequent alignment).

Longitudinal Overlap of Laser Pathways:

The effect of longitudinal misalignment of the laser pathways was mentioned previously in this section. Because the electron beam is still intercepting the same number

of photons as in the fully-overlapped case, the average Compton signal is not reduced; however, the nulling of the photon intensity at the troughs of the interference pattern is weakened. A worst-case of this situation is one in which the laser pathways do not overlap longitudinally at all at the location of the electron beam, resulting in a total absence of the sinusoidal interference pattern. This loss of alignment results in an increase in the measured beam size.

In practice, this source of systematic error has been quite significant, especially in measurements made during September of 1994, when changes in ambient temperature drove diurnal variations in the laser transport line from the source to the Laser-Compton BSM housing in the tunnel. The reduced transverse size of the laser at the FP served to exacerbate this problem. While it is possible to correct this problem during tuning by frequently optimizing the longitudinal overlap, it is crucial that this additional tuning item be executed faithfully.

Rotations of the BSM housing:

Two significant rotations of the housing of the Laser-Compton BSM are possible: a *roll*, resulting in xy mixing; and a *pitch*, resulting in the electron beam path leaving the plane of the fringes.

The case of a roll results in a straightforward enlargement of the measured beam size by $\sigma_x^* \sin\theta$, where θ is the relative rotation angle. The horizontal beam size of 1.7 - 2.0 μm, and a vertical size of 70 nm where 45 nm was the best possible, results in an upper bound of 32 mrad for θ. Experience with the tight rotation tolerance of the WS3 yoke (Section 3.3) suggests that installation tolerances which are a fraction of 32 mrad can be met; therefore, this seems like an unlikely candidate. A multiknob which uses QSM2 and QS3 to roll the FP spot has been prepared, and will be used in future runs to measure this effect. A more reasonable estimate of 3 mrad installation error for roll results in a contribution of 5 nm added in quadrature, which is negligible.

Figure 5.10 illustrates the effect of pitching the beam relative to the plane of the interference pattern. In this case, scanning a perfectly flat electron beam across
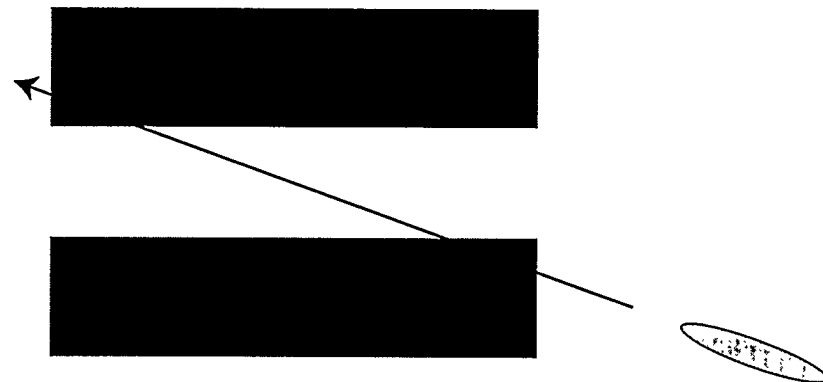
Figure 5.10: When the electron beam trajectory does not lie in the plane of the interference fringes, a beam which nominally passes through the null of the pattern intersects the higher-intensity regions at the longitudinal extremes of the laser. This results in higher scattering rates at the "nulls" and lower rates at the "peaks", and increases the measured spot size.

the pattern will no longer produce a full-modulation Compton signature, and the measured beam size of such a beam will no longer be identically zero as a result. The measured beam size in this case is related to the actual vertical size by:

$$\sigma_{meas}^2 = \sigma_{real}^2 + \phi^2 \sigma_{z,laser}^2, \tag{5.11}$$

where ϕ represents the relative angle between the beam trajectory and the plane of the fringes. For $\sigma_{real} = 40$ nm and $\sigma_{z,laser} = 25$ μm, a pitch angle of 500 μrad results in a 5% correction to the beam size. Such a relative angle is not inconceivable.

Typically, the beam size seen by the BSM is measured as a function of FP angle during the course of FFTB operations. No meaningful correlations have been seen during these scans, indicating that the beam and the fringes were probably coplanar at the level of 500 μrad.

Beam Jitter and Magnet Vibration

Pulse-to-pulse jitter in the incoming beam will result in consecutive pulses not passing through the interference pattern at the same location. This in turn will

result in a variation of the signal at the PMT. While jitter at the midpoint between a peak and a trough of the interference pattern will average out to the jitter-free signal level, jitter at the peak will reduce the measured peak intensity (which is obtained by averaging 6 pulses), and will increase the measured trough intensity by a similar amount. This will increase the beam size at the FP. The increase in the beam size can be determined by adding the pulse-to-pulse jitter amplitude in quadrature with the actual spot size.

K. Oide has pointed out [61] that while the FP spot size averages over jitter in this fashion, the incoming emittance is also measured over a large number of pulses; since the measured size on a wire scanner is increased by jitter in a mathematically-identical fashion to the Laser-Compton BSM, the emittance measurement should already include the effect of the beam jitter. While the wire scanner software is capable of eliminating incoming jitter pulse-by-pulse, the software is not sufficiently sophisticated to do so correctly during a Quad Emit scan, when the optics through the FFTB are changing due to quad scans. At this time it is impossible to eliminate jitter from the Laser- Compton BSM's measurement, as there are no BPMs with sufficient resolution at the correct phase. Therefore the emittance measurements made at WS1 should also be configured to leave residual jitter in the measurement for a more accurate comparison.

While the incoming beam jitter contributes to both the WS1 measured emittance and the BSM measured spot size, only the BSM sees the effects of jitter introduced by vibrations of the FFTB magnets. Therefore, any jitter coming from the FFTB's own vibrational properties will cause a spot size increase which is not consistent with the WS1 emittance value. The Laser-Compton BSM is attached securely to the doublet table, in the expectation that the monitor and the doublet quads will oscillate coherently and the effects of vibration will cancel out. While the expectation is certainly correct at the level of microns, the nanometer level is less certain. Measurements of the vibrational properties of the doublet/table/BSM system are ongoing; in addition,

a high-resolution RF BPM has been installed in the Laser-Compton BSM housing, in an effort to directly measure the relative vibration between the incoming beam and the BSM.

5.2.3 Gas-Ion Beam Size Monitor

In order to measure the beam size via the Gas-Ion Beam Size Monitor, several aspects of the monitor must be properly tuned first. These aspects are: the microchannel plate (MCP) high voltage; the injected gas pressure; the longitudinal slit which defines the device's longitudinal acceptance; and the beam-induced backgrounds.

MCP High Voltage:

The relative sensitivity of the MCP ion detectors is a nonlinear function of the applied voltage. Therefore, increasing the voltage increases the percentage of ions detected, which in turn improves the statistical precision of all measurements. However, the MCPs also detect particulate backgrounds which are created by the beam; the signals from such particles scale identically to the ion signals. As the high voltage increases, these backgrounds become amplified, and the danger that a transient "spike" in the backgrounds can damage the MCP amplifiers also increases.

In general practice, the tuning of the backgrounds (see below) and the tuning of the high voltages is performed as a cycle: the voltages are increased until the backgrounds induce a certain number of counts in the MCPs; at that time the backgrounds are reduced until the number of background events is too small to accurately measure, and the high voltage is then increased again until the background counts reach the maximum tolerance. Once the high voltage is set above 1900 volts, measurements can generally be performed.

Injected Gas Pressure:

The appropriate quantity of gas to be injected for each pulse of beam is a function of the beam size, particularly the vertical size. A relatively high pressure of gas increases the number of ions, and therefore the counting rate improves. For measuring

the size of flat beams, however, the relative ion counting rates around the azimuth of the device must be measured with minimal systematic skewing between horizontal and vertical MCPs. During flat beam measurements, it was observed that the beam size was dependent upon the pressure of injected gas, with lower pressures reducing the measured vertical beam size.

The pressure-dependent beam size phenomenon was explained by ion"pile-up:" the number of ions incident upon the horizontally-mounted MCPs exceeds their maximum counting rate, and therefore the number of ions in these MCPs is reduced, while no such reduction occurs in other MCPs which have fewer ions incident upon them. This reduces the measured ratio between the number of horizontally-kicked ions and the number of vertically-kicked ions, resulting in a larger measured beam size. Note that reducing the gas pressure reduces the number of counts on each electron beam crossing, which increases the number of electron beam pulses needed to make a measurement with the desired (10%) precision.

Longitudinal Acceptance:

The ionization and acceleration effects measured by the Gas-Ion BSM occur at all points at which the electron beam and injected gas interact; in particular, the beam ionizes and accelerates helium atoms both at the focus and at short distances away from the focus. The ions from the focal point, where the beam size is at its smallest, are therefore counted along with ions produced where the beam has diverged and is no longer as small. These latter ions tend to dilute the measurement of the small spot. This problem is exactly analogous with the difficulty encountered in the Laser-Compton BSM when the laser longitudinal size is large compared to the vertical betatron function (see previous Section).

In the Gas-Ion BSM, the problem is resolved by placing a narrow slit between the beam-gas interaction point and the MCPs. The slit aperture can be remotely enlarged or contracted, in particular it can be reduced down to a total size of 200 μm. The vertical betatron function is on the order of 100 μm. For the very smallest

beams, therefore, some residual systematic dilution can be expected, due to the large acceptance of the slit.

Early tests revealed an installation defect with the longitudinal slit: the slit was installed with an angle relative to the beam trajectory. This resulted in unequal longitudinal acceptance on the North and South sides of the BSM. In particular, at a setting of 200 μm, the ions on one side of the device were almost completely excluded. This defect was corrected mechanically, and tested by measuring the intensity of a fluorescent light mounted inside the BSM as a function of azimuth angle, for slit apertures down to the minimum.

Backgrounds:

Because the Gas-Ion BSM does not require detectors which are, essentially, directly in the path of the electron beam, background tuning for this device is less difficult than for the Laser-Compton BSM and WS6A. The large scattering-angle required for a beam particle to enter the MCPs directly generally indicates that only electrons and secondaries far from the beam core will be relevant, and backgrounds incurred in the extraction line can be ignored. In addition, the upstream face of the BSM is shielded locally with approximately 1 inch of lead. Nonetheless, the background counting rates of the Gas-Ion BSM (rates measured with no gas injection) are like other backgrounds around the FFTB in that a "happy medium" of collimator settings is required: far enough in to eliminate long beam tails while not so far into the beam pipe that the core begins to scrape on the collimators and repopulate the tails.

Result of Gas-Ion BSM tuning

Figure 5.11 shows the azimuthal distribution of ions obtained during a beam size measurement made in the September 1994 FFTB experiments. Coupled with the time of flight measurement (not shown), the estimated beam size is 1.6 μm in the horizontal and 80 nm in the vertical. The estimated total error (systematic and

7952A1
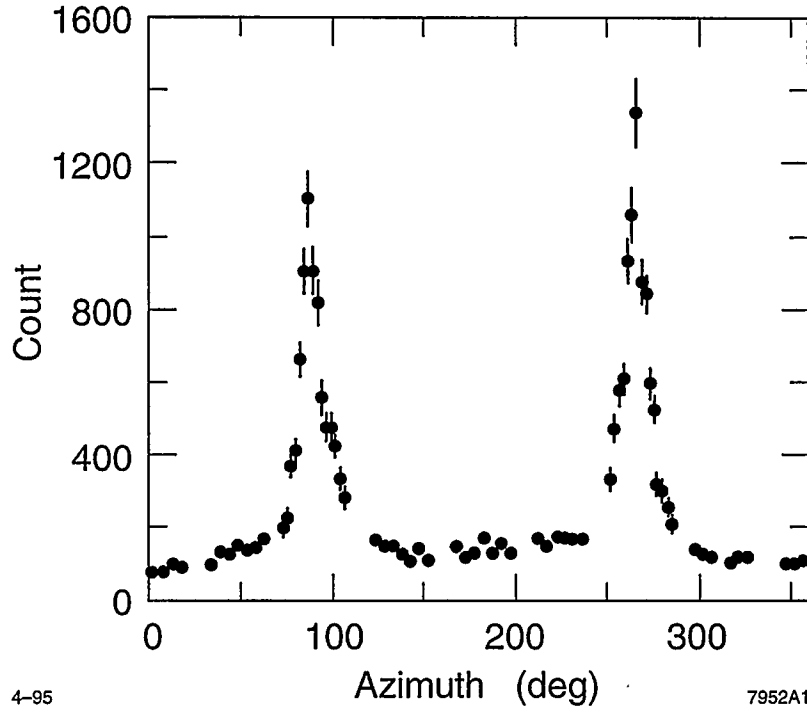
Figure 5.11: Azimuthal distribution of detected ions at Gas-Ion BSM. The amplitude of the "twin-peak" anisotropy coupled with the time-of-flight measurement (not shown) yield a beam size of 1.6 μm by 80 nm.

statistical) on the measurement is 0.1 μm (x) and 25 nm (y). The error on the horizontal beam size is dominated by statistical error, while the error in the vertical is dominated by uncertainties in the correction for ion "pile-up." In future runs, the gas pressure will be reduced to relieve this source of systematic error.

Note that the measured beam size above is larger than the best vertical beam size measured at the Laser-Compton BSM, by about 10%. However, in order to move the vertical waist downstream by 52 centimeters, it proved necessary to reduce the overall demagnification of the Final Transformer by that same factor, such that the expected

value of β_y^* is now 120 μm. When this fact is considered, the beam size measured by the Gas-Ion BSM is completely consistent with that measured with the Laser-Compton BSM. While this improves confidence that the various systematic errors of the two devices are under control and not arbitrarily enlarging the measured beam size, it also strengthens the case for a beam-optics explanation of the unexpectedly large beam size.

Two advantages of the Gas-Ion BSM are worth mentioning here. The first is that the device can measure a relatively large beam size (up to 40 μm has been achieved) with fairly simple adjustment of its parameters. In particular, the time window of the acquisition system is stepped until ions appear: while a large beam will produce only very slow ions, these can nonetheless be detected by forcing the acquisition electronics to delay measurement. The second advantage is that the Gas-Ion BSM has full azimuthal detection of the azimuthal ions, and therefore is able to unambiguously measure the roll angle of a focused beam. In particular, the azimuthal angles of the centers of the "horizontal" peaks in Figure 5.11 determine such a roll angle. On one measurement performed in September of 1994, a roll angle in excess of 4° was seen [62]. Another measurement, made in March of 1995, found that the roll angle was varied when the vertical dispersion was scanned, and that the roll angle was zero when the minor axis of the beam ellipse was minimized. The conclusion drawn at the time was that the horizontal dispersion was not optimized, and that introducing vertical dispersion caused a spurious horizontal-vertical correlation. This is assumed to be the culprit in the September 1994 measurement as well.

Since the beamline tuning errors do not differ between the two BSMs, it is more fruitful to consider at this time the possible systematic errors in the Gas-Ion BSM measurement.

Possible Systematic Errors

Acceptance Anisotropy
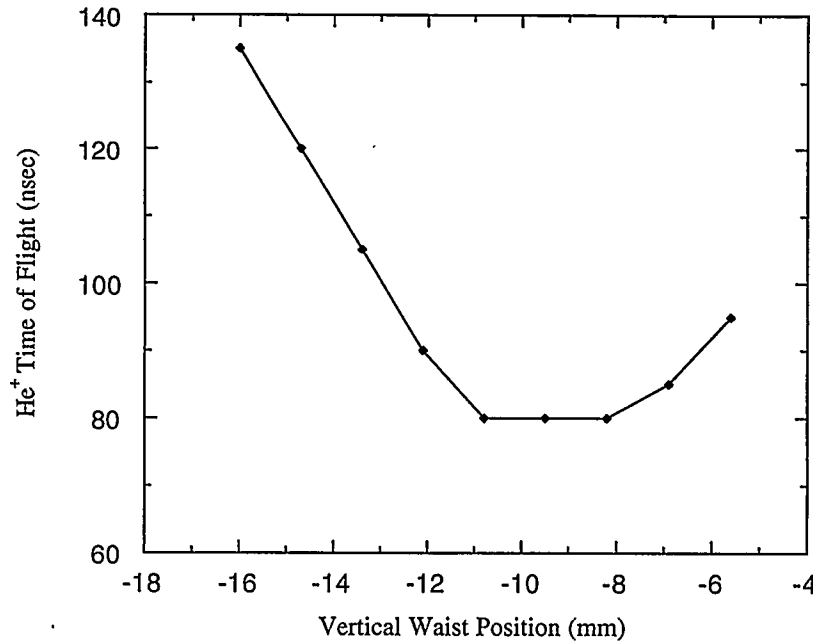
Figure 5.12: Gas-Ion BSM measured time-of-flight versus vertical waist position. In the linear regions of the curve the $\sigma_y^* > \sigma_x^*$ and thus the TOF measures primarily σ_y^*; in the flat section $\sigma_x^* > \sigma_y^*$ and the constant value of σ_x^* dominates the TOF measurement.

The possible effects of an anisotropy in the acceptance of ions was discussed previously in this section. The principal effect is to reduce the number of horizontally-accelerated ions relative to the number of vertically-accelerated ones, increasing the measured spot size. This effect was measured and corrected down to a few percent prior to the September 1994 FFTB run.

Unequal Detector Responses

Unequal calibrations of the MCP and amplifier gains around the azimuth of the BSM would result in distortions in the azimuthal distribution of the ions, and could either increase or decrease the measured beam size. In order to eliminate this source of error, a radioactive source (specifically, Americium) was used to calibrate the system:

a source placed in the center should generate equal numbers of counts in all MCP readouts. (An exception is the horizontal MCP units: in these, the width of the individual readout strips is reduced by half, and the number doubled, to improve resolution of the beam roll angle [63]. In this case, the number of counts should be reduced by exactly 50%.) An additional calibration was made by setting a round beam in the center of the BSM. Because the time of flight corresponds to the size of the beam's major axis, a vertical waist scan against the time of flight will produce a constant time of flight when $\sigma_y \leq \sigma_x$, and the time of flight will increase linearly when $\sigma_y > \sigma_x$. At the turnover between constant TOF and linearly-increasing TOF, the beam sizes in x and y are equal. This situation is shown in Figure 5.12.

Bunch Length and Intensity

Because the Gas-Ion BSM measurement couples to the beam size through the electric field of the bunch, knowledge of the intensity and bunch length are essential in correctly measuring the absolute beam size. A sufficiently low charge density will eliminate the ion trapping which produces the two-peak anisotropy in Figure 5.11, and will increase the time of flight as well, resulting in increases in both measured numbers relative to the actual beam size.

The bunch intensity is measured by a series of toroids in the linac, with an expected accuracy of 2% at FFTB currents. The bunch length is measured by activating the SLC main bend, 50B1, bending the beam into the SLC arcs, and reflecting the resultant synchrotron light into a streak camera. Such a measurement cannot be made during FFTB running; it is necessary to tune up FFTB beam conditions, make the bunch length measurement, and assume that the conditions do not change overmuch during the FFTB run. Such a measurement was made in March of 1995, and showed an RMS bunch length of 600 μm, with precision of approximately 15%. It is therefore expected that charge-density effects could result in as much as 1% systematic error in the horizontal measurement.

Non-Gaussian Beams
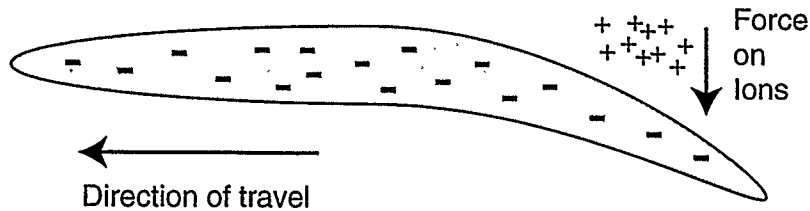
Figure 5.13: Head-tail effect in Gas-Ion BSM. The positive ions produced by the head and core of the beam are attracted to the negative charge in the offset tail; this results in a change in the azimuthal distribution of ions at the MCPs.

By far the most significant measurement difficulty observed with the Gas-Ion BSM is non-Gaussian distributions in the beam. In particular, a head-tail effect of the type shown in Figure 5.13, in which the tail has some transverse deflection relative to the head, can produce a serious error in measurement. This is because the tail tends to attract ions which form when the head passes through the injected helium. A vertically-deflected tail will therefore pull all the ions in the vertical direction, causing the horizontal peaks of Figure 5.11 to appear with a separation different from 180°; while a horizontal tail will enhance one of the peaks and reduce the other. Simulation studies show that the sum of the two peak amplitudes is also reduced; therefore, such tails introduce an increase in the measured spot size which cannot easily be corrected to yield the correct value [64].

In actual operation, it is often possible to correct the head-tail effect by steering the beam in the linac. Offsetting the electron beam in the disc-loaded wave guide can produce a head-tail effect which cancels the effect seen at the BSM. Another tuning technique which has some effect in this case is changing the relative phase of the bunch and the linac RF, the "phase-ramp". Changing the setting of "phase-ramp" works to introduce a head-tail energy correlation; this can cause the tail to be overbent and/or over-focused relative to the head, which can then have the effect of straightening out the bunch (BNS damping works in a similar fashion). Neither of these techniques is guaranteed to work, however. In particular, the stability of the

linac is a key to their success or failure, and the probability of success for tuning such head-tail effects is known to be higher during the night than during the day (due, presumably, to thermal changes in the linac itself).

5.3 Other FP Tuning Issues

In addition to the phenomena described in the preceding sections, a handful of other issues relating to global spot-size tuning have been observed and/or measured. These include the alignment of the FT sextupole magnets, and the doublet settings required to set the waist at each of the BSMs.

5.3.1 FT Sextupole Alignment

The sextupoles in the FT are only 3% as strong as the CCS sextupoles, and do not have magnet movers. Therefore, the technique described in Chapter 4 for CCS sextupole alignment is not applicable. Instead, a variation of the SLC technique [53] is used, in which the sextupole strength is changed and the movement of the waists and change in coupling at the FP are measured by scanning the appropriate knob. In this case, each of the sextupoles is powered independently, which is an advantage over the SLC FF technique.

The sextupole alignments were measured during January of 1995, using the WS6A wires. In order to perform the measurement, it was therefore necessary to maintain a large amount of x'y coupling in order to prevent wire breakage. When the x'y coupling knob was scanned, a step size was selected which would reduce the likelihood of breaking a wire. In any event, no wire was broken. The waist and coupling measurements were performed at 5 values of the sextupole strength. In order to estimate the amount of natural drift of the waists and coupling, the waists and coupling were scanned before the sextupole was first stepped, and then again after it was set back to zero strength. The RMS drift of each knob, over the four hours of the experiment,
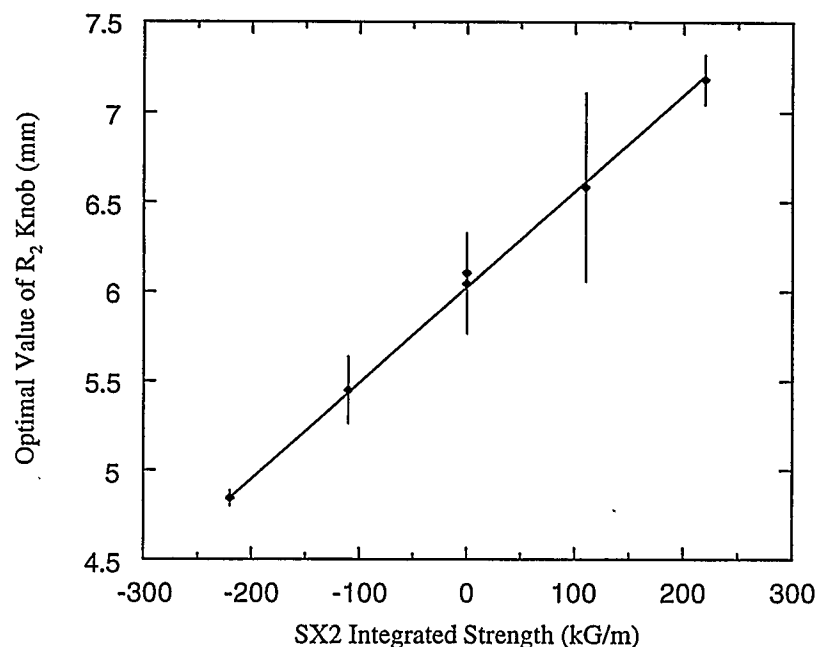
Figure 5.14: Optimal value of the x'y coupling knob versus SX2 strength. The linear correlation indicates a vertical misalignment of the sextupole, of approximately 1 mm.

was added in quadrature to the statistical uncertainty of the optimal knob setting. Knob scans were performed using the sextupole mover knobs.

Both the horizontal and vertical waist scans coupled to horizontal misalignment of the normal sextupoles and vertical misalignment of the skew sextupoles (i.e., normal quadrupole aberration). While the two measurements agreed within their errors, the vertical measurement was generally of too low a precision to use. Misalignments were therefore determined by horizontal waist measurements using the horizontal wire, and coupling measurements using the vertical wire.

Figure 5.14 shows the optimal knob setting versus sextupole strength for the best correlation observed in this procedure, specifically SX2 strength versus coupling. The

slope of the line indicates a vertical misalignment of over 1 mm, with a statistical error of approximately 50 μm. Several misalignments between 500 μm and 1 mm were observed, and statistical errors ranged from 50 to 150 μm. The misalignments over 500 μm were corrected mechanically. It is, however, expected that the operation of beam-based alignment will cause systematic shifts in the beamline larger than the alignment tolerances of the small sextupoles (tolerances are as low as 25 μm). Therefore, the multi-dimensional scan described in Section 5.2.2 will be used in the future.

5.3.2 Reproducibility of Final Doublet Optics

The excitation of the final doublet magnets is the most crucial factor in positioning the horizontal and vertical waists. It has been our unfortunate experience that the exact excitations which provide the desired positionings do not reproduce at all well. The design value for QC1's integrated gradient for Laser-Compton BSM operation is 1,848 kilogauss; during FFTB runs, the actual value used has varied from 1,700 to 1,950 kilogauss. This range of values corresponds to over 10 cm of waist motion (as the doublet waist knob uses approximately 20 kG to move the waist 1 cm). Such large shifts in the required doublet strength cannot be explained by mismatch of the incoming beam, or by CCS sextupole positioning. At this time, the only explanation for the strange variability of the doublet is that these magnets are extremely sensitive to hysteresis effects; because the precise history of scans and optics changes is not the same from run to run, the doublet value needed changes from run to run. This is a problem which causes some annoyance; however, simulation studies reveal that scale-factor errors in the doublet quads cause principally waist shifts, which can be corrected easily. It is worth noting, however, that magnets with such sensitivity to excitation history may also experience higher multipole aberrations due to differential sensitivities in the different pole-pieces.

Chapter 6

System Performance and Chromatic Considerations

One of the crucial design issues for a linear collider final focus is its tunability. Specifically, the final focus system must converge upon an optics suitable for delivering luminosity in a length of time which is small compared to the length of the luminosity run, and also small compared to the length of time needed for the various drifts in the system (magnet strength and placement, etc.) to drive the system away from its high-luminosity optics. If either of these conditions are not met, the collider in question will spend most of its operational lifetime struggling to achieve luminosity and almost none will be delivered to the experiment. This is true for any collider; the extremely tight tolerances on emittance and aberrations in a linear collider, and the absence of equilibrium states which are present in storage rings, imply that the linear collider will have a more complex tuning sequence which must be repeated, in part or in full, more often. Therefore, the tunability requirements on linear collider subsystems are significantly more demanding than those on a storage ring.

Another significant design issue for a linear collider final focus is the overall chromatic behavior of the system. As shown in Chapter 2, the leading aberration which works to prevent arbitrarily small beam sizes is the chromaticity, and the correc-

tion of this aberration essentially drives the entire design process for such systems. Chromatic aberrations are of greater concern than geometric aberrations for several reasons. First, final-focus correction of chromatic aberrations requires placement of high-order correction magnets in the dispersive regions of the system, and this is a much more difficult endeavor than correction of geometric aberrations with magnets in non-dispersive regions. Second, the longitudinal emittance of the beam emerging from the linac is generally quite difficult to reduce: while all linear collider designs call for transverse emittances which are one or two orders of magnitude smaller than those in the SLC, the IP energy spread is expected to be larger by a factor of two than for the existing SLC. Finally, the chromatic properties of a final focus system are also connected to the luminosity delivered when the energy centroid is marginally off the expected. When one considers the thousands of klystrons expected to provide the energy in most linear collider designs, the need for good performance if a few fail to fire, or the beam is fractionally off in timing, becomes clear.

In the case of the FFTB, measurement of the chromatic properties serves an additional purpose. Chapter 5 showed that the minimum spot size achieved was not the expected spot size. One possible reason for this failure is improper tuning of chromatic aberrations, or presence of unexpected chromatic aberrations. Measurement of the chromatic properties of the FFTB would shed some light on this possibility.

In this chapter, therefore, both the overall tunability and the chromatic properties of the FFTB are considered as indicators of the overall merit of the FFTB design approach.

6.1   Tunability and Convergence Speed

The Final Focus Test Beam has never run for a continual period in excess of 19 days. It is therefore imperative that the total tune-up time from the "cold iron" state to small-spot running be as short as possible. In particular, the days which precede each

run must be maximally utilized for pre-beam checks, and the early run-time must be as thoroughly packed with principal and "parasitic" programs as possible.

The run sequence for pre-operation and the first four days of the September, 1994 run is summarized below. This was the longest FFTB run.

- Verify operational status of all hardware (movers, power supplies, rack electronics, lasers, gas injectors, computers, etc.) to the extent possible without beam (3-7 days). Select/generate configurations for initial operation.

- Establish beam to main dump in low-divergence optics. Use ST62 to establish coarse energy match, steer to dump, establish semi-fine energy match; beam checkout of BPMs, safety devices; establish launch and energy feedback (4 Hours).

- Verify wire scanner functionality, tune linac emittances (4 hours).

- Measure incoming beam matrix (14 hours).

- Local Diagnostics (23 hours).

- Match incoming beam and verify on Beta Exchanger wire scanners; tune backgrounds for Laser-Compton BSM (17 hours).

- Tune horizontal beam size on Laser-Compton BSM (2 hours).

- Tune vertical beam size on Laser-Compton BSM (8.5 hours).

- Commission Gas-Ion BSM and tune beam sizes (16 hours).

The list of activities above comprises approximately 4 days of beam time, at the end of which global tuning has been performed on each of the BSMs. In addition, many of the activities listed above would be further optimized as the procedures were debugged and the FFTB optics more fully understood in an operational sense. Measuring the incoming beam, for example, was performed on 3 occasions during the September 1994 experiments. The first time, shown above, required 14 hours; the second required 5 hours; the third required 1.5 hours. Matching the incoming beam required 17 hours when first attepmted, and later in the run was accomplished in 2 hours. In addition, many activities are not repeated during the FFTB run. For example, beam based alignment and lattice diagnostics are performed only at the beginning of the FFTB run. Experience with recovering small spots after various outages (described below) indicates that this is sufficient.

An additional test of the tunability of the FFTB is the recovery speed from other programs in the FFTB area, or from long outages of the accelerator complex due to other systems. For example, during the 13th and 14th of September, the FFTB was used to perform other experiments and also entered for adjustment of the Laser-Compton BSM. Recovery from tunnel access to tuned horizontal and vertical spot sizes at the Laser-Compton BSM on this occasion required a total of 16 hours.

It must be noted that the initial tune-up of the FFTB is performed with an energy spread $\frac{\sigma_E}{E}$ below 5×10^{-4}, as measured by the ESM in the extraction line. This is an aid in tuning the FFTB which is not likely to be available to future linear colliders. However, the short startup period of 4 days, coupled with the outtage recovery time of 16 hours, indicates that the FFTB can be tuned and retuned in adequately short periods of time.

6.2   Chromatic Studies

In the FFTB, the only chromatic aberrations which may be tuned directly are dispersion and chromaticity in each plane, and horizontal second-order dispersion. These are tuned by steering the beam at the final doublet to introduce dispersion, changing the sextupole strengths to change the chromaticity of the beamline, or moving the CCS sextupoles to introduce first- or second-order dispersion. The first-order dispersion corrections are routinely applied during FP global tuning, while second- order dispersion has not been tuned.

Because the sextupole alignment tolerances for scanning the sextupole strength are tighter than the procedure in Chapter 4 can achieve, scans of the sextupole strengths are not generally performed. Instead, two techniques for more globally measuring the chromatic properties of the FFTB are utilized. The first is a monochromatic technique, in which the energy centroid is moved; the second requires increasing the energy spread of the incoming beam and retuning the FP spot. Each technique is
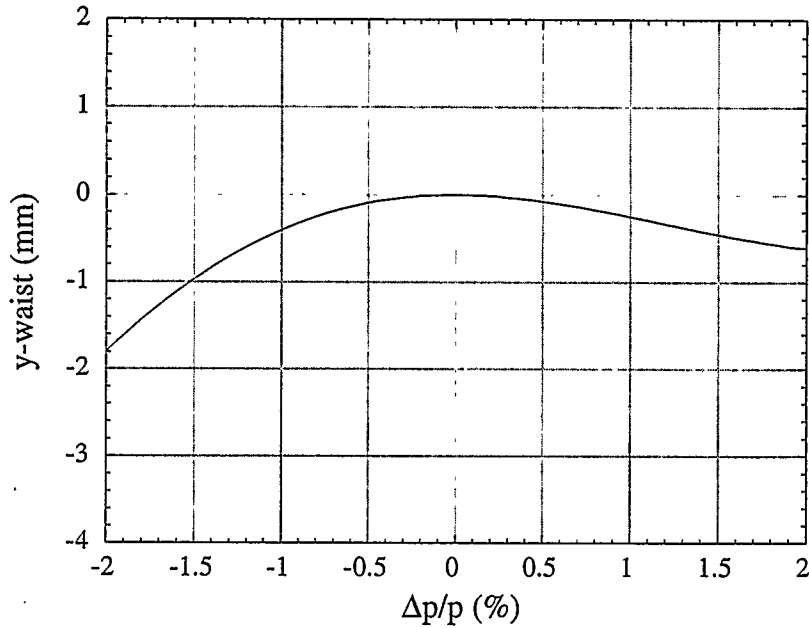
Figure 6.1: Design curve for waist shift with changing centroid energy, for a monochromatic beam with properly-set sextupoles.

described below.

6.2.1 Energy Centroid Scans [64]

Figure 6.1 shows the expected waist position as a function of energy centroid offset for a nearly-monochromatic beam. The waist clearly shifts by an appreciable amount over a small range of energy change: at -1% the vertical waist shift is over 200 μm, twice the betatron function in that plane. It should therefore be possible to measure the vertical waist shift as a function of the energy offset and verify the predicted behavior of Figure 6.1.

Figure 6.2 shows the results of the first such scan in September 1994. The measured data, given by the squares, do not agree with the design curve. A previous
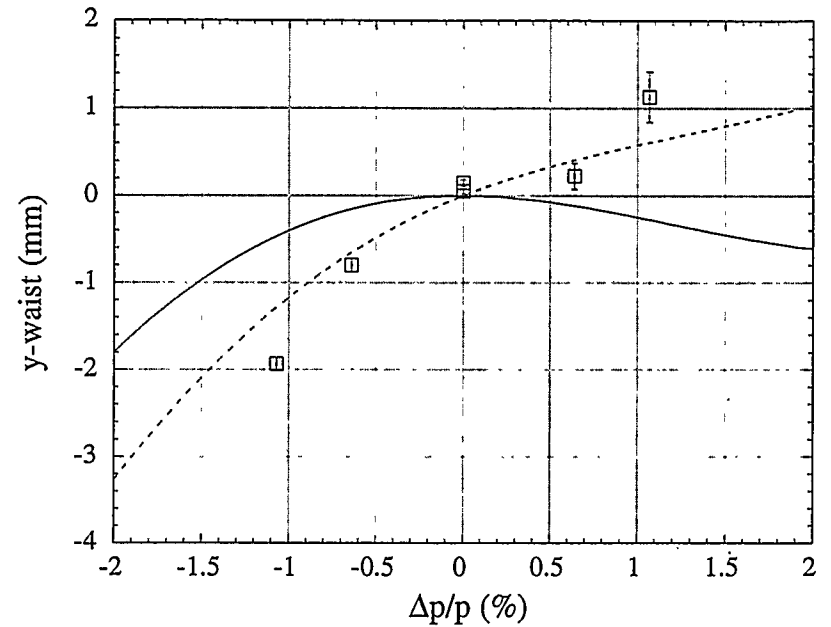
Figure 6.2: Measured waist position as a function of beam centroid energy (boxes). The solid curve is the design prediction; dashed line shows the prediction when the energy centroid error of +0.73% is included.

measurement of the beam energy via the lattice-diagnostics technique described in Chapter 4 had indicated an energy error of +0.73%, which had not yet been corrected. Including the energy offset, but making the assumption that the waist had been adjusted to the FP via the final doublet at the higher energy, yielded the dashed curve. The agreement between the dashed curve and the data was taken as verification of the lattice measurement of energy matching, and the energy was subsequently lowered by 300 MeV.

Once the new energy centroid had been tuned, a second set of data was taken by moving the energy centroid and measuring the waist shift. This data is indicated by crosses in Figure 6.3, along with the first set of data. While the crosses fall closer to the design curve, some discrepancy remains. At this time, it was noted that the SD1
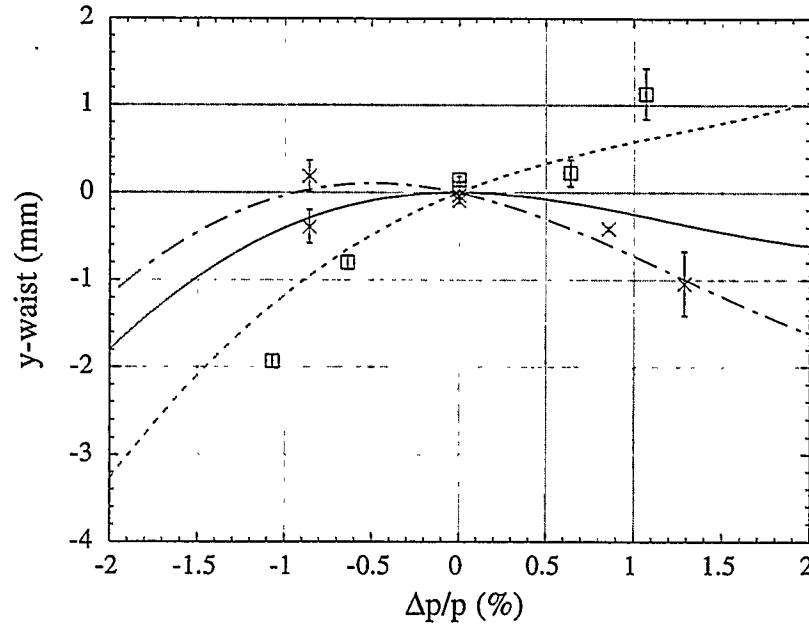
Figure 6.3: Measured waist position as a function of beam centroid after correcting the energy match between the beam and the magnets (crosses); the dot-dash line represents a fit which includes a 3% error in SD1 settings, as described in the text. Previous scan data and fit are also shown for comparison.

sextupoles had been set to a value 3% above design due to an early misinterpretation of the data from Figure 6.2. The dot-dash curve in Figure 6.3 represents the model prediction of such a situation, and this curve fits the data well. Based upon Figure 6.3, the chromatic properties of the FFTB, in the vertical plane, are felt to be well-matched to the expected. Further studies of the chromatic waist shift, with energy and sextupoles properly matched, are a priority for future runs.
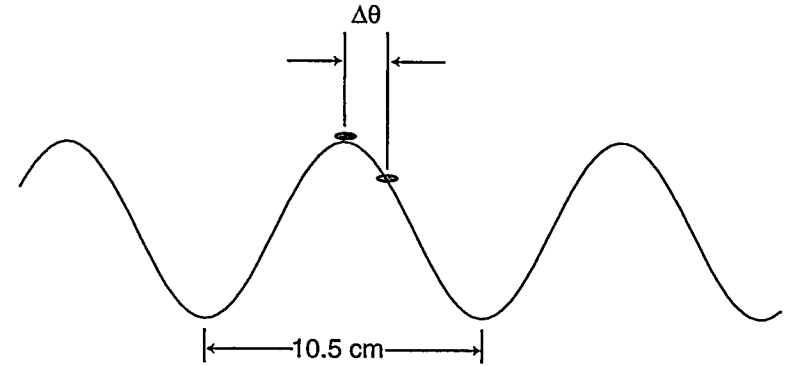
Figure 6.4: Changing the energy spread via phase-ramp. The bunch riding the crest of the RF will have relatively little energy spread because the bunch length σ_z is small relative to the RF wavelength; a bunch which is off crest by $\Delta\theta$ will have a head-tail energy difference due to different gradients at the head and tail.

6.2.2 Increased Energy Spectrum Measurements

A further test of the chromatic properties of the FFTB is to directly increase the energy spread of the beam, and measure the FP spot size. This is a more direct test of the properties as they affect the main parameter of interest, namely the beam size.

In order to tune the energy spread, the extraction-line optics are first retuned to provide a vertical waist at the ESM; once the beam size is measured with these optics, the "low-noise" optics is restored, and the ESM beam size is measured again. This allows the geometric and dispersive contributions to the "low-noise" beam size to be determined; this in turn allows the geometric beam size to be subtracted in quadrature from the ESM measurements, and permits the energy spread to be estimated directly. The "low-noise" optics produced a beam at the ESM with an RMS of 635 μm, while the waist optics yielded a beam with an RMS of 298 μm. This indicates that the narrow energy spread was 5×10^{-4} (RMS), while the monochromatic contribution to the beam size in the "low-noise" optics was 561 μm.

Once the various contributions to the ESM beam size are determined, the energy

spread is increased by changing the relative timing between the linac RF and the beam in the waveguides, via the "phase-ramp" knob. Figure 6.4 shows two bunches in the linac. The bunch riding the crest of the RF will have a center-head and center-tail energy difference due to the head and tail recieving slightly less acceleration than the core; however, the RMS bunch length is 600 μm, while the wavelength of the SLAC linac ($\nu = 2856$ MHz) is 10.5 cm. The energy difference over 600 μm is therefore expected to be only cos(600μm/10.5cm), or 99.998%. The bunch which is off-crest, however, will have a centroid energy which is proportional to $\cos \Delta \theta$. The centroid-tail energy difference can be estimated, in the limit where $\lambda_{RF} \Delta \theta \gg \sigma_z$, as the energy derivative with respect to longitudinal position multiplied by the bunch length:

$$\frac{\Delta_E}{E} \approx 2\pi \frac{\sigma_z}{\lambda_{RF}} \sin \Delta \theta. \tag{6.1}$$

To increase the centroid-tail energy spread by the FFTB design value of 3×10^{-3}, therefore, will require a phase-ramp change of 83.7 mrad, or 4.8°. Note that increasing the energy-spread via phase-ramp is likely to produce transverse head-tail offsets; such offsets increase the complexity of measuring the beam size via the Gas-Ion BSM, which in turn means that this measurement is best made with the Laser-Compton BSM.

Figure 6.5 shows the measured beam spot at the ESM with phase-ramp tuned off its optimum by 5°. The RMS bunch size is 2.26 mm, corresponding to an RMS energy spread of 3.6×10^{-3} once the geometric contribution is subtracted in quadrature and the 60 cm vertical dispersion is divided out. Note that the resultant energy spread is quite non-Gaussian; in fact, it is the twin-horned distribution which is expected to be seen at future linear colliders [1].

Prior to enlarging the energy spread, the weighted average of 11 beam size measurements was 82.3 ± 1.4 nm, with a χ^2/ν of 1.27. At this time, the Sector 28 vertical emittance was measured to be $\gamma \epsilon_y = 2.2 \times 10^{-6}$ meter-radians. After the enlargement, beam size was tuned with the standard set of knobs (waist, dispersion, coupling).

---

[1] Sometimes known as the "Steining Distribution," after R. Steining, or the "Batman Distribution," for obvious reasons.
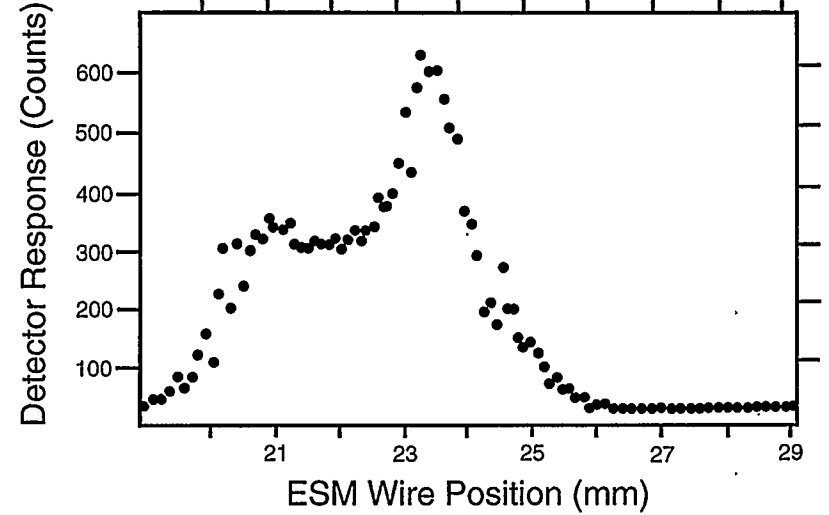
Figure 6.5: Vertical beam profile at the ESM after moving phase-ramp off its optimal setting by 5°.

The weighted average of 7 beam size measurements performed after this tuning was 96.0 ± 2.7 nm, with a χ^2/ν of 1.16. The measured Sector 28 vertical emittance at this time was $\gamma \epsilon_y = 1.8 \times 10^{-6}$ meter-radians. Assuming that these measurements are valid, we can estimate the maximum possible chromatic contribution to the beam size at the lower energy spread. Consider an aberration which scales as σ_δ^n, where n is a positive integer:

$$A_{chrom} = A_0 \sigma_\delta^n. \tag{6.2}$$

The ratio of the aberration content at the lower energy spread to the content at the higher energy spread is therefore just the ratio of the energy spreads raised to the power n:

$$\frac{A_{chrom}^{small}}{A_{chrom}^{big}} = \left(\frac{\sigma_\delta^{small}}{\sigma_\delta^{big}} \right)^n. \tag{6.3}$$

The worst case dilution of the beam size at the low energy spread occurs when the RHS of Equation 6.3 is maximized for $\sigma_\delta^{small} < \sigma_\delta^{big}$. This occurs when $n = 1$. As a

worst-case for the low-energy-spread case, we can therefore state that:

$$(\sigma_y^{big})^2 = \sigma_0^2 + A_0^2(\sigma_\delta^{big})^2, \tag{6.4}$$
$$(\sigma_y^{small})^2 = \sigma_0^2 + A_0^2(\sigma_\delta^{small})^2.$$

Subtracting the two measured beam sizes in quadrature yields an expression for A_0:

$$(96.0\text{nm})^2 - (82.3\text{nm})^2 = A_0^2 \left[(\sigma_\delta^{big})^2 - (\sigma_\delta^{small})^2\right]. \tag{6.5}$$

Therefore, $A_0^2 [(3.6 \times 10^{-3})^2 - (5 \times 10^{-4})^2] = 2400\text{nm}^2$; this gives a value of A_0 of 1.37×10^4 nanometers. The chromatic beam size growth at the smaller energy spread is therefore 6.9 nanometers added in quadrature with the geometric spot size, which is negligible. Note that this is an *upper bound* on the chromatic contribution. The chromatic contribution could be smaller if the measured contribution at large energy spread scales with a higher power than the first power, especially since the bunch charge appears to be concentrated in the wings of the distribution, as shown in Figure 6.5. It is possible that the Sector 28 emittance measurements are not valid at the required level of precision, since such small emittances frequently result in beam sizes which are comparable to the default step size of the wire scanner stepper motors. The larger energy spread could also cause a head-tail effect in the beam which enlarged the measured RMS; it is not inconceivable that such an effect would be too small to affect the linac emittance measurement but large enough to dilute the Laser-Compton beam size measurement. Finally, the size with the larger energy spread could be a systematic effect of a non-Gaussian distribution affecting the Laser-Compton BSM [66].

While the result above absolves the chromatic aberrations from the diluted beam size at low energy spread, the total contribution at the large energy spread is 49 nanometers added in quadrature to the geometric spot size. This is much larger than the largest expected uncorrected aberration of the FFTB beamline (See Chapter 2). For the design vertical beam parameters ($\epsilon_y = 3.0 \times 10^{-11}$ m · rad, $\beta_y^* = 100$ μm), the

linear monochromatic beam size is 55 nm; an aberration of 49 nm added in quadrature would enlarge this to 74 nm.

Assuming for the moment that the beam size increase is due to chromatic aberration, what could possibly cause such an effect? In Chapter 4, evidence was presented for a persistent vertical dispersion, up to several millimeters in amplitude, through the CCSY and FT sections. Such a vertical dispersion would cause both high-order chromatic aberrations, such as T_{366}, and high-order chromo- geometric aberrations, such as T_{316} and T_{326}. In simulation studies, a beam with a Gaussian energy distribution with RMS of 5×10^{-4} and the typical FFTB emittances can still be tuned to 45 nanometers; however, a beam with a flat energy distribution with RMS of 3.6×10^{-3} and the typical FFTB emittances can only be tuned to a minimum spot size of 78 nanometers. The difference in quadrature between these two beam sizes is 65 nanometers, quite close to the observed 49 nanometer aberration. Such a result is another strong argument for attempting to null the observed η_y through the FFTB during the next run; nonetheless, the overall chromatic properties of the FFTB must be considered well under control, since the uncorrected chromaticity (as shown in Figure 2.11) is enormous.

With chromatic aberrations of the FFTB eliminated as suspects in the dilution of the beam size, it is necessary to turn to geometric aberrations, and systematic errors in the beam size measurement processes. In an attempt to measure the collective geometric aberrations, the beam size was measured as a function of the vertical emittance. Vertical emittance was varied by introducing dispersion in the damping ring extraction line, which filamented into additional emittance at the end of the linac; emittance was measured by the Sector 28 multi-wire station. The measurement was inconclusive: a line through the data points with no monitor offset and a line with a 40 nanometer offset fit the data equally well [65]. Such measurements are difficult because at the default scan ranges and step sizes for the Sector 28 wires change as a function of emittance, as do the linac and FFTB collimation settings required to elim-

inate backgrounds on the Laser-Compton BSM. A more careful set of measurements of this type is planned for the future FFTB experiments.

Chapter 7

Conclusions and

Recommendations for NLC

The Final Focus Test Beam has achieved and measured a vertical beam size of 70 nanometers at the Focal Point while operating with a vertical emittance of 2×10^{-11} m·rad, and an RMS energy spread of 5×10^{-4}. This beam size has been achieved repeatedly, recovered after interrruptions in FFTB running of up to four months, and measured by two extremely different beam size monitors. Measurements over time have indicated that the spot size is stable over many hours, to the extent that scans of the waists, dispersions, and coupling serve to restore the measured beam size. A direct test of the chromaticity of the beamline indicates that the chromatic correction is behaving as predicted; a global test of the chromatic aberrations indicates a maximum chromatic contribution to the beam size of 49 nanometers at an energy spread of 3.7×10^{-3}, which is added in quadrature to the monochromatic beam size.

What is the significance of such a result for the Next Linear Collider design effort? The optics and tolerances of the FFTB were mapped out in advance of construction through an array of theoretical studies and simulations; these tolerances and other requirements were engineered into the systems *a priori*, and factored into the local and global tuning strategies; and the resultant system very quickly converged

upon a final state near to its design goals. The indication is that the physics of large demagnifications, and small β^* values, is understood at the level which permits calculation of optics and tolerances – we can be confident that we know enough to specify many of the tolerances of NLC Final Focus systems. Furthermore, the FFTB has demonstrated the feasibilty of a new generation of precision correction and diagnostic elements which have immediate application in the NLC design process. Micronresolution BPMs, laser-compton and gas-ion beam size monitors, 3 DOF magnet movers with submicron step size and placement accuracy, part-per-million high- current power supplies – such devices are no longer conjectural or theoretical in nature. Already BPMs with FFTB-level performance are being included in tuning simulations of the NLC [67], and several different variations of beam size measurement by Compton scattering are expected to be used in such a collider. The experience of the FFTB leads to the expectation that such devices can be made to work because they already have.

7.1   Future Issues for the FFTB

The Final Focus Test Beam cannot be counted as a completed endeavor. Many areas remain to be explored, or explained. The most obvious of these is the fact that the FFTB has never achieved the beam size anticipated when the vertical emittance is considered. At the reduced vertical emittance and low energy spread quoted above, beam sizes under 50 nanometers should have been achieved. Furthermore, spot size dilution observed when the energy spread is increased is not consistent with the FFTB design. These difficulties represent a modest increase in the FFTB beam size, but would completely dominate the NLC colliding beam size. It is therefore imperative that these phenomena be understood.

At the lower energy spread, Chapter 6 argues that the sum of all chromatic aberrations is far too small to account for the discrepancy between the expected beam

size and the achieved. The remaining suspects are: geometric aberrations; wakefields; vibration-induced beam jitter; and unexplained measurement errors. Chapter 5 discussed the difficulties in using the Final Transformer geometric sextupoles to eliminate manufacturing defects in the quadrupoles and imperfections in the CCS sextupole system (strength imbalances and rolls). These difficulties have been understood, and will be addressed procedurally in future runs via a multi- dimensional scan. In addition, a careful measurement of the focused beam size as a function of incoming emittances will give more general information on the presence or absence of higher-order aberrations. Using the beta match steering magnets to introduce position and angle changes at the FP has also been suggested [68] as a means of probing the aberrations of the FFTB. Wakefields seem an unlikely source of spot size dilution in this case, because the bunch charge in the FFTB is small and the FFTB itself is relatively free of tight vertical apertures. One probe of this possibility would be to reduce the incoming intensity and measure the beam size again. This is a touchy measurement, because so many other parameters typically couple to bunch charge (extraction emittance from the damping ring, for example). Decoupling any beam size change from the intensity alteration from other effects would be difficult, but in principle such a measurement is possible. Finally, the arena of unmeasured measurement systematics is one which is dominated by the imagination of the experimenters. Such measurement errors which have already been considered are routinely measured and found minimal. This remains an area of ongoing concern. One recommendation which has been made is the inclusion of a 6° vertical measurement mode in future versions of the Laser-Compton BSM. Such a mode would measure the vertical beam size from 750 nm to 3.4 μm. The advantage of such a mode is that beam sizes of 350 nm, which are routinely and easily achieved in the FFTB, would appear in such a vertical mode as 100% modulation. Thus, measuring a 100-300 nm beam in a 6° mode would ensure that no error in the laser or the mirror system was causing the nulls in the interference pattern to fill in.

While the chromatic aberrations appear to pose no problem at the lower energy spread, understanding their source and determining a correction scheme are imperative for the NLC, which will not have the option of operating in such a condition. The vertical dispersion noted in Chapter 4 begs the question: is the vertical dispersion measured by varying the centroid energy related to the FP chromatic aberration observed at large energy spread? Such an aberration could occur via the interaction of vertical dispersion with the chromaticity of the quadrupoles and the sextupole field of the sextupoles, resulting in additional second-order effects. The measured vertical dispersion ray could result from rolled bend magnets or rolled quadrupoles. In either case, the first-order effects (dispersion and coupling at the FP) can be corrected by magnets in the FFTB, but no scheme for tuning the second-order "feed-down" effects exists, save to eliminate the incoming dispersion directly. Should the dispersive ray in Chapter 4 prove to be a result of rolled quadrupole magnets, it is also conceivable that the coupling observed at WS3 in Chapter 3 arises from the same source. The possibility that all of these phenomena are generated by simple uncorrected rolls in the bends and/or quads of the FFTB highlights the utility of maximal local correction of beamline errors: much time has been spent understanding the rolled spot at WS3, and any situation which invites untunable feed-down errors is to be discouraged.

In addition to the correction of aberrations already observed, the FFTB remains the best possible test-bed for prototyping NLC Final Focus technology and algorithms. As noted above, scanning the global aberration knobs is a standard means to reduce the spot size, and this is periodically necessary to retain good beam conditions. A significant improvement in the overall final focus system would be elimination of this requirement. A major driving force in the slow dilution of the beam size is believed to be the CCS central quadrupoles, QN1 and QM2. Misalignment of these magnets causes steering errors between the SF1 or SD1 sextupoles, resulting in waist, dispersion, or coupling at the FP. The alignment tolerances for these elements is on the order of 0.3 μm [69]. Two approaches which address this difficulty are the use

of feedback loops which use the BPM readings from quadrupole BPMs on either side of each sextupole to correct for any such steering effect, and a stretched-wire system which monitors the relative position of each magnet in the FFTB [70]. Tests of both systems are in the works, and a feedback similar to the CCS feedback described above has been proposed for NLC [67].

While the FFTB BPMs are capable of 1 μm resolution at the design intensity, measurement of far smaller variations in beam position is expected to be a necessity at any collider with beam sizes in the nanometer range. The FFTB currently contains a BPM which uses C-band (5712 MHz) trapped modes to measure beam position [71], and this device has achieved an approximate resolution of 100 nm. Future runs will include several such devices, with a goal of demonstrating a resolution in the vicinity of 1 nanometer. Such high-resolution BPMs will allow pulse-to-pulse jitter correction at the Laser-Compton BSM, and tests of a feedback which maintains collisions at the nanometer level (in this case, the BPM functions as a "target," substituting for the other beam). In addition, the small beam sizes at FFTB, coupled with such BPMs, will allow the contribution to beam jitter from magnet vibration to be measured directly, and compared to theoretical predictions. This is vital, as a well-developed theory of ground-motion effects is essential to the NLC Final Focus. Ultimately, a "feed-forward" based upon geophone measurements of magnet vibration may be possible.

7.2 Recommendations for NLC

The experiences of the Final Focus Test Beam, both positive and negative, make possible several specific recommendations for the design of the NLC Final Focus. These recommendations are summarized below.

Clean living pays off: In general, the more vigorously any particular tolerance could be attacked, the less beam-time was required by it. Magnet strengths,

for example, were determined and fixed very carefully: magnets were given independent power supplies which regulate with significantly greater precision than required; magnet hysteresis curves were carefully measured, and the standardization cycle included in the control system; power supplies were equipped with dual readbacks to reject common-mode failures of the transductor. Because of the effort invested in all the systems which connect the magnet strength in the control system with the real magnet strength, relatively little beam time has been lost to quad strength errors, drifting power supplies, etc. On the other hand, planarizing the FFTB (ensuring that magnets, movers, and BPMs all conform to the same x-y coordinate system) received somewhat less attention, and this remains a possible problem with implications for tuning and chromatic aberration; and FT sextupole alignment was not considered a priority, and many hours of beam-time were used to determine that the magnets could not be scanned with impunity.

More specifically, pre-beam tests to ensure the planarity of the final focus at the sub-milliradian level are clearly in order for the NLC. Independent power supplies for each magnet, with independent controllers, are highly recommended. A recurring problem in the SLC, and many other accelerators, is turn-to-turn magnet shorts, which are invisible to constant-current power supplies. A system which monitors the resistance of the magnets (via the power supply voltage) and flags any units which are significantly different from some reference value, would reduce the incidence of this problem.

Magnet Movers: The ability to align FFTB magnets remotely with high precision has been tremendously helpful. In addition, unexpected benefits of the system have included the sextupole mover knobs, mover-based lattice diagnostics, and steering knobs whose absolute accuracy is coupled to the accuracy of quadrupole strengths and mover LVDT readbacks. The final focus is likely to be the area of the NLC which demands the greatest flexibility in tuning and diagnostics, and movers for every final focus magnet will likely be required.

Specific improvements to the beam-based alignment algorithm are discussed in Chapter 4. A system which uses the more sophisticated cuts described in Chapter 4 would be useful if sufficient computer time were available. In addition, a system which aligns the entire beamline to a common line is conceivable. Such a system might use a more optimized algorithm for quadrupole alignment, and possibly a "soft constraint" on amplitude of magnet motion [47]. Adaptation of the sextupole alignment algorithm to the NLC appears feasible, especially since the beam jitter at the IP phase will be quite small, and therefore a relatively simple jitter-correction algorithm will enhance this procedure.

Coupling Correction and Beta Matching: Both the FFTB and SLC Final Focus are constructed with beta-matching normal quadrupoles, skew quadrupoles, and emittance measurement wires interleaved in one region. This arrangement is vulnerable to many difficulties, including "dead zones" in the parameter space where correction is not possible. Furthermore, both final focus systems have been constructed such that only two out of four coupling terms may be corrected. This leads to difficulties interpreting beam size measurements and beam matching. For example, the angular divergence measurement will be coupled by $(x'y')^*$, leading to a rolled beam on the divergence wire. This begs the question: is the roll angle equal to what the emittance measurement predicts?

Because of these difficulties, the NLC Final Focus is envisioned to fully separate coupling cancellation, emittance measurement, and beta matching [72]: a set of four skew quads, set in betatron phase to cancel all four coupling terms, are followed by a set of beam size measurement devices which are phased to perform a full 4D beam reconstruction; these are then followed by the normal quads which match the beam. In this way the emittance wires can make a measurement of the residual, post-correction coupling, eliminating the difficulty of measuring emittance when the projected vertical emittance is dominated by coupling. Furthermore, the final focus beamline is then uncoupled from the last correction skew quad to the IP, with the possible exception

of a single QS3-style doublet skew quad, and can be easily understood and tuned. All final focus wires should produce an unrolled image. Finally, the use of multi-wire emittance eliminates the invasive quad emit measurements.

Some consideration of the beta matching quads is required to determine whether the system is sufficiently flexible. The FFTB beta matching region includes one quad with a reversible power supply (QA0), which has been essential in operating in conditions other than the design (such as low-divergence). The use of such quads for "Irwin Knobs" (see Chapter 3) should be anticipated.

A beta matching program which is more fully optimized to the optics of the final focus is conceivable. Specifically, such a program should use the fully- correlated error matrix from beam measurement to predict the uncertainty in the beam parameters at the tune-up points. Even more useful would be a system which predicts uncertainty in beam sizes and waist positions rather than in β and α.

Beam Size Measurements: The inclusion of an IP image point in a final focus system is a topic of considerable debate. In the FFTB, the wire scanners at the image points (WS2 and WS3) have been used to verify the match of the incoming beam. Because WS2 and WS3 are easier to use and less exotic than the BSMs, these gave the fastest and most reliable information about changes in incoming beam conditions. While inclusion of an image point is therefore recommended, it is further suggested that an image point with an aspect ratio greater than 50:1 be avoided if possible, due to the unreasonable constraints on beam size monitor installation thus engendered.

In practice, it is not always possible to design an aberration-free final focus system which contains an image point upstream of the IP. Chapter 5 mentions the possibility of a tune-up image point downstream of the IP. Such an arrangement has some advantages: it relieves a constraint on the design of the body of the final focus, places the image downstream of both sextupole families so that horizontal and vertical chromaticities can be simultaneously cancelled at the image, and couples the beam size at the image to all the magnets upstream of the IP. The main disadvantage is

that the extraction line must also accept beamstrahlung photons, e^+e^- pairs from the IP, and the disrupted outgoing beam; to add an IP image to the design is a potentially serious complication. This is nonetheless an idea which deserves some consideration.

Because of the extreme power density of the NLC beam, it is unlikely that any 4 μm or 7 μm wires can be used at all. More exotic devices such as the laser-Compton BSM or a simpler laser-wire [73] will be necessary for the smaller beam sizes, while divergence monitors can probably make do with larger and less vulnerable wires. A laser-wire device is currently in development for SLC; in all likelihood, both types of measurement devices will be well-understood by the time NLC is built.

At the present time, no technology seems likely to permit direct single-beam measurements of the NLC's focused spot in the vertical. The presence of a particle physics detector complicates the design of any such device in any case; while a beam size monitor can be envisioned which allows tuning of the various knobs (via finding the center of a parabola), only beam-beam scans seem usable for measuring and tuning the final spot. The advisability of installing a tuning beam size monitor for "capture" tuning depends upon the anticipated difficulty of measuring beam sizes via beam-beam deflection scans when both beams are large and poorly tuned. This situation makes the desirability of IP image beam size monitors even greater for NLC than for FFTB.

Global Knobs: The FFTB uses two sets of global knobs: a magnet set for implementing relatively large changes, and a sextupole mover set for scanning and implementing small changes. This is an approach which has worked well, and is recommended for NLC. One difficulty is that NLC Final Focus designs have extremely strong sextupoles and large $R_{12,34}$ matrix elements from the sextupole to the IP, such that the necessary motions of the CCS sextupoles are sometimes as small as tens of nanometers. One sensible workaround is to fashion the CCS sextupoles as a pair of magnets: One large, powerful magnet which supplies almost all of the chromatic correction, and a set of trim sextupoles at the same locations, which are short and have large apertures and supply perhaps 1% of the chromatic correction. Such an

arrangement also allows the two main sextupoles in each CCS (x and y) to be powered or controlled together even if they are poorly matched: if the desired strength of each sextupole is K_s, then the sextupoles can be set to $K_s + dK_s$ and $K_s - dK_s$, while the trim sextupoles are set to $-dK_s$ and $+dK_s$, respectively. Finally, this arrangement decouples the chromatic correction from the sextupole mover knobs, by allowing the main sextupoles in a pair to be scanned in strength (assuming they are sufficiently well aligned). Such a scheme – using small sextupoles on movers for scanning – must be well-integrated into the overall scheme of the final focus. Specifically, the time needed for the movers to converge on a new position must be short, preferably comparable to a single wire scan (a few tens of seconds).

Alternatively, a set of weak quadrupoles at the CCS sextupole locations can be used for scanning. This would require two quads at each sextupole (normal and skew), and introduce some concern about thermal effects and hysteresis of the magnets.

The inclusion of geometric sextupole correctors is recommended, in that the tolerances on magnet construction errors may be loosened significantly. Geometric octupoles seem unneccessary at this time.

Feedbacks: The FFTB operates with two principal feedback loops (in addition to all the loops which operate in the injector, the damping ring, and the linac): launch and energy. The launch feedback utilizes only 2 BPMs upstream of the FFTB, while the energy loop utilizes 5 BPMs in the extraction line. In general a system with more BPMs is advisable, in the event that one or more fail. The launch feedback, for example, could use every BPM in the emittance section of the NLC Final Focus, which would redundantly span the entirety of (x,x') and (y,y') space.

Placing critical feedbacks in the downstream end of the NLC Final Focus is not recommended. For one thing, the post-collision beam in NLC will have an enormous emittance due to disruption; this will weaken confidence in any sensitive electronic systems (such as low-noise stripline BPM readout electronics), and also dilute the correlation between post-collision and pre-collision beam conditions. Furthermore,

the FFTB energy feedback cannot operate when the tune-up stopper ST62 is inserted; ideally, the energy feedback should be upstream of any major tune-up stopper or single-beam dumper. Finally, the BPMs which have the largest dispersion (and the largest coupling to the energy) should have the smallest coupling to geometric beam jitter which can be arranged. An ideal location for the energy feedback sensors in NLC is in the upstream half of the bend region which bends the beam from the linac into the final focus, as shown in Figure 1.1. This is similar to the SLC energy feedback, which uses BPMs in the upstream portion of the arcs [74].

Extraction Line: One of the most difficult experiences in the FFTB has been learning anything about the final doublet, or the magnets which immediately precede it in the Final Transformer. It is essential to remember that the extraction line will include all diagnostics which look at these critical magnets, and detectors for any IP beam size monitors which are installed. It is therefore crucial that the extraction line optics be arranged to allow such devices to do their job. Because the extraction line must also accomodate the disrupted outgoing beam and the e^+e^- pairs created at collision, adequate design of the extraction line may be one of the most challenging aspects of NLC.

Because the final doublet strength determines the location of the focused beam in z, it is recommended that the doublet be tunable to some degree, rather than using permanent magnets for the entire doublet.

Appendix A

Machine and Beam Parameters in a Telescopic System

The familiar machine and beam parameters of accelerator physics were derived to describe the properties of closed, periodic storage rings and the beams which were recirculating therein. These derivations do not apply directly to the Final Focus Test Beam, nor to any other beam line constructed primarily of modules with pure-telescopic behavior. Nonetheless, the familiar β, α, etc., are used throughout the FFTB design and operation. The purpose of this Appendix is to provide an understanding of how such parameters may be adapted to the unfamiliar environment of telescopic, single-pass lattices.

A.1 Twiss Parameters: $\beta, \alpha, \gamma, \epsilon$

The most commonly-used parameters in accelerator physics (other than energy, luminosity, and cost) are the Twiss parameters: β, α, γ; and the emittance, ϵ. In order to fully understand the issues of these parameters, it is necessary to consider briefly the single-particle dynamics of a beam transport line. The following derivation is heavily influenced by several sources, most notably Wiedemann [75], Brown and Servranckx

[5], and Carey [76].

A.1.1 Derivation of the Twiss Parameters

Let us consider for this derivation a beam transport system with no cross-plane coupling and no dispersion. In such a transport line, we may write the equation of motion for a charged particle as:

$$u''(s) + K(s)u(s) = 0, \tag{A.1}$$

where u is taken to be a generalized position coordinate (x or y), prime designates differentiation along the longitudinal axis of the beam line, and $K(s)$ is the focusing field strength at the location s, as defined in chapter 1. For a constant $K(s)$, of course, Eq. A.1 is a simple harmonic oscillator. Let us therefore assume a solution of the form

$$u(s) = \sqrt{W\beta(s)} \cos(\psi(s) - \psi_0), \tag{A.2}$$

in which the oscillator's amplitude β and phase ψ are functions of s, and W and ψ_0 are constants of integration. We may test the validity of this solution by differentiating twice and substituting back into Eq. A.1. For notational simplicity, the variables u, β, and ψ are all understood to be functions of s.

$$u' = \sqrt{W} \left[\frac{\beta'}{2\sqrt{\beta}} \cos(\psi - \psi_0) - \sqrt{\beta} \sin(\psi - \psi_0)\psi' \right], \tag{A.3}$$

$$u'' = \sqrt{W} \left[\frac{\beta\beta'' - \frac{1}{2}(\beta')^2}{\beta^{3/2}} \cos(\psi - \psi_0) - \frac{\beta'\psi'}{\sqrt{\beta}} \sin(\psi - \psi_0) \right.$$
$$\left. - \sqrt{\beta} \sin(\psi - \psi_0)\psi'' - \sqrt{\beta} \cos(\psi - \psi_0)(\psi')^2 \right]. \tag{A.4}$$

When we substitute this expression back into Eq. A.1, and require that the solution hold for all values of ψ, we find that we get two conditions, one from the sine terms and one from the cosine. From the sine terms, the relevant condition is:

$$\sqrt{W} \left[\sqrt{\beta}\psi'' + \frac{\beta'\psi'}{\sqrt{\beta}} \right] = 0, \tag{A.5}$$

while the condition from the cosine terms is:

$$\sqrt{W}\left[\frac{\beta\beta'' - \frac{1}{2}(\beta')^2}{\beta^{3/2}} - \sqrt{\beta}(\psi')^2 + K\sqrt{\beta}\right] = 0. \tag{A.6}$$

Note that Equation A.5 can be used to derive a relationship between the amplitude and phase angle in Eq. A.2. We can multiply through by $\sqrt{\frac{\beta}{W}}$, obtaining:

$$\beta\psi'' + \beta'\psi' = 0, \tag{A.7}$$

and, recognizing that $(\beta\psi')' = \beta\psi'' + \beta'\psi'$, we write

$$(\beta\psi')' = 0. \tag{A.8}$$

We can therefore assert that $\beta\psi'$ is a constant. If we define this constant to be 1, then

$$\psi = \int_0^s \frac{ds}{\beta}. \tag{A.9}$$

The parameter ψ, when defined as in Equation A.9, is known as the *phase advance* between the initial point and point s.

Equation A.6 can now be used to determine the evolution of β from one point of the beamline to another. Defining the addtional Twiss parameters, α and γ as follows:

$$\alpha \equiv -\frac{1}{2}\beta', \quad \gamma \equiv \frac{1+\alpha^2}{\beta}, \tag{A.10}$$

we can solve Eq. A.6:

$$\alpha' = K\beta - \gamma. \tag{A.11}$$

In addition, let us explicitly write u' in this notation:

$$u' = \sqrt{W}\left[\frac{-\alpha}{\sqrt{\beta}}\cos(\psi - \psi_0) - \frac{1}{\sqrt{\beta}}\sin(\psi - \psi_0)\right]. \tag{A.12}$$

We may now solve for W using our epxressions for u (Eq. A.2) and u' (Eq. A.12) to find:

$$\gamma u^2 + 2\alpha u u' + \beta u'^2 = W. \tag{A.13}$$
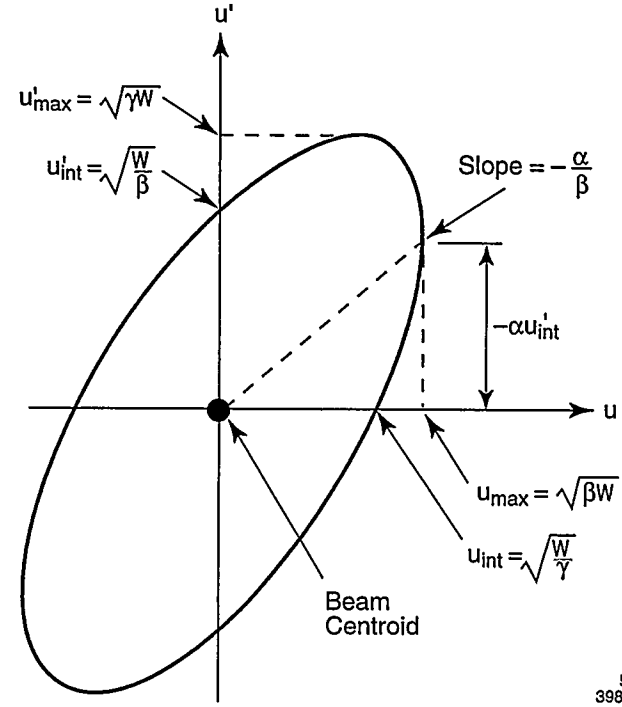
Figure A.1: Ellipse defined by parameters β, α, γ, with area πW, in the uu' plane.

The significance of Equation A.13 shall be discussed shortly. For the time being, it is worthwhile to note that Equation A.13 is the functional form of an ellipse the uu' plane. The arrangement of this ellipse in the uu' plane is shown in Figure A.1. Any particle at longitudinal location s in this beamline, whose motions are described by a given set of parameters W, β, α, will have position and angle relative to the reference line of the system which satisfies Equation A.13. The parameters β, α, γ are collectively known as the *Twiss parameters*.

A.1.2 Twiss Parameters and the R Matrix

Let us now perform two transformations to the particle trajectory determined by A.2 and A.12. First, let us eliminate the arbitrary phase angle ψ_0 and instead write both position and angle in terms of a component proportional to $\sin\psi$ and a component proportional to $\cos\psi$:

$$u = a\sqrt{\beta}\cos\psi + b\sqrt{\beta}\sin\psi, \tag{A.14}$$

$$u' = a\left(\frac{-\alpha}{\sqrt{\beta}}\cos\psi - \frac{1}{\sqrt{\beta}}\sin\psi\right) + b\left(\frac{-\alpha}{\sqrt{\beta}}\sin\psi + \frac{1}{\sqrt{\beta}}\cos\psi\right). \tag{A.15}$$

Second, let us define an arbitrary starting point, $s = 0$, at which we define:

$$u(0) \equiv u_0, \quad u'(0) \equiv u'_0, \quad \beta(0) \equiv \beta_0, \quad \alpha(0) \equiv \alpha_0, \quad \psi(0) \equiv 0. \tag{A.16}$$

Once this is done, we can define a and b in terms of the initial position and angle of the particle in question:

$$a = \frac{u_0}{\sqrt{\beta_0}}, \quad b = u'_0\sqrt{\beta_0} + \frac{\alpha_0 u_0}{\sqrt{\beta_0}}. \tag{A.17}$$

We can now use A.17 to eliminate a and b in Equations A.14 and A.15. This gives us $u(s)$ and $u'(s)$ as a function of u_0 and u'_0. Note that this is simply the definition of the R matrix between the initial point and the point s. Thus we can write the 2×2 R matrix as a function of the Twiss parameters at the beginning of the line, the end of the line, and the phase advance between the two:

$$R^{0 \to s} = \begin{pmatrix} \sqrt{\frac{\beta}{\beta_0}}(\cos\psi + \alpha_0\sin\psi) & \sqrt{\beta\beta_0}\sin\psi \\ \frac{\alpha_0 - \alpha}{\sqrt{\beta\beta_0}}\cos\psi - \frac{1 + \alpha\alpha_0}{\sqrt{\beta\beta_0}}\sin\psi & \sqrt{\frac{\beta_0}{\beta}}(\cos\psi - \alpha_0\sin\psi) \end{pmatrix}. \tag{A.18}$$

Once this is done, we can invert this relationship, writing a matrix for Twiss parameters at the point s as a function of the Twiss parameters at the starting point and the R matrix between the two:

$$\begin{pmatrix} \beta \\ \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} R_{11}^2 & -2R_{11}R_{12} & R_{12}^2 \\ -R_{11}R_{21} & 1 + 2R_{12}R_{21} & -R_{12}R_{22} \\ R_{21}^2 & -2R_{21}R_{22} & R_{22}^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \alpha_0 \\ \gamma_0 \end{pmatrix}. \tag{A.19}$$

A.1.3 Matched Twiss Parameters

Let us return to Equation A.18, which relates the uncoupled R matrix to the Twiss parameters at the endpoints of the segment of interest and the phase advance from one to another. Let us now consider the R matrix which represents a single turn around a storage ring, from a selected point S to itself, which we shall denote by $\mathcal{M}$. Does there exist a set of Twiss parameters which is transported by $\mathcal{M}$ back to itself? We can rewrite Equation A.18 for such a special case:

$$\mathcal{M} = \begin{pmatrix} \cos\mu + \alpha\sin\mu & \beta\sin\mu \\ -\gamma\sin\mu & \cos\mu - \alpha\sin\mu \end{pmatrix} \tag{A.20}$$

where we have used μ to indicate the phase advance due to one turn around the ring.

What is the significance of Equation A.20? For one thing, for the special case of the Twiss parameters which are transported back to themselves, these Twiss parameters are a function of the R matrix alone, and not of any other conditions. While the derivation of the betatron function, etc., was done in complete generality, we see here that there is a special relationship between this set of Twiss parameters and the R matrix of the storage ring.

In addition, consider Equation A.13, in which the Twiss parameters define an ellipse in the uu' phase plane. We see that the ellipse whose Twiss parameters are given by Equation A.20 is also transported into itself by a single turn around the storage ring. This ellipse is called the *machine ellipse* at point S, and its significance is that particles which lie on the machine ellipse at one time will remain on the ellipse in subsequent turns. Note that the position of an individual *particle* in the phase plane is not stationary; rather, a particle whose position and angle lie on the ellipse at one time will, after one turn, have a position and angle which is a different point on the same ellipse.

Now let us consider a large ensemble of particles which are all described by the machine ellipse at S. On subsequent turns around the storage ring the particles will exchange their positions on the phase plane with one another, but no particle will

migrate to a point off the ellipse. Because the particles are not distinguishable from one another, an observer measuring the phase space distribution of such particles would note that the distribution at S remains the same after each turn. Such a stationary distribution is called the *matched beam* at S; as we have seen, the Twiss parameters of such a distribution can be obtained solely from the R matrix around the storage ring from S back to itself.

There remains one free parameter in Equation A.13. This is the parameter W, which carries the information about the total area of the machine ellipse. Particles described by the parameters of the machine ellipse, but with different values of W, will lie on concentric similar ellipses. It is possible to choose a characteristic value of W which will describe the overall properties of the beam, and the total phase volume it inhabits. For example, we see from Figure A.1 that the ellipse is bounded in spatial extent by $\pm\sqrt{W\beta}$. Let us define ϵ to be that value of W such that the RMS distance of the particles in a bunch from the reference line is given by $\sqrt{\epsilon\beta}$. This is the definition of a beam's *emittance* in use at SLAC. We can also see from Figure A.1 that the RMS angular divergence of the beam is given by $\sqrt{\epsilon\gamma}$. Note that the definition of ϵ is arbitrary up to an overall scaling factor. Fermi National Accelerator Laboratory, for example, defines the area of the machine ellipse to be ϵ, and not $\pi\epsilon$, and defines the ellipse such that $\sqrt{\frac{\epsilon_u \beta_u}{\pi}} \equiv 2\sigma_u$[78].

A matched beam injected into a storage ring at a given point will remain matched (ie, described by the machine Twiss parameters) for all time. What happens if a mismatched beam is injected instead? Such a situation is shown diagrammatically in Figure A.2. Note that while the phase volume of the beam is not similar to the machine ellipse, *is still possible to draw a machine ellipse, of arbitrary area, which encloses the beam in phase space.* In this situation, the beam's phase volume is less than the volume of the machine ellipse needed to enclose it – the machine ellipse is not full. The ratio between the machine ellipse which is needed and the actual phase volume of the beam is a parameter known as *BMAG* in SLC notation. BMAG is a
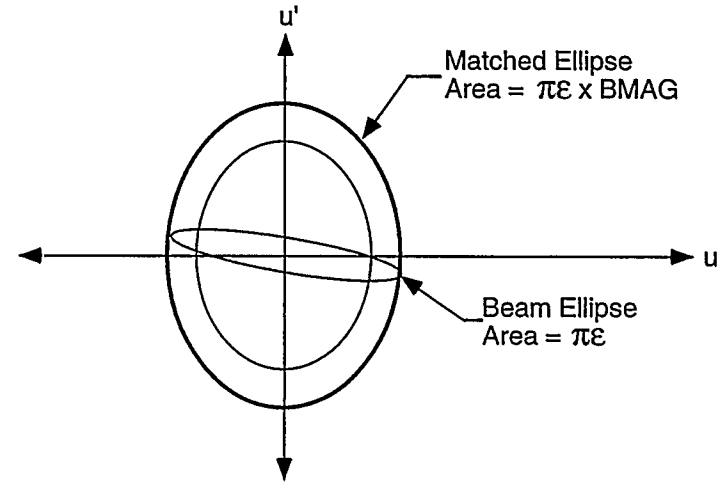
Figure A.2: Injection of a mismatched beam into a storage ring. The "maximal machine ellipse" needed to enclose the beam ellipse is shown. The particles within the beam ellipse will, on repeated turns, trace out the ellipses concentric to the maximal machine ellipse.

"mismatch parameter," such that a BMAG of 1.0 means a perfectly matched beam, and a BMAG greater than unity represents some severity of mismatch.

Note that, for any particle in the real beam, it is possible to draw an ellipse similar and concentric to the "maximal machine ellipse" which passes through that particle's position in phase space. We can regard that particle as lying on the machine-like ellipse, in which case it will be transported on subsequent turns to other points on that machine-like ellipse. This exercise can be carried out for each particle in the mismatched beam. In the case of perfectly linear transport, this will result in the ellipse containing the real beam precessing around within the maximal machine ellipse without distortion.

The real world is seldom as simple as purely linear beam transport, however. For one thing, the beam in a real storage ring will generally have a finite energy spread. At injection, we can think of this as injecting several bunches, each of which has

initially the same phase space coverage but at different centroid energies. As these bunches circulate around the ring, the leading effect of chromatic aberrations of the storage ring will be to cause the bunches at different energies to precess about the phase plane at different rates. This will result in an equilibrium situation in which the maximal machine ellipse has been completely filled. This process, known as *emittance filamentaton*, essentially guarantees that any bunch injected into a properly tuned storage ring will conform to the matched beam parameters in short order. The only parameter which is purely a function of the beam and not set by the machine is the emittance[1]. Because of this phenomenon, when the Twiss parameters are referred to by accelerator physicists, it is generally understood that the matched machine parameters are meant.

Let us return our attention to Equation A.20. While we have defined the matched Twiss for a circular, closed beam line, Equation A.20 places no such requirements upon us. Any beamline with a known R matrix can be decomposed as shown, and yield a "matched beam" which is the same at the exit as at the entrance. For a sufficiently long system (such as the SLC linac), the mechanism of emittance filamentation will even guarantee that the beam is matched at the end of the system, albeit with a potentially enormous emittance. The principal limitation is that R matrix be amenable to the decomposition of Equation A.20. Note, for example, that the trace of the matrix will yield the phase advance as follows:

$$Tr(R) = 2\cos\mu. \qquad (A.21)$$

Consequently, only a beamline whose R matrix has a trace between -2 and 2 can be decomposed in this fashion. One type of beamline for which this is not the case is a pure magnification, which has trace $M + \frac{1}{M}$. In this case, some simple algebra shows that $|Tr(R)| = 2$ for the case of $M = \pm 1$ and is larger for all other values of M, which agrees with our commonsensical notion that a *magnification* should not be capable of

[1]In fact, in electron storage rings, even the emittance is a function of the machine lattice, as synchrotron damping forces the beam to the ring's equilibrium emittance.

transporting *any* phase space distribution into itself [2]. While a $-I$ transform has a trace of -2, this requires a matched phase advance of π. According to Equation A.20, therefore, $R_{12} = 0$ for all values of the betatron function. This, again, agrees with our expectation that the $-I$ transform will transport *any* phase space distribution without distortion.

If our formalism indicates that pure magnification modules of any kind do not have a betatron function as we have come to understand the term, then this begs the question: what do we mean by "the betatron function" in such a context? In order to answer this question, it is necessary to define the sigma matrix of a beam.

A.1.4 The Sigma Matrix and the Beam Ellipse.

In contrast to the Twiss parameters, the sigma matrix of a beam can be quite simply defined. Consider a set of n particles distributed about some centroid (mean position), and let the distance of the particle from the centroid be given by coordinates $(x_i)_{1 \to n}$. We shall let i vary from 1 to 6, and use the standard beam physics notation: coordinate 1 is horizontal position (in meters or microns); coordinate 2 is horizontal angle ($\frac{dx_1}{ds}$), in radians or microradians; coordinates 3 and 4 are the vertical position and angle; coordinate 5 is longitudinal distance from the centroid, in meters or microns; and coordinate 6 is the fractional energy deviation, which is dimensionless. The sigma matrix is the set of second moments of the beam in this coordinate set, defined as follows:

$$\sigma_{ij} \equiv\; <x_i x_j> - <x_i><x_j>$$
$$= \frac{1}{n}\sum_{q=1}^{n}(x_i)_q(x_j)_q - \frac{1}{n^2}\sum_{q=1}^{n}(x_i)_q\sum_{q=1}^{n}(x_j)_q. \qquad (A.22)$$

There is another notation used to refer to the beam second moments, which I define here for completeness. This is the *correlation matrix* notation, which defines a vector

[2]This is not to imply that a magnifying module does not have any phase advance, but only that the inconsistent requirements that it magnify and yet leave the phase space distribution unchanged leads to a nonsensical result.

of beam sizes:

$$\sigma_i \equiv \sqrt{\sigma_{ii}}, \tag{A.23}$$

and a matrix of normalized correlation coefficients, denoted by r:

$$r_{ij} \equiv \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}, i \neq j. \tag{A.24}$$

Returning our attention to the sigma matrix, we can easily determine the rules by which a sigma matrix is propagated by a linear transport system. This is done by considering the transport rules for the coordinates of the constituent particles, namely $x_i(2) = R_{ij}^{1\to2}x_j(1)$, in which summation over repeated indicies is implicit. We can apply this rule to the particle coordinates in Equation A.22 to obtain:

$$
\begin{aligned}
\sigma_{ij}(2) &= \frac{1}{n}\sum_{q=1}^{n}x_i(2)x_j(2) - \frac{1}{n^2}\sum_{q=1}^{n}x_i(2)\sum_{q=1}^{n}x_j(2) \\
&= \frac{1}{n}\sum_{q=1}^{n}R_{ik}^{1\to2}x_k(1)R_{jl}^{1\to2}x_l(1) - \frac{1}{n^2}\sum_{q=1}^{n}R_{ik}^{1\to2}x_k(1)\sum_{q=1}^{n}R_{jl}^{1\to2}x_l(1) \\
&= R_{ik}^{1\to2}R_{jl}^{1\to2}\left[\frac{1}{n}\sum_{q=1}^{n}x_k(1)x_l(1) - \frac{1}{n^2}\sum_{q=1}^{n}x_k(1)\sum_{q=1}^{n}x_l(1)\right] \\
&= R_{ik}^{1\to2}R_{jl}^{1\to2}\sigma_{kl}(1), \tag{A.25}
\end{aligned}
$$

where the positions x_i, etc., are understood to be the positions of the qth particle. This propagation rule can be rewritten as a matrix equation by the non-intuitive step of replacing R_{jl} with $(R^t)_{lj}$:

$$\sigma_{ij}(2) = R_{ik}\sigma_{kl}(1)(R^t)_{lj} \tag{A.26}$$

$$\sigma(2) = R\sigma(1)R^t. \tag{A.27}$$

Let us consider, again, a beamline which contains no coupling elements (skew quadrupoles, solenoids, etc.), and also an initial sigma matrix which contains no cross-plane correlations. In this case we can reduce the both the sigma matrix and the R matrix to block-diagonal forms, three 2×2 matrices along the diagonal of the 6×6 matrix. In such a system, it has been shown [77] that the determinant of each

of the uncoupled sub-matrices is always unity. Combining this fact with Equation A.27, we see that the determinant of the uncoupled sub-matrices of the sigma matrix is unchanged by linear, uncoupled transport[3].

Let us consider now only one of the sub-matrices of the sigma matrix, which shall be arbitrarily denoted as the xx' submatrix, using coordinates x_1 and x_2. Since the determinant is constant under the beam transport considered here, the matrix can be factored into its determinant, given here by $\mathcal{E}^2$, and a matrix of parameters with determinant 1. Note that Equation A.22 requires the sigma matrix to be symmetric, so that:

$$
\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \mathcal{E}\begin{pmatrix} b & a \\ a & c \end{pmatrix} \equiv \mathcal{E}\mathcal{T}, \tag{A.28}
$$

where $\mathcal{T}$ is defined as the matrix of parameters whose determinant is unity. From Equation A.28 and the definition of the sigma matrix, we can see that the mean-squared distance of particles in the bunch is given by $\mathcal{E}b$, and the mean-squared angle by $\mathcal{E}c$. The determinant $\mathcal{E}$ can now be eliminated from the rule for sigma matrix transport and Equation A.27 rewritten in terms of the $\mathcal{T}$ matrix:

$$\mathcal{T}(2) = R\mathcal{T}(1)R^t. \tag{A.29}$$

Because the $\mathcal{T}$ matrix is symmetric, it contains only 3 distinct parameters [4]. It is therefore desirable to write a transport equation for the parameters of the $\mathcal{T}$ matrix which takes this into account. This yields a matrix equation for the parameters:

$$
\begin{pmatrix} b(2) \\ a(2) \\ c(2) \end{pmatrix} = \begin{pmatrix} R_{11}^2 & 2R_{11}R_{12} & R_{12}^2 \\ R_{11}R_{21} & 1+R_{12}R_{21} & R_{12}R_{22} \\ R_{21}^2 & 2R_{21}R_{22} & R_{22}^2 \end{pmatrix}\begin{pmatrix} b(1) \\ a(1) \\ c(1) \end{pmatrix}. \tag{A.30}
$$

---

[3]In fact, even in the case of a *coupled* sigma matrix, it can be shown that the 2×2 block-diagonal matrices have constant determinants when transported through an uncoupled linear beam line.

[4]Because $\mathcal{T}$ is symmetric and has unit determinant, it only contains 2 *free* parameters; the third "distinct" parameter is necessarily a linear combination of the other two.

The equation for transporting the $\mathcal{T}$ matrix parameters is hauntingly familiar. If, for example, we replace b with β, a with $-\alpha$, and c with γ, we see that Equation A.30 is identical to A.19.

What is the significance of all this? Remember that the Twiss parameters, and their transport laws, were originally derived without any assumptions about the type of transport system in use, other than that it be linear and uncoupled. The decomposition of Equation A.28 reduces the beam sigma matrix to an invariant component and three parameters; these three parameters correspond to the beam size, its angular divergence, and the correlation between position and angle within the bunch, in the same fashion as the Twiss parameters. In other words, for an arbitrary beam with a known distribution function, *it is possible to decompose the sigma matrix into a set of parameters which are mathematically identical to the emittance and Twiss parameters*. Thus, we can adapt the *machine* parameters β, α, γ into *beam* parameters in an environment in which the matched Twiss parameters are not defined. The advantage of this decomposition, and this notation, is that it automatically separates the portion of the beam matrix which cannot be manipulated by the linear transport system (emittance) from the parameters which can be so manipulated. This is helpful in that it implicitly prevents us from trying to independently adjust the beam size and angular divergence at the waist. It is also helpful because it allows us to adapt other concepts from storage rings, such as *chromaticity*, which are derived in terms of the betatron function. In sum, the machine ellipse is a special case of the sigma matrix decomposition shown here, with the caveat that a periodic system will force the beam sigma matrix to conform to the machine ellipse.

A.2 Twiss Parameters: Alternate Approach

In recent years, Irwin [79] has developed an alternate approach to the development of matched Twiss parameters. His approach bypasses the equations of motion entirely,

and relies only upon the linear transport matrix around one turn of the storage ring. In addition, the Irwin formalism is more easily extended to the case of a storage ring with x-y coupling in its lattice. In order to lay the groundwork for deriving the *Irwin Parameters*, which specify the intensity and phase of such coupling, his approach is summarized here.

As in Section A.1.3, $\mathcal{M}(s)$ is defined to represent the linear transport matrix around one turn of the storage ring, beginning and ending at point s, and we require (for the time being) that $\mathcal{M}$ be uncoupled and have zero dispersion ($\mathcal{M}_{i6} = 0$ for all $i \neq 6$). In this case, as before, we can consider $\mathcal{M}$ to be a 2×2 matrix around the storage ring. Consider now the eigenvalues and eigenvectors of $\mathcal{M}$. Because $\mathcal{M}$ is a real symplectic matrix, if λ is an eigenvalue then λ^* and $\frac{1}{\lambda}$ are also eigenvalues (where in this context λ^* means complex conjugate of λ, not the value of λ at the IP). We can express all the values of λ, in the most general case, as:

$$\lambda = e^{\pm i\mu}, \quad \lambda = e^{\pm\mu}, \qquad \text{(A.31)}$$

where μ is a *real, positive* value. Because we want to limit our consideration to stable storage rings, we reject the real eigenvalues (which would cause the beam to explode out of the ring after some number of turns) and turn to the complex eigenvalue. The complex eigenvalue corresponds to a complex eigenvector, which we shall denote v:

$$\mathcal{M}v = e^{i\mu}v. \qquad \text{(A.32)}$$

From this equation and the requirement that $\mathcal{M}$ be real, we can find the eigenvalue of $e^{-i\mu}$ to be v^*. Using these relations and general relations between complex conjugates, we can write the following:

$$\mathcal{M}(v + v^*) = e^{i\mu}v + e^{-i\mu}v^* = (v + v^*)\cos\mu + i(v - v^*)\sin\mu, \qquad \text{(A.33)}$$

$$i\mathcal{M}(v - v^*) = i(e^{i\mu}v - e^{-i\mu}v^*) = -(v + v^*)\sin\mu + i(v - v^*)\cos\mu. \qquad \text{(A.34)}$$

Let us, for notational convenience, make the following definition:

$$\mathcal{U} \equiv v + v^*, \quad \mathcal{V} \equiv -i(v - v^*), \qquad \text{(A.35)}$$

where both $\mathcal{U}$ and $\mathcal{V}$ are explicitly real vectors. We can now rewrite Equations A.33 and A.34 as the following:

$$\mathcal{M}_{11}\mathcal{U}_1 + \mathcal{M}_{12}\mathcal{U}_2 = \mathcal{U}_1 \cos\mu - \mathcal{V}_1 \sin\mu,$$

$$\mathcal{M}_{21}\mathcal{U}_1 + \mathcal{M}_{22}\mathcal{U}_2 = \mathcal{U}_2 \cos\mu - \mathcal{V}_2 \sin\mu,$$

$$-\mathcal{M}_{11}\mathcal{V}_1 - \mathcal{M}_{12}\mathcal{V}_2 = -\mathcal{U}_1 \sin\mu - \mathcal{V}_1 \cos\mu,$$

$$-\mathcal{M}_{21}\mathcal{V}_1 - \mathcal{M}_{22}\mathcal{V}_2 = -\mathcal{U}_2 \sin\mu - \mathcal{V}_2 \cos\mu. \tag{A.36}$$

Let us now define a transformation matrix, A_0, based upon the vectors $\mathcal{U}$ and $\mathcal{V}$:

$$A_0 \equiv \begin{pmatrix} \mathcal{U}_1 & \mathcal{V}_1 \\ \mathcal{U}_2 & \mathcal{V}_2 \end{pmatrix}, \tag{A.37}$$

and a rotation matrix, $\mathcal{R}(\mu)$:

$$\mathcal{R}(\mu) \equiv \begin{pmatrix} \cos\mu & \sin\mu \\ -\sin\mu & \cos\mu \end{pmatrix}. \tag{A.38}$$

With these definitions, we find that Equation A.36 is equivalent to a simple matrix equation:

$$\mathcal{M}A_0 = A_0\mathcal{R}(\mu). \tag{A.39}$$

We can now go one step further, and define $\mathcal{M}$ in terms of $\mathcal{R}$ and A_0:

$$\mathcal{M} = A_0\mathcal{R}(\mu)A_0^{-1}. \tag{A.40}$$

It is worthwhile to note at this point that the matrix $\mathcal{R}(\mu)$ represents a rotation in a single plane, and can therefore be rewritten:

$$\mathcal{R}(\mu) = \mathcal{R}(\theta)\mathcal{R}(\mu)\mathcal{R}(\theta)^{-1}. \tag{A.41}$$

This allows us to rewrite Equation A.40 as:

$$\mathcal{M} = A_0\mathcal{R}(\theta)\mathcal{R}(\mu)\mathcal{R}(\theta)^{-1}A_0^{-1} = A(\theta)\mathcal{R}(\mu)A(\theta)^{-1}, \tag{A.42}$$

in which we have defined $A(\theta) \equiv A_0\mathcal{R}(\theta)$.

Because $\mathcal{M}$ has a unit determinant, there are three independent parameters in Equation A.42. One is the rotation angle μ, which completely determines $\mathcal{R}(\mu)$. This means that $A(\theta)$ can have only two parameters which are determined by the value of $\mathcal{M}$, out of a total of four entries. This implies that there are two parameters in $A(\theta)$ which may be altered without changing the value of $\mathcal{M}$. One of these is the determinant: because $|A(\theta)| = (|A(\theta)^{-1}|)^{-1}$, the value of $|A(\theta)|$ cancels out in Equation A.42. Therefore we can, without loss of generality, require $|A(\theta)| = 1$. The remaining degree of freedom is θ, which is set below.

When we multiply out Equation A.42, we find:

$$\mathcal{M} = \begin{pmatrix} \cos\mu - \sin\mu(A_{11}A_{21} + A_{12}A_{22}) & \sin\mu(A_{11}^2 + A_{12}^2) \\ -\sin\mu(A_{21}^2 + A_{22}^2) & \cos\mu + \sin\mu(A_{11}A_{21} + A_{12}A_{22}) \end{pmatrix}. \tag{A.43}$$

We can now use our degree of freedom, θ, to adjust $A(\theta)$ such that $A(\theta)_{12}=0$. When this is done, Equation A.43 becomes:

$$\mathcal{M} = \begin{pmatrix} \cos\mu - \sin\mu(A_{11}A_{21}) & \sin\mu(A_{11}^2) \\ -\sin\mu(A_{21}^2 + A_{22}^2) & \cos\mu + \sin\mu(A_{11}A_{21}) \end{pmatrix}. \tag{A.44}$$

At this point, we can use Equation A.20 and obtain the values of $A(\theta)$ by inspection:

$$A(\theta) = \begin{pmatrix} \sqrt{\beta} & 0 \\ -\dfrac{\alpha}{\sqrt{\beta}} & \dfrac{1}{\sqrt{\beta}} \end{pmatrix}. \tag{A.45}$$

In sum: from a consideration of the mechanics of eigenvalues for a one-turn matrix for a storage ring, we can decompose the matrix into a component which carries the information about the phase advance, $\mathcal{R}$, and a component which carries the Twiss parameter information, $A(\theta)$. We can perform one additional decomposition to $A(\theta)$, specifically into a component which relates the correlation between position and angle in the beam (α), and a component which carries beam size information (β):

$$A(\theta) = \begin{pmatrix} \sqrt{\beta} & 0 \\ 0 & \dfrac{1}{\sqrt{\beta}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}. \tag{A.46}$$

Now let us consider the sigma matrix. From our previous experience, we have seen that the matched sigma matrix can be characterized by an emittance, ϵ, and the Twiss parameters, β, α, γ, without any consideration of phase advance. Consequently, we see that, in this notation, the sigma matrix must be a function of the emittance and $A(\theta)$ only. If we begin with the sigma matrix:

$$\sigma = \begin{pmatrix} \epsilon\beta & -\epsilon\alpha \\ -\epsilon\alpha & \epsilon\gamma \end{pmatrix}, \tag{A.47}$$

then we can, almost by inspection, decompose this into:

$$\sigma = A(\theta)\mathcal{E}A(\theta)^t, \tag{A.48}$$

where we have defined $\mathcal{E}$ to represent the "emittance matrix:"

$$\mathcal{E} \equiv \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix}. \tag{A.49}$$

A.3   Irwin Parameters: a, b, c, d

The approach taken in Section A.2 can be extended in a straightforward fashion to the case of a fully-coupled one-turn matrix around a storage ring. Consider the 4×4 transport matrix $\mathcal{M}$. In this case we have two degrees of freedom, conventionally x and y. We therefore expect to have two sets of eigenvalues which satisfy the same constraints set forth for the two-dimensional case. We can define the eigenvalues thus:

$$\lambda_1 = e^{\pm i\mu_1}, \quad \lambda_2 = e^{\pm i\mu_2}, \tag{A.50}$$

where 1 and 2 (rather than x and y) represent the degrees of freedom to avoid pre-judging the situation. We can define the eigenvectors of these eigenvalues as $v^{(1)}$ and $v^{(2)}$, and their complex conjugates. This then leads us to define the linear combinations of the eigenvectors:

$$\mathcal{U}^{(1)} \equiv v^{(1)} + v^{(1)*}, \quad \mathcal{V}^{(1)} \equiv -i(v^{(1)} - v^{(1)*}),$$
$$\mathcal{U}^{(2)} \equiv v^{(2)} + v^{(2)*}, \quad \mathcal{V}^{(2)} \equiv -i(v^{(2)} - v^{(2)*}). \tag{A.51}$$

a transformation matrix A_0:

$$A_0 \equiv \begin{pmatrix} \mathcal{U}_1^{(1)} & \mathcal{V}_1^{(1)} & \mathcal{U}_1^{(2)} & \mathcal{V}_1^{(2)} \\ \mathcal{U}_2^{(1)} & \mathcal{V}_2^{(1)} & \mathcal{U}_2^{(2)} & \mathcal{V}_2^{(2)} \\ \mathcal{U}_3^{(1)} & \mathcal{V}_3^{(1)} & \mathcal{U}_3^{(2)} & \mathcal{V}_3^{(2)} \\ \mathcal{U}_4^{(1)} & \mathcal{V}_4^{(1)} & \mathcal{U}_4^{(2)} & \mathcal{V}_4^{(2)} \end{pmatrix}, \tag{A.52}$$

and a rotation matrix for two degrees of freedom, $R(\vec{\mu})$:

$$R(\vec{\mu}) \equiv \begin{pmatrix} R(\mu_1) & 0 \\ 0 & R(\mu_2) \end{pmatrix}, \tag{A.53}$$

where $\vec{\mu} = (\mu_1, \mu_2)$ and $R(\mu)$ is the 2×2 rotation matrix as defined in Section A.2. We can write

$$\mathcal{M} = A_0 \mathcal{R}(\vec{\mu}) A_0^{-1} \tag{A.54}$$

and, defining $A(\vec{\theta})$ in analogy with Section A.2, we can write:

$$\mathcal{M} = A(\vec{\theta})\mathcal{R}(\vec{\mu})A(\vec{\theta})^{-1}. \tag{A.55}$$

In this notation, the full rotation matrix $\mathcal{R}(\vec{\mu})$ represents the phase advance of the storage ring in each of the *normal-mode* degrees of freedom, and $A(\vec{\theta})$ includes the usual scaling and rotation of the normal-mode phase planes *and the transformation between the normal modes and the coordinate system used in defining the real beamline*. This transformation between the normal modes and the coordinates in which $\mathcal{M}$ is represented is the coupling of the $\mathcal{M}$ matrix.

A.3.1   Factorization of $A(\vec{\theta})$

The state of affairs is now similar to that for the single degree of freedom case, in that we have a matrix, $A(\vec{\theta})$, which contains within it the beam transport properties of the single-turn matrix $\mathcal{M}$; and this matrix has two degrees of freedom, represented by $\vec{\theta} = (\theta_1, \theta_2)$. In addition, let us assert that the matrix $A(\vec{\theta})$ is *symplectic* [79] , ie:

$$A(\vec{\theta})^t J A(\vec{\theta}) = J, \tag{A.56}$$

where J is defined:

$$J \equiv \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{A.57}$$

Let us consider the properties of a generic 4×4 symplectic matrix, m. If m is known to be symplectic, then we can decompose it into four 2×2 submatrices:

$$m = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}, \tag{A.58}$$

where $|m_1| = |m_4| \equiv g^2$ and $|m_2| = |m_3|$. I can now define a pair of matrices with unit determinant based upon m_1 and m_4:

$$N_1 \equiv \frac{m_1}{g}, \quad N_2 \equiv \frac{m_4}{g}, \tag{A.59}$$

and install these as the block-diagonal components of a 4×4 matrix N:

$$N \equiv \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix}. \tag{A.60}$$

It is not difficult to show that I can now write m in a fascinating new form:

$$m = N(N^{-1}m) = \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix} \begin{pmatrix} gI & N_1^{-1}m_2 \\ N_2^{-1}m_3 & gI \end{pmatrix}$$
$$= NS, \tag{A.61}$$

where N represents all the "in-plane" phenomena of the two normal modes, and S represents all the "skew," or "cross-plane" effects between the normal modes and the coordinate system.

Now that we have isolated the "in-plane" effects of the two degrees of freedom, we are free to define $A(\vec{\theta})$ such that N_1 and N_2 have zeroes for their (1,2) elements, as does the $A(\theta)$ matrix in Section A.2. The properties of $N_{1,2}$ – namely, unit determinant

and a (1,2) element which is zero – permit us to factor them into betatron and alpha sub-matrices as we did before:

$$N_{1,2} = \begin{pmatrix} \sqrt{\beta_{1,2}} & 0 \\ 0 & \frac{1}{\sqrt{\beta_{1,2}}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\alpha_{1,2} & 1 \end{pmatrix}.$$
$$= \tilde{B}_{1,2}\tilde{A}_{1,2}. \tag{A.62}$$

We can therefore write:

$$N = \begin{pmatrix} \tilde{B}_1 & 0 \\ 0 & \tilde{B}_2 \end{pmatrix} \begin{pmatrix} \tilde{A}_1 & 0 \\ 0 & \tilde{A}_2 \end{pmatrix}$$
$$= \tilde{B}\tilde{A}. \tag{A.63}$$

As a final note, while we have chosen here to factor the normal component to the left of the skew component, it is easy to demonstrate that factoring the normal component to the right will result in a different skew matrix, but leaves the normal matrix unchanged.

A.3.2 The Structure of the S-Matrix

We are now left with the problem of determining the most general possible coupling matrix of the form in Equation A.61 above. In order to determine this transformation, we can simplify matters by using the Lie Transformation approach to accelerator problems used by Irwin [80] and Roy [6]. In this approach, we begin by writing the generator of the most arbitrary coupling in the x-y basis [5]:

$$S = -axy - bxy' - cx'y - dx'y', \tag{A.64}$$

where a, b, c, d can arbitrarily vary from $-\infty \to \infty$. The transformation given by the generator S can be determined by allowing the related *Lie Transform* to operate on

---

[5]There are several variants of the Irwin coupling notation, which vary in their assignment of the terms a, b, c, d and the overall sign. Here I define the terms as they have been defined in the SLC control system.

an arbitrary 4-vector, $\vec{x}$:

$$\vec{x}_2 = e^{:S:}\vec{x}_1, \tag{A.65}$$

where the definition of $:S:$ is as given in [80]. While the effects of a Lie transform can in general be quite messy, in this case we find that the expansion of $e^{:S:}\vec{x}_1$ converges upon a simple form. This allows us to write the Lie transform $e^{:S:}$ as a matrix.

Let us define D as:

$$D \equiv bc - ad, \tag{A.66}$$

and define g and h in two alternate forms:

$$g = \cosh\sqrt{D}, \quad h = \frac{\sinh\sqrt{D}}{\sqrt{D}}, \quad D > 0,$$

$$g = \cos\sqrt{|D|}, \quad h = \frac{\sin\sqrt{|D|}}{\sqrt{|D|}}, \quad D < 0. \tag{A.67}$$

We can now write the matrix S given by the generator $\mathcal{S}$:

$$S = \begin{bmatrix} gI & h\begin{pmatrix} c & d \\ -a & -b \end{pmatrix} \\ h\begin{pmatrix} b & d \\ -a & -c \end{pmatrix} & gI \end{bmatrix}. \tag{A.68}$$

A.3.3 The Fully-Coupled 4×4 Sigma Matrix

We can now complete our examination of the matrix $A(\vec{\theta})$. Using the form of S determined in Equation A.67, we can now fully express $A(\vec{\theta})$ as the product of three matrices:

$$A(\vec{\theta}) = \tilde{B}\tilde{A}S. \tag{A.69}$$

In analogy with section A.2, we can define the *normal-mode emittances,* ϵ_1 and ϵ_2. These are the invariant emittances obtained when the beam is considered in the normal-mode coordinate system. We can now write the sigma matrix in terms of

$A(\vec{\theta})$ acting on the emittance matrix:

$$\mathcal{E} = \begin{pmatrix} \epsilon_1 & 0 & 0 & 0 \\ 0 & \epsilon_1 & 0 & 0 \\ 0 & 0 & \epsilon_2 & 0 \\ 0 & 0 & 0 & \epsilon_2 \end{pmatrix}. \tag{A.70}$$

In analogy with the uncoupled case, we can write:

$$\begin{aligned} \sigma &= A(\vec{\theta})\mathcal{E}A(\vec{\theta})^t \\ &= \tilde{B}\tilde{A}S\mathcal{E}S^t\tilde{A}^t\tilde{B}^t. \end{aligned} \tag{A.71}$$

Without further ado, the fully-expressed elements of the sigma matrix in terms of the normal-mode emittances, Twiss and Irwin parameters are:

$$\begin{aligned} \sigma_{11} &= \beta_1[g^2\epsilon_1 + h^2(c^2 + d^2)\epsilon_2] \\ \sigma_{22} &= \frac{1}{\beta_1}\left\{g^2(1 + \alpha_1^2)\epsilon_1 + h^2[(a + \alpha_1 c)^2 + (b + \alpha_1 d)^2]\epsilon_2\right\} \\ \sigma_{12} &= -g^2\alpha_1\epsilon_1 - h^2[c(a + \alpha_1 c) + d(b + \alpha_1 d)]\epsilon_2 \end{aligned} \tag{A.72}$$

$$\begin{aligned} \sigma_{33} &= \beta_2[g^2\epsilon_2 + h^2(b^2 + d^2)\epsilon_1] \\ \sigma_{44} &= \frac{1}{\beta_2}\left\{g^2(1 + \alpha_2^2)\epsilon_2 + h^2[(a + \alpha_2 b)^2 + (c + \alpha_2 d)^2]\epsilon_1\right\} \\ \sigma_{34} &= -g^2\alpha_2\epsilon_2 - h^2[b(a + \alpha_2 b) + d(c + \alpha_2 d)]\epsilon_1 \end{aligned} \tag{A.73}$$

$$\sigma_{13} = \sqrt{\beta_1\beta_2}gh(b\epsilon_1 + c\epsilon_2)$$

$$\sigma_{14} = \sqrt{\frac{\beta_1}{\beta_2}}gh[-(a + \alpha_2 b)\epsilon_1 + (d - \alpha_2 c)\epsilon_2]$$

$$\sigma_{23} = \sqrt{\frac{\beta_2}{\beta_1}}gh[-(a + \alpha_1 c)\epsilon_2 + (d - \alpha_1 b)\epsilon_1]$$

$$\begin{aligned} \sigma_{24} = \frac{1}{\sqrt{\beta_1\beta_2}}gh[&(a\alpha_1 - d\alpha_2 + b\alpha_1\alpha_2 - c)\epsilon_1 \\ &+ (a\alpha_2 - d\alpha_1 + c\alpha_1\alpha_2 - b)\epsilon_2]. \end{aligned} \tag{A.74}$$

The preceding derivation was done for a closed storage ring, using the one-turn transport matrix $\mathcal{M}$. However, we have already seen that the betatron and alpha functions may be adapted from machine parameters to beam parameters, and thus $\tilde{B}$ and $\tilde{A}$ can be written for an arbitrary beam in any environment; and that the emittance matrix $\mathcal{E}$ is explicitly a function of the beam alone. Finally, the matrix S is the most general coupling matrix; consequently, we expect that the decomposition of coupling it represents does not require a closed storage ring to be valid (ie, a zero-length coordinate rotation between x and y can be represented by an S matrix). We can therefore state that any beam matrix can be decomposed into two emittances, two betatron functions, two alpha functions, and four Irwin paramters; and that the relationship between the parameters and the sigma matrix is as given in Equations A.72 – A.74.

A.3.4 Properties of the Irwin Parameterization

There are several useful properties of the Irwin parameters, and of the sigma matrix decomposition shown in Equations A.72–A.74. We have already encountered D, which is the negative of the determinant of the upper-right submatrix of the S matrix. We can also define r:

$$r^2 \equiv a^2 + b^2 + c^2 + d^2. \qquad (A.75)$$

It is possible to show that both r and D are invariant under uncoupled beam transport. If we define the *projected emittances* via the determinants of the upper-left and lower-right submatrices of the sigma matrix:

$$\epsilon_x \equiv \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}, \qquad (A.76)$$

$$\epsilon_y \equiv \sqrt{\sigma_{33}\sigma_{44} - \sigma_{34}^2}, \qquad (A.77)$$

then we find the following relation holds between the projected and normal-mode emittances[6]:

$$\epsilon_x^2 = g^4\epsilon_1^2 + h^4 D^2 \epsilon_2^2 + g^2 h^2 r^2 \epsilon_1 \epsilon_2,$$

$$\epsilon_y^2 = g^4\epsilon_2^2 + h^4 D^2 \epsilon_1^2 + g^2 h^2 r^2 \epsilon_1 \epsilon_2. \qquad (A.78)$$

From Equation A.78, it can be shown that the projected emittance can never be smaller than the smallest of the two normal-mode emittances. The converse is not true: the projected emittances can be made arbitrarily large, and can be larger than the sum of the normal-mode emittances. Consider the case of emittances given in Equation A.78, in which the values of g and h are given by their hyperbolic forms. In this case all three terms contributing to each projected emittance are positive-definite, and $g^4 \geq 1$.

Finally, let us consider the significance of the betatron and alpha functions determined in the decomposition of $A(\vec{\theta})$. It was mentioned in Section A.3.1 that the normal component of $A(\vec{\theta})$ could be factored to the right of the skew component, and this would result in a different skew matrix (with the same functional form) and an identical normal matrix. Let us call this skew matrix $\tilde{S}$. In this case, the sigma matrix decomposition of Equation A.71 becomes:

$$\sigma = \tilde{S}\tilde{B}\tilde{A}\mathcal{E}\tilde{A}^t\tilde{B}^t\tilde{S}^t. \qquad (A.79)$$

In Equation A.79, an uncoupled beam sigma matrix is constructed by the innermost matrix product, $\tilde{B}\tilde{A}\mathcal{E}\tilde{A}^t\tilde{B}^t$, and the coupling applied to this matrix after the fact. The innermost matrix product therefore represents the normal-mode beam matrix, with normal-mode emittances *and* normal- mode Twiss parameters, and the $\tilde{S}$ matrix represents the requisite coordinate transformation between this and the external coordinate system. Because the $\tilde{B}$ and $\tilde{A}$ matrices are identical to those extracted in Equation A.71, we can conclude that the Irwin recipe for decomposing the sigma

---

[6]These relations were originally worked out by P. Raimondi.

matrix extracts the normal-mode Twiss parameters, as well as the normal mode emittances.

A.3.5 Alternate Parameterizations: The Spence Parameters

Another parameterization of the coupling in common use at SLAC was developed by W. Spence. The approach is similar, but the matrix S defined in Equation A.68 in the following way:

$$
S = \begin{bmatrix} b_{\text{Spence}}I & \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\ \begin{pmatrix} -B_{22} & B_{12} \\ B_{21} & -B_{11} \end{pmatrix} & b_{\text{Spence}}I \end{bmatrix}. \tag{A.80}
$$

The upper-right 2×2 matrix is known as the "B-matrix". By comparing the Spence and Irwin parameters, we arrive at the following relations:

$$
\begin{aligned}
|B| &= -\sinh^2\sqrt{D}, \quad D > 0 \quad (|B| < 0), \\
&= \sin^2\sqrt{|D|}, \quad D < 0 \quad (|B| > 0), \\
b_{\text{Spence}} &= \sqrt{1 - |B|} = g.
\end{aligned} \tag{A.81}
$$

A general purpose conversion between the Spence and Irwin parameters can easily be deduced from these relations.

Each of the two parameterizations has its advantages. For example, in the Spence approach, $|B| \leq 1$, which puts implicit constraints on the values of B_{ij}. A large, negative value of B_{21} coupled with a small positive value of B_{12} will result in a $|B|$ value greater than 1. Consequently, an arbitrarily-selected set of B_{ij} values will not necessarily correspond to a physically realizable transformation, whereas any combination of Irwin parameters will yield a real S matrix. On the other hand, it is possible to select multiple sets of Irwin parameters which yield the same S matrix.

Consider the case of $c = -b = \phi$. In this case, $D = -\phi^2$, $g = \cos(\phi)$, $h = \frac{\sin(\phi)}{\phi}$, and we can write the S matrix:

$$
S_\phi = \begin{pmatrix} \cos\phi & 0 & \sin\phi & 0 \\ 0 & \cos\phi & 0 & \sin\phi \\ -\sin\phi & 0 & \cos\phi & 0 \\ 0 & -\sin\phi & 0 & \cos\phi \end{pmatrix}. \tag{A.82}
$$

If we now let $c = -b = \phi + 2\pi$, we find the S matrix changes as follows:

$$
S_{\phi+2\pi} = \begin{pmatrix} \cos(\phi+2\pi) & 0 & \sin(\phi+2\pi) & 0 \\ 0 & \cos(\phi+2\pi) & 0 & \sin(\phi+2\pi) \\ -\sin(\phi+2\pi) & 0 & \cos(\phi+2\pi) & 0 \\ 0 & -\sin(\phi+2\pi) & 0 & \cos(\phi+2\pi) \end{pmatrix}, \tag{A.83}
$$

which equals S_ϕ by inspection. In sum, it is easier to go from Irwin parameters to a proper S matrix, but less ambiguous to go from an S matrix to Spence parameters.

A.4 Dispersion Parameters: D and η

So far we have considered only a linear system with no energy-dependent transport properties. Let us now consider the effect of lowest-order energy dependence. In a bending field which produces a radius of curvature $\rho(s)$, the off-energy particles will be over- or under-bent relative to the on-energy particles. If we define the reference orbit to be the closed orbit of the particles at the design trajectory, then this differential bending will result in the off-energy particles deviating from the reference orbit. We can quantify this effect, which leads to a modification to the equation of motion, Equation A.1:

$$
u'' + K(s)u = \frac{\delta}{\rho(s)}, \tag{A.84}
$$

where $\delta = \frac{\Delta E}{E}$, and $\rho(s)$ is the instantaneous radius of curvature at s[7]. The complete solution to Equation A.84 is given by [81]:

$$u(s) = u(0)C(s) + u'(0)S(s) + \delta D(s),$$
$$u'(s) = u(0)C'(s) + u'(0)S'(s) + \delta D'(s), \qquad (A.85)$$

where $C(s)$ and $S(s)$ are the conventional sinelike and cosinelike rays from the initial point to point s, and D can be determined by the method of Green's functions from C, S, and ρ between the initial and final points. C and S are, of course, the familiar R_{11} and R_{12} elements; C' and S' are the R_{21} and R_{22}. If we identify δ as the sixth coordinate of the particle (the fifth is longitudinal position from the centroid, which is not a crucial parameter in final focus systems), then the characteristic rays D_x, D'_x, D_y, D'_y translate to the $R_{16}, R_{26}, R_{36}, R_{46}$.

Let us consider only the xx' plane, which is the usual bend plane, with the understanding that all discoveries there can be equally well applied to the yy' plane if necessary. Let us now consider the 3×3 matrix R:

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{16} \\ R_{21} & R_{22} & R_{26} \\ 0 & 0 & 1 \end{pmatrix}. \qquad (A.86)$$

Furthermore, let us define $\mathcal{M}(s)$ to be the one-turn value of R from s back to itself:

$$\mathcal{M} = \begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} & \mathcal{M}_{16} \\ \mathcal{M}_{21} & \mathcal{M}_{22} & \mathcal{M}_{26} \\ 0 & 0 & 1 \end{pmatrix}. \qquad (A.87)$$

For a particle with coordinates $(x, x', \delta) = (0, 0, 0)$, we can see trivially that the particle will return to the same position after one turn, and this defines the reference orbit. A particle with an energy offset only will begin at $(0, 0, \delta)$ and return to

---

[7]In reality, the inhomogeneous term in Equation A.84 should indlude $\bar{\delta} \equiv \frac{\delta}{1+\delta}$, and not δ. The use of δ is traditional, and a good enough approximation for most cases.

$(\mathcal{M}_{16}\delta, \mathcal{M}_{26}\delta, \delta)$. This begs the question: for a given value of δ, does there exist a combination of position and angle at s which is transformed into itself? If so, this would constitute the closed orbit for off-energy particles, or (putting it another way) would show how the reference orbit is deformed for particles of different energies. Let us define this position to be $(\eta\delta, \eta'\delta, \delta)$ for the case of linear transport. We can determine the values of η and η' by expressing the stability condition:

$$\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} & \mathcal{M}_{16} \\ \mathcal{M}_{21} & \mathcal{M}_{22} & \mathcal{M}_{26} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta\delta \\ \eta'\delta \\ \delta \end{pmatrix} = \begin{pmatrix} \eta\delta \\ \eta'\delta \\ \delta \end{pmatrix}. \qquad (A.88)$$

From Equation A.88, we see that in general the characteristic functions $\eta(s)$ and $\eta'(s)$ depend upon all the elements of the matrix $\mathcal{M}$.

One special case of Equation A.88 is the case in which, for some location s_0, both $\mathcal{M}_{16}$ and $\mathcal{M}_{26}$ are zero. In this case, $\eta(s_0)$ and $\eta'(s_0)$ are both required to be zero as well, and the reference orbit for on-energy and for off-energy particles will coincide at point s_0. Let us now define $R_{16}^{s_0 \to s}$ and $R_{26}^{s_0 \to s}$ to be the R_{16} and R_{26} from the aforementioned s_0 to some other location s. A particle with coordinates $(0, 0, \delta)$ at point s_0 will be transported to $(R_{16}^{s_0 \to s}\delta, R_{26}^{s_0 \to s}\delta, \delta)$ at point s. Since the actions of the storage ring must transport this particle back to its starting position, it follows that the path of the particle around the ring constitutes the closed orbit for particles at its energy. Therefore, The condition described above requires that, at any point in the machine,

$$\eta(s) = R_{16}^{s_0 \to s},$$
$$\eta'(s) = R_{26}^{s_0 \to s}. \qquad (A.89)$$

A.4.1 Dispersion Matrix Decomposition

Before considering the application of the dispersion formalism to linear collider final focus systems, it is worthwhile to note a property of the dispersive transport matrices

in the previous section. Specifically, any of the matrices shown can be factored trivially into a geometric matrix and a chromatic matrix:

$$\begin{pmatrix} R_{11} & R_{12} & R_{16} \\ R_{21} & R_{22} & R_{26} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & R_{16} \\ 0 & 1 & R_{26} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\equiv \tilde{D}\tilde{M}. \tag{A.90}$$

A.4.2 Dispersion and the Beam Matrix

We saw in Section A.1.4 that it is possible to define a beam matrix which carries all the information pertaining to the second moments of the beam's distribution. We can extend this technology trivially to the case of systems with energy-dependent beam transport, using the same propagation law, Equation A.27. For simplicity, let us consider here only storage rings which have a point at which the on-energy and off-energy closed orbits coincide, as described above.

At the point of coincidence, we can see by inspection that the stationary beam distribution is given by the monochromatic Twiss parameters, which can be determined by application of Equation A.20. To obtain the matched beam at an arbitrary location downstream, we can apply the transport rule, Equation A.27, and the decomposition shown in the previous section:

$$\sigma_2 = R\sigma_1 R^t$$
$$= \tilde{D}\tilde{M}\sigma_1\tilde{M}^t\tilde{D}^t. \tag{A.91}$$

It is worthwhile to note that $\tilde{M}\sigma_1\tilde{M}^t$ is the beam transport for a non-dispersive system. Furthermore, it can be shown easily that $\sigma_{16} = \sigma_{26} = 0$ at the point of coincidence. Therefore, if we define $\sigma^{(D)}$ to be the full beam matrix including dispersion, and $\sigma^{(M)}$ to be the monochromatic matrix determined by use of $\tilde{M}$, then we find:

$$\begin{aligned} \sigma_{11}^{(D)} &= \sigma_{11}^{(M)} + R_{16}^2\sigma_{66}^{(M)} = \sigma_{11}^{(M)} + \eta^2(s)\sigma_{66}^{(M)}, \\ \sigma_{22}^{(D)} &= \sigma_{22}^{(M)} + R_{26}^2\sigma_{66}^{(M)} = \sigma_{22}^{(M)} + \eta'^2(s)\sigma_{66}^{(M)}, \\ \sigma_{12}^{(D)} &= \sigma_{12}^{(M)} + R_{16}R_{26}\sigma_{66}^{(M)} = \sigma_{12}^{(M)} + \eta(s)\eta'(s)\sigma_{66}^{(M)}, \\ \sigma_{16}^{(D)} &= R_{16}\sigma_{66}^{(M)} = \eta(s)\sigma_{66}^{(M)}, \\ \sigma_{26}^{(D)} &= R_{26}\sigma_{66}^{(M)} = \eta'(s)\sigma_{66}^{(M)}, \\ \sigma_{66}^{(D)} &= \sigma_{66}^{(M)}. \end{aligned} \tag{A.92}$$

Furthermore, if we take the step of directly substituting $\sigma_2^{(M)} \equiv \tilde{M}\sigma_1\tilde{M}^t$ into Equation A.91, we find:

$$\sigma_2^{(D)} = \tilde{D}\sigma_2^{(M)}\tilde{D}^t. \tag{A.93}$$

We can now use the sigma matrix decomposition of Equation A.71 to substitute for $\sigma_2^{(M)}$, to find:

$$\sigma_2^{(D)} = \tilde{D}\tilde{B}\tilde{A}S\mathcal{E}S^t\tilde{A}^t\tilde{B}^t\tilde{D}^t, \tag{A.94}$$

where we have extended the definition of the emittance matrix thus:

$$\mathcal{E} = \begin{pmatrix} \epsilon_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \epsilon_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \epsilon_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_l^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \left(\frac{\sigma_E}{E}\right)^2 \end{pmatrix}. \tag{A.95}$$

A.4.3 Dispersion in Linear Collider Final Focus Systems

Now let us consider a linear collider such as that depicted in the schematic of Figure 1.1, and let us assume that the design is such that, at the nominal entry to the final focus, R_{16} and R_{26} from the exit of the damping ring are both zero. We can define this

entry location to be a "treaty point" at which η and η' are also zero. While no "closed orbit" exists for a linear, single-pass system, we conventionally define the functions $\eta(s)$ and $\eta'(s)$ for this region as though the treaty point was a dispersion-free point in a storage ring, ie, as in Equation A.89. This is known as the *lattice dispersion* of the beamline.

We saw in the previous section that, at points where the on-energy and off-energy reference orbits coincide, the energy-position correlation terms ($\sigma_{i6}, i \neq 6$), are identically zero. While we have *defined* the dispersion functions at the entrance of the final focus to be zero, this does not affect the incoming beam, which may have nonzero energy-position correlations. This arises from imperfections in the real lattice. At this location, however, orbit deviations with changes in the centroid energy are not reliable measures of dispersion. For one thing, the presence of steering feedbacks will tend to eliminate the effects of such deviations. For another, the beam energy is generally changed by changing the accelerating properties of the linac at a given location; this location may be upstream or downstream from the source(s) of unwanted dispersion. Instead, we can examine the beam matrix itself, and use Equation A.92 to extract the *residual dispersion*:

$$\eta_x^{(0)} = \frac{\sigma_{16}}{\sigma_{66}},$$

$$\eta_x'^{(0)} = \frac{\sigma_{26}}{\sigma_{66}},$$

$$\eta_y^{(0)} = \frac{\sigma_{36}}{\sigma_{66}},$$

$$\eta_x^{(0)} = \frac{\sigma_{16}}{\sigma_{66}}. \tag{A.96}$$

These values can be used to determine a $\tilde{D}$ matrix, as in Equation A.94.

A.5 Chromaticity

The quantity known in closed storage rings as chromaticity is usually defined as the change in the tune around the ring with energy:

$$\xi \equiv \frac{\Delta\nu}{\Delta\delta} = \frac{\Delta\psi}{2\pi\Delta\delta}. \tag{A.97}$$

It is also legitimate, therefore, to define ξ as the change in phase advance of a section of beamline with energy. Using this definition, it can be shown [82] that:

$$\xi_{x,y} = -\frac{1}{4\pi} \int K(s)\beta_{x,y} ds, \tag{A.98}$$

where $K(s)$ is here understood to be positive for focusing in the plane of interest (in other words, for computing vertical chromaticity, horizontally-focusing quads are considered to have negative $K(s)$, vertically-focusing positive).

Let us consider the beam shape and phase advance at the IP or another waist point of the final focus system. At this point, $\alpha = 0$ and consequently $\gamma^* = \frac{1}{\beta^*}$. Using the Twiss-transport shown in Equation A.19, we can write an equation for the betatron function a distance s from the IP:

$$\beta(s) = \beta^* + \frac{s^2}{\beta^*}. \tag{A.99}$$

The expression for $\beta(s)$ can then be substituted into the defining equation for phase advance, Equation A.9, to yield a relation between the drift distance from the IP and the phase advance:

$$\Delta\psi = \tan^{-1}\frac{s}{\beta^*}, \quad \text{or} \quad s = \beta^* \tan\Delta\psi. \tag{A.100}$$

The value of $\Delta\psi$ can be replaced as shown in Equation A.97

$$s = \beta^* \tan(2\pi\xi\Delta\delta). \tag{A.101}$$

A simple reading of Equation A.101 would indicate that the waist shift is a periodic function of the energy offset of a given particle, and therefore that the waist

should be restored to its nominal position for particles with $\delta = \xi/2$. In practice, however, the waist shift is a monotonic function of energy offset. The reason for this is that as the energy of the beam changes, the betatron function in the quadrupoles changes. As Equation A.98 shows, this results in a value of ξ which is a function of energy. In practice, therefore, the energy-dependent waist shift can be approximated by Equation A.101 about the $\delta = 0$ point:

$$s \approx 2\pi\beta^*\xi\Delta\delta. \tag{A.102}$$

Note that Equation A.99 can also be used to obtain an expression for the linear beam size off the waist by a distance s:

$$\sigma^2 = \epsilon(\beta^* + \frac{s^2}{\beta^*}). \tag{A.103}$$

We can now combine Equations A.102 and A.103, and determine the beam size at the IP. location due to changes in the focal point from chromaticity:

$$\sigma^2(\delta) = \epsilon\beta^*[1 + (2\pi\xi\delta)^2]. \tag{A.104}$$

If we convolve $\sigma^2(\delta)$ with a Gaussian energy distribution of RMS width $\frac{\sigma_E}{E}$, we obtain an expression for the beam size at the nominal IP as a function of the emittance, betatron function, and energy width:

$$\sigma^2 = \epsilon\beta^* + 4\pi^2\xi^2(\frac{\sigma_E}{E})^2. \tag{A.105}$$

Equation A.105 is identical to Equation 1.5, with on exception: the definition of ξ used by Roy is greater than the definition of ξ used here by a factor of 2π. Consequently, a factor of $4\pi^2$ appears in Equation A.105 which is not present in Equation 1.5.

Appendix B

Measurements of FFTB Magnets

A considerable insight into the properties of accelerator electromagnets and their applicatons can be extracted from the two-dimensional multipole expansion of the magnetic field. In this Appendix, the expansion is presented in both polar and Cartesian coordinate systems. A description of the two techniques used to measure the multipole characteristics of the FFTB standard quadrupole and sextupole follows, with a discussion of the results of the techniques. The author is indebted to J.K. Cobb of SLAC for many fine papers on the theory and practice of magnetic measurements [83], [84], and also for his description of the rotating- coil technique [85].

B.1   Expansion of the Magnetic Field

A beamline magnet of the type used in the FFTB can be approximated as follows: between $l = 0$ and $l = l_0$, $\vec{B} = \vec{B}(r, \theta)$, while outside of these bounds $\vec{B} \equiv 0$; and $B_z = 0$ at all points. Within the z-axis bounds of the magnet, therefore, an appropriate coordinate system to use is planar polar coordinates. If we postulate that the magnetic field inside the aperture is constant in time at all points, and that the aperture contains no currents, then we are justified in defining a *magnetic scalar*

potential, Φ. This leads to the familiar equations for the magnetic field [1]:

$$\vec{B} = \vec{\nabla}\Phi, \text{ and} \tag{B.1}$$

$$\nabla^2\Phi = 0. $$

The equation for Φ can be solved in a straightforward fashion by the separation of variables technique; rejecting solutions which diverge at the origin of the coordinate system, we find:

$$\Phi = \sum_{n=1}^{\infty} K_n r^n \sin(n\theta - \alpha_n). \tag{B.2}$$

The magnetic field vector at any point inside the magnet can then be found by application of $\vec{\nabla}$:

$$\begin{aligned}\vec{B}(r,\theta) &= \hat{r} \sum K_n n r^{n-1} \sin(n\theta - \alpha_n), \\ &+ \hat{\theta} \sum K_n n r^{n-1} \cos(n\theta - \alpha_n).\end{aligned} \tag{B.3}$$

The relative strengths of the K_n, and the values of α_n, are determined by the boundary conditions. In order to understand these, it is instructive to rewrite $\sin(n\theta - \alpha_n)$ as a separate sine and cosine term, and to rewrite the first three terms of Φ in cartesian coordinates:

$$\begin{aligned}\Phi &= \sum K_{n,N} r^n \sin n\theta + \sum K_{n,S} r^n \cos n\theta \tag{B.4} \\ &= K_{1,N} y + K_{1,S} x \\ &+ 2K_{2,N} xy + K_{2,S}(x^2 - y^2) \\ &+ K_{3,N}(3x^2 y - y^3) + K_{3,S}(x^3 - 3y^2 x) + ...\end{aligned}$$

The components of Φ can be controlled by forcing the magnet to have an *equipotential surface* whose geometry corresponds to the desired term in Equation B.4. For example, a magnet which contains a nonzero $K_{1,N}$, and no other components, must have an equipotential surface at $y = \pm y_0$, where y_0 is some constant. Similarly, a

---

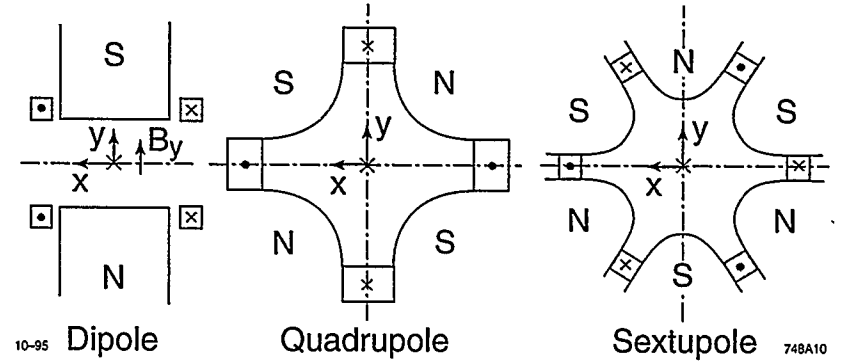[1]This treatment defines Φ opposite in sign to the usual convention.

Figure B.1: Shaping of iron pole-pieces required to generate pure dipole, quadrupole, or sextupole magnetic fields.

magnet which contains only $K_{2,N}$ must have an equipotential surface at $xy = \pm q_0^2$, where q is a constant. The equipotential surface in a warm-iron magnet is formed by shaping the iron pole-pieces, since to good approximation $\Phi = 0$ at the surface of the iron. Thus a magnet with hyperbolic pole-pieces will be a pure $n = 2$ pole, or a *quadrupole* magnet. Figure B.1 shows the pole-face shaping required to obtain a pure dipole, quadrupole, or sextupole magnet. All the magnets shown in Figure B.1 are so-called "normal" magnets, in which $K_{n,S} = 0$ for all n. To form a magnet with non-zero $K_{n,S}$ terms, a "skew" magnet, it is necessary to roll a normal magnet by an angle π/n, as Equation B.4 shows.

Let us now apply the $\vec{\nabla}$ operator in cartesian coordinates to the first 3 terms of Φ in Equation B.4, to obtain the magnetic field:

$$\begin{aligned}\vec{B}(x,y) &= \hat{y}\left[K_{1,N} + 2K_{2,N}x - 2K_{2,N}y + 3K_{3,N}(x^2 - y^2) - 6K_{3,S}xy + ...\right] \tag{B.5} \\ &+ \hat{x}\left[K_{1,S} + 2K_{2,N}y + 2K_{2,S}x + 6K_{3,N}xy + 3K_{3,S}(x^2 - y^2) + ...\right].\end{aligned}$$

Equation B.5 demonstrates immediately the essential properties of accelerator magnets. The $K_{1,N/S}$ terms constitute a constant kick in the horizontal and vertical, respectively; the $K_{2,N}$ terms provide the focusing effect of quadrupole magnets (in which the kick received by the particle is proportional to its distance from the origin),

and also contain the polarity of the focusing (in which a horizontallly-focusing magnet is vertically defocusing, and vice versa); the $K_{2,S}$ terms constitute a skew focusing, in which the horizontal kick is proportional to the vertical offset and vice versa. Finally, it can be shown from Equation B.5 that a beam offset horizontally in a sextupole will receive additional normal quadrupole kicks, while a beam offset vertically in a sextupole will receive additional skew quadrupole kicks.

B.1.1 Additional Definitions

Real magnets in accelerators act on the beam over a finite length. For this reason the *integrated field*, $\int \vec{B}dl$, is usually used to characterize such magnets. We can take advantage of the longitudinal symmetry postulated above and rewrite Equation B.5:

$$\int \vec{B}(x,y)dl = \hat{y}\left[B_{0,y}L + G_N x - G_S y + S_N(x^2 - y^2) - 2S_S xy\right] \quad (B.6)$$
$$+ \hat{x}\left[B_{0,x}L + G_N y + G_S x + 2S_N xy + S_S(x^2 - y^2)\right],$$

where the length of the magnet is L. The coefficients $nLK_{n,N/S}$ in Equation B.5 have been replaced in Equation B.6 with a series of integrated quantities: the *integrated gradient*, $G_{N/S}$, and the *integrated sextupole*, $S_{N/S}$. The integrated sextupole is defined here such that, at a radius r_0, the $\int Bdl$ from the sextupole field is given by $S_N r_0^2$. The *integrated second derivative* of the magnet is twice the value of $S_{N/S}$, and this quantity is also frequently used to specify the strength of a sextupole magnet, or the sextupole content of a multipole magnet.

The choice of coordinate system also provides three degrees of freedom: the coordinate system in which the magnetic field expansion is performed may be selected by moving the origin and rotating about the longitudinal axis, yielding x, y, and θ degrees of freedom. For magnets which are intended to function as quadrupoles, the x and y coordinates are typically chosen such that the dipole field component vanishes, and the roll is chosen such that the value of G_S vanishes:

$$\int \vec{B}(x,y)_{quad}dl = \hat{y}\left[G_N x + S_N(x^2 - y^2) - 2S_S xy\right] \quad (B.7)$$
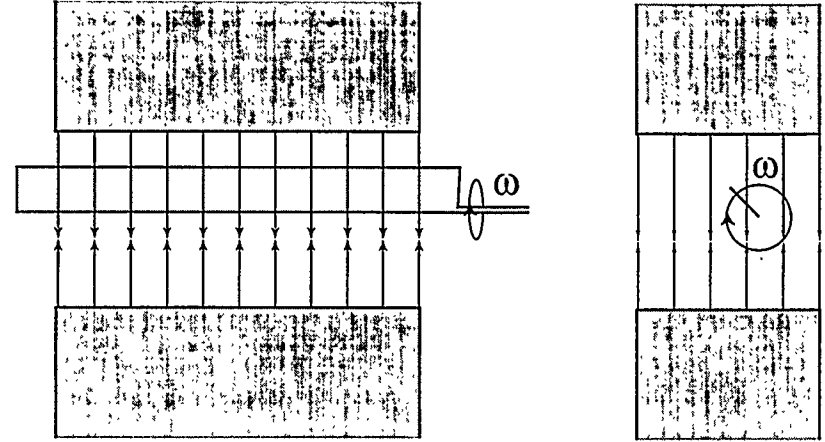
Figure B.2: A rotating coil in an arbitrary magnetic field.

$$+ \hat{x}\left[G_N y + 2S_N xy + S_S(x^2 - y^2)\right].$$

For sextupole magnets the x and y coordinates are chosen such that the G_N and G_S vanish, while the roll is chosen such that S_S vanishes:

$$\int \vec{B}(x,y)_{sext}dl = \hat{y}\left[B_{0,y}L + S_N(x^2 - y^2)\right] \quad (B.8)$$
$$+ \hat{x}\left[B_{0,x}L + 2S_N xy\right].$$

With these definitions in hand, we may now examine the two methods used to measure the FFTB magnets.

B.2 Rotating Coil Measurements

Consider a situation as depicted in Figure B.2: A rotating coil of length l_1 and radius r_1 is inserted into the aperture of a magnet of length L, and rotated asymmetrically (i.e., one end of the coil moves while the other is stationary) with frequency ω. If we define the stationary leg of the coil to be at the (x,y) origin of coordinates, and leave

the $\theta = 0$ rotation arbitrary, then the total magnetic flux through the coil when it sits at angle θ_1 can be computed from Equation B.3:

$$\Phi_B(\theta = \theta_1) = \int_{l=0}^{l=L} dl \int_{r=0}^{r=r_1} dr B_\theta(r, \theta_1) \qquad (B.9)$$
$$= L \sum K_n r_1^n \cos(n\theta_1 - \alpha_n).$$

Because the coil is rotating with frequency ω, we can replace θ_1 with ωt_1. Furthermore, the rotating coil gives rise to an induced EMF:

$$V = -\frac{d\Phi_B}{dt} = L \sum n K_n r_1^n \omega \sin(n\omega t - \alpha_n). \qquad (B.10)$$

Equation B.10 shows that the induced EMF is a function of time, and that the the $2n$-pole moment of the magnet generates an EMF which oscillates with a frequency of $n\omega$. The relative contributions of the different multipole moments can be deduced by separating the different frequency components of the EMF signal as the coil rotates in the magnet aperture.

The extraction technique for the coefficients used for the FFTB magnets has been in regular use at SLAC for many years [84]. The loop is rotated at a frequency which is carefully maintained at a constant value, and the potential difference around the loop is measured at M intervals equally spaced around the full circle of its rotation. Because the rotation speed is constant, the M equally-spaced measurements of V constitute a series of samples equally separated in time. By performing a fast Fourier transform (FFT) of the voltages measured at the M intervals, therefore, the multipole components up to $n = M/2$ can be computed.

In an ideal situation, the $M/2$ Fourier coefficients can be extracted from the rotating-coil measurement without any interference between coefficients. In a real-world situation, however, precision becomes limited by the precision with which V can be measured. In particular, a very large signal from a low- frequency component can force the use of a coarse voltmeter setting, which in turn results in poor signal-to-noise performance for the smaller signals of high-frequency components. This is especially a problem if the coil center does not precisely match the desired center
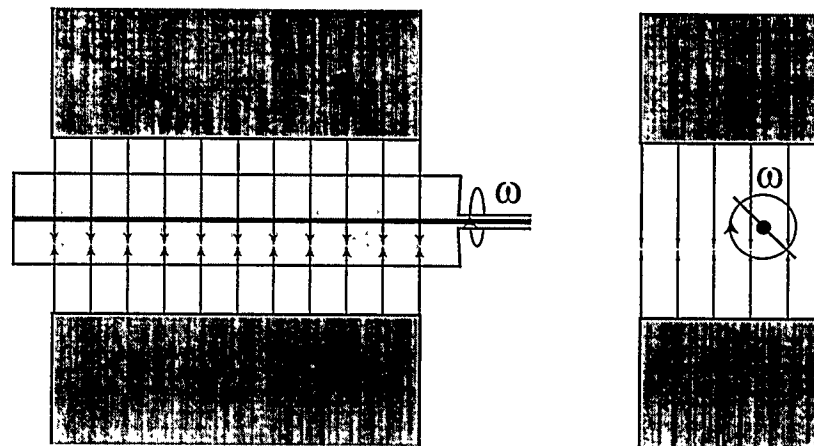
Figure B.3: A two-coil system used to null EMF from odd-harmonic magnetic fields; such a coil gives higher precision measurement of the quadrupole field.

of the magnet (null-bend point for quads or null-gradient point for sextupoles). If a particular harmonic is to be measured with high precision, an appropriate coil design can improve the signal- to-noise behavior. For example, precision quadrupole measurements can be made by constructing two coils with a common axis, as shown in Figure B.3. In this case, the sum of the two coil voltages is measured. In this case, the voltage measured will be:

$$V = L \sum n K_n r_1^n \omega \left[\sin(n\omega t - \alpha_n) + \sin(n(\omega t + \pi) - \alpha_n) \right]. \qquad (B.11)$$

For even harmonics, such an arrangement doubles the signal, while for odd harmonics the sum of the two sinusoidal terms cancels. For measuring the quad term, this arrangement rejects the misalignment background and any other backgrounds at the rotation frequency (such as motor-induced backgrounds), but also rejects the sextupole, decapole, and other odd harmonic multipole terms. This is the arrangement used to measure the quadrupole excitations (G_N versus current) for the FFTB quads.

Figure B.4 shows a typical result for the measurd G_N for an FFTB quadrupole, as a function of the excitation current I_{ex}. A coil of radius 8.66 mm was used for
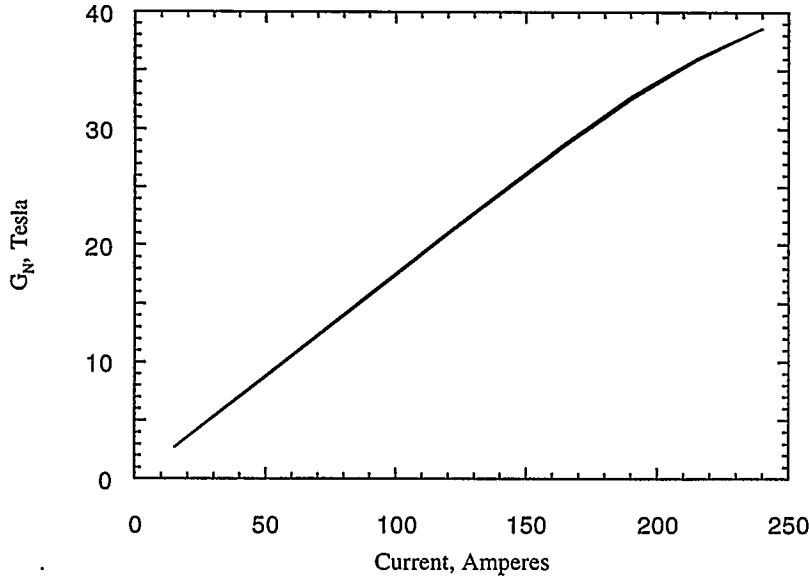
Figure B.4: Magnetic field hysteresis curve for a typical FFTB standard quadrupole. The lower curve represents the rising-current measurement.

the measurement. The expected statistical error on the measurement of each value of G_N is 0.02% of the value itself, and comes from the resolution of the voltmeter, the precision of the rotating coil frequency, and the stability of the quadrupole power supply; while the systematic error is expected to be 0.04% of the value of G_N, arising primarily from possible calibration errors in the coil and the power supply transductor [85]. Figure B.5 shows the RMS distribution of G_N values at each current used for measurement.

Let us consider the first two terms in the harmonic expression of Equation B.10, which in the case of the two-coil system are quadrupole and octupole:

$$V = L \left[2K_2 r_1^2 \omega \sin(2\omega t - \alpha_2) + 4K_4 r_1^4 \omega \sin(4\omega t - \alpha_4) \right] + \dots \qquad (B.12)$$

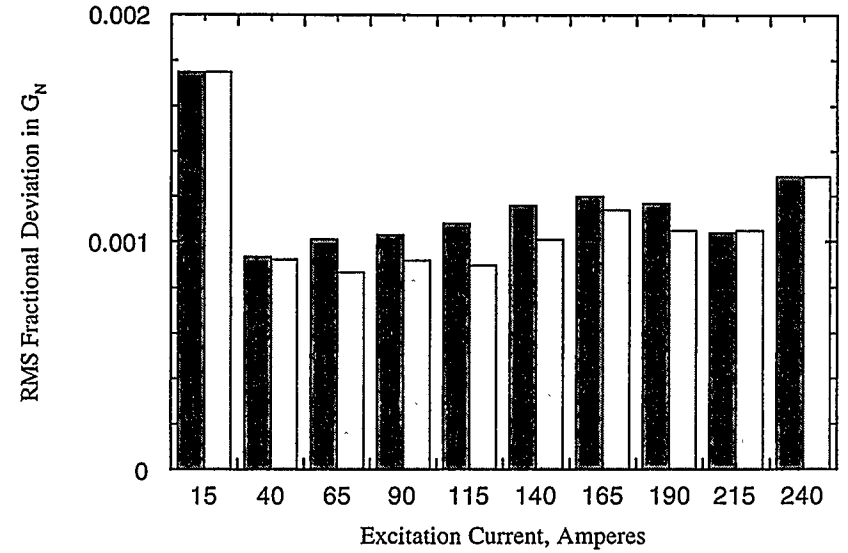We can rewrite this with G_N and the integrated octupole, O, and define the coordinate

Figure B.5: RMS deviations of the quadrupole field measurements for all currents. Deviations are shown as fraction of average strength at the given current. Both rising-leg (dark) and falling-leg (light) are shown.

system such that $\alpha_2 = 0$:

$$V = G_N r_1^2 \omega \sin 2\omega t + O r_1^4 \sin(4\omega t - \alpha_4) + \dots \qquad (B.13)$$

The magnitude of the integrated quadrupole field at $r = r_1$ is given by $G_N r_1$, and that of the integrated octupole field is given by $O r_1^3$. Let us assume that the resolution of 2×10^{-4} given for G_N corresponds to a 2×10^{-4} resolution of the integrated quadrupole field at r_1. Furthermore, let us assume that this limitation implies the overall statistical limitation of the measurement of the integrated magnetic field at r_1. It therefore follows that the magnetic field due to octupole moment at $r = r_1$ can be measured to 0.02% of the magnetic field due to quadrupole moment at $r = r_1$. The tolerance on higher-order multipole fields for the standard quads is 0.1% of the quadrupole field at 70% of the aperture. The coil is capable of measuring a higher-order even multipole field equal to 0.02% of the quadrupole field, and its radius is

75% of the aperture. Because the multipole fields are expressed as a ratio between two fields, the systematic errors cancel out in this case. Out of 174 multipole content measurements on the FFTB standard quads (6 multipoles on each of 29 quads), only 3 were found to be at the tolerance stated above, with all others significantly lower.

The primary purpose of the rotating-coil measurements was to determine the quadrupole strength as a function of power supply current, in order to derive a polynomial for use in the control system. Secondarily, the measurement was to measure the even-harmonic multipole moments, especially the "allowed" harmonics. "Allowed" multipole moments are those which arise from construction phenomena which affect all poles symmetrically. For example, the pole-faces in the FFTB quadrupoles are not infinite in the vertical or horizontal. The truncation of those pole-faces leads to 12-pole, 20-pole, and 28-pole moments. As described above, the "allowed" multipole contents were found to be quite small relative to the construction tolerances. The sextupole component not an allowed harmonic, and is therefore expected to be small. However, the sextupole field is proportional to r^2, which is the lowest power or r after the quadrupole. Therefore, relatively small construction errors at the pole-face will translate to larger sextupole errors near the center of the magnet, where the beam is present.

After the rotating-coil measurements described above, each FFTB standard quad was disassembled in order to insert its BPM, then reassembled. The reassembled quad was then placed on a test-stand which was used to simultaneously calibrate the BPM, measure the quadrupole field to ensure that the magnet had reassembled properly, measure the magnetic center relative to the mechanical center of the magnet, and measure the sextupole component. This set of measurements is described below, save for the BPM calibration which is described in Appendix C.
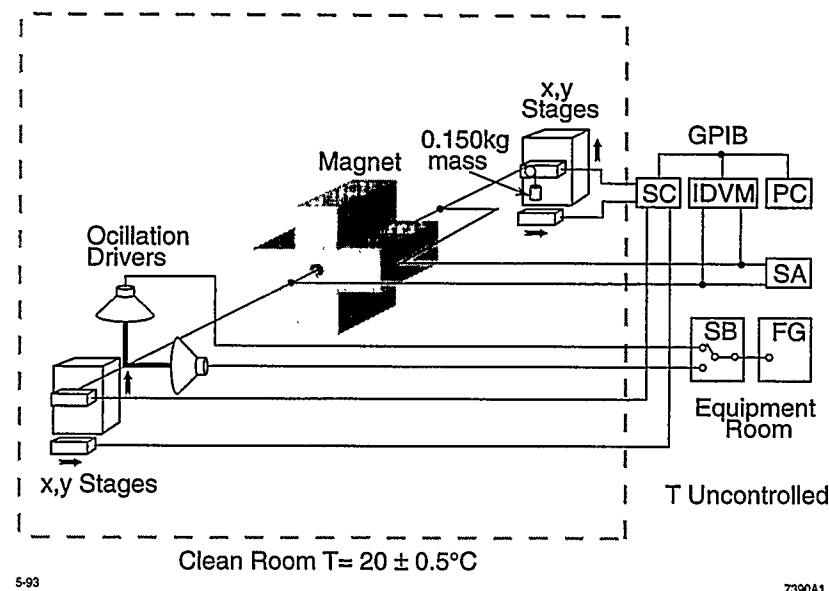
Figure B.6: Schematic of the stretched-wire measurement apparatus, including Stage Controller (SC), Integrating Digital Voltmeter (IDVM), Spectrum Analyzer (SA), Frequency Generator (FG) and horizontal/vertical driver Switch Box (SB). The mounting table and Coordinate Measuring Machine (CMM) are not shown.

B.3 Stretched-Wire Measurements

Figure B.6 shows a schematic diagram of the apparatus used for the stretched-wire measurement. A gold-plated tungsten wire 35 μm in diameter is strung through the magnet aperture, and attached to a 150 gram mass; the resulting wire tension is near the breaking tension. The resulting wire sag over the full 1.8 meter length of the wire is 12 μm; however, the wire position is measured at stations at either end of the magnet, and the sag between the end of the magnet and its center is close to 1 μm. The wire is made parallel to the magnet at the 0.05 mrad level via a Coordinate Measuring Machine (CMM), not shown in Figure B.6.

Once the wire is installed, the oscillation drivers are activated, and driven by a

frequency-generator at the wire resonant frequency of 83 Hz. With the quadrupole magnet excited by 165 amperes, the vibration results in an EMF across the wire. Like the rotating-coil method described above, the EMF at the driving frequency is due to the dipole term in the multipole field, and the second harmonic (166 Hz) is due to the quadrupole term. The EMF is monitored by attaching the wire to a spectrum analyzer (SA). By exciting the vertical driver and moving the wire in the horizontal until the 83 Hz component of the signal is nulled, the horizontal center of the quadrupole can be located. Similarly the horizontal driver allows determination of the vertical center. A modest extension of this technique [86] allows determination of the magnetic center in x, y, and z. The coordinate system of the computer-controlled motion stages is then set to the point determined by this procedure.

In order to measure the magnetic field and sextupole content, the stretched wire is moved (via the stages) by a known vertical or horizontal distance about a central point (x, y). The wire is in this case connected to an integrating digital voltmeter (IDVM). Given a magnetic field described by Equation B.7, the integrated voltage can be written as a function of the sweep center (x, y), the total sweep distance δ_x in the horizontal and δ_y in the vertical, and the field components:

$$
\begin{aligned}
\int V dt \;=\; & \delta_x \left[G_N x + S_N (x^2 - y^2) - 2 S_S xy + \frac{S_N \delta_x^2}{12} \right] \\
- \; & \delta_y \left[G_N y + 2 S_N xy + S_S (x^2 - y^2) + \frac{S_S \delta_y^2}{12} \right].
\end{aligned}
\tag{B.14}
$$

The grid of measurements used is shown in Figure B.7. A grid of 25 center points, from -4 mm to +4 mm in x and y, with a horizontal sweep of 2 mm and a vertical sweep of 2 mm, gives a total of 50 measurements, with the wires never more than 5 mm from the center of the magnet in x or y. The total grid of 50 measurements was repeated 10 times in order to determine the repeatability of individual measurements and to allow sufficient redundancy to perform cuts. Once cuts to the data had been applied and appropriate error-bars for each data point established, the map of $(x, y, \delta_x, \delta_y, \int V dt)$ data was fitted to Equation B.14 via a MINUIT-mediated χ^2 minimization. In the
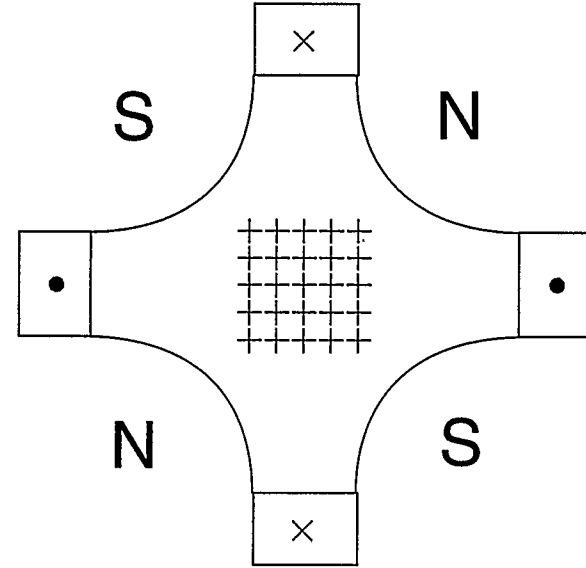
Figure B.7: Arrangement of the 50 stretched-wire sweep measurements in the quadrupole aperture.

fit, horizontal and vertical offsets, and roll angles, are included as fit parameters.

In order to determine some possible sources of noise in the system, several test measurements were performed. These tests are described below.

B.3.1 Test of the Stretched-Wire System

The first test of the stretched-wire system was a null-field test: the system was installed as shown in Figure B.6, the vibrating-wire operation was performed to locate the center of the magnet, and the magnet was then turned off. After a suitable period of time, the measurement sequence described above was initiated. This would allow determination of the background noise in the system with no magnetic field present.

For each of the 50 points in the $(x, y, \delta_x, \delta_y)$ grid, the average $\int V dt$ over 10 measurements and the RMS were determined. These are plotted in Figure B.8. Because
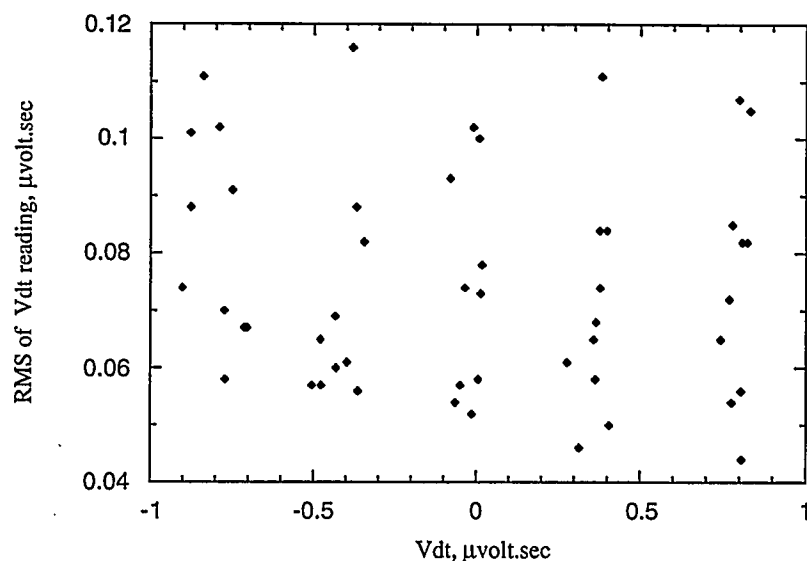
Figure B.8: Distribution of average integrated voltage and RMS deviation from same over the grid of null-field measurements described in the text.

all the values are clustered about an RMS width of 0.080 $\mu V \cdot$ sec, it follows that the background noise in the system, with the magnet power supply deactivated, is approximately 0.080 $\mu V \cdot$ sec.

A second, complementary test of the system was also performed. In this test the magnet power supply was activated, the magnet excitation current was set to 165 amperes, and the wire moved to the center of each sweep on the grid, but no actual sweep was performed. Instead the IDVM integrated while the wire was nominally stationary, and the 250 resulting $\int V dt$ measurements were expected to be consistent with zero, with some RMS.

When the stationary-wire test was performed, it was determined that the average of the measurements was indeed close to zero, but the RMS in this case was 0.38 $\mu V \cdot$ sec, significantly larger than in the zero-field case. Why this should be the case is not known. One possibility is that the supposedly-stationary wire is in fact moving

in small increments as the feedback loop for the wire stages operates, but in this case the signal strength at each point should correlate with its position, which was not seen. A second stationary-wire test, with the magnet power off, revealed similar results but a smaller, 0.26 $\mu V \cdot$ sec RMS value for $\int V dt$.

An additional feature of the powered stationary-wire test was a series of "noise" events, in which the value of a measurement at a given location was several $\mu V \cdot$ sec different from the average at that location. This feature was not seen with the magnet power supply off, and was believed to be caused by other electrical devices in the vicinity briefly drawing power away from the magnet power supply. The noise events were found to appear with a simple periodicity in the measurement process. When the length of time required for a measurement was increased, the frequency of the noise events was observed to decrease, indicating that the events were periodic in time and not in measurement number.

B.3.2 Results of the Stretched-Wire measurement

The FFTB standard quads were measured in the stretched-wire technique at an excitation of 165 amperes, to be consistent with one of the data points taken with the rotating coil. At each point in the grid, 10 x-sweep measurements and 10 y-sweep measurements were taken, for a total of 500 points. The 10 x-sweep measurements were then combined to form an average and an RMS, after filtration of noise events such as those described above. This reduced the total dataset to 50 measurements and 50 RMS values.

A MINUIT fitting engine was then employed to fit the parameters in Equation B.14 to the 50 data points, using the RMS values divided by $\sqrt{m}$ as uncertainties, where m is the number of readings averaged to give a single data point. Because of the filtration of noise events, m is not constant for all data points. It is interesting to note that the RMS values observed at each data point were in general consistent with the 0.08 $\mu V \cdot$ sec observed for the zero-field swept-wire measurement described previously,

and not with the larger values observed for the stationary-wire measurements.

Early results with the fit as described were not encouraging. A search for systematic errors determined that the G_N value obtained using only horizontal sweep data was different from the value when only vertical sweep data was used. The fit was expanded to include a seventh parameter, a global scale factor by which horizontal position data was multiplied to make it consistent with the vertical data. This value was found to oscillate about 0.997, and the convergence was improved with its inclusion, but the value of χ^2/ν was still too high.

An additional source of error was speculated to be the stages and/or their controller. It was presumed that the stages might have some overall positioning accuracy over millimeters of motion which was larger than the micron-sized precision of the system: over the 10 measurements of a given $(x, y, \delta_x, \delta_y)$ sweep, the stages were reproducibly returning the stages to the same location, but that location was not the location expected by the measurement program or reported by the stage controller. An *ad hoc* assumption of a 3 μm positioning error in each plane was found to fit the data relatively well, and also to fit the data obtained in the CCS sextupole measurements (see below). This overall accuracy number is a systematic error in that it cannot be reduced through additional measurements of the same data; consequently it forms the main source of error in the measurement. It is worth noting that magnets which were measured under cool-weather conditions reached χ^2/ν values much lower than 1, indicating that the temperature in the equipment room was a factor in this source of systematic error.

Standard Quadrupole Measurements

In addition to the systematic errors listed above, six of the FFTB standard quads measured by this procedure had extremely anomalous results: while nothing qualitatively wrong appeared in the data, these magnets reported a very large sextupole component, a scale-factor quite different from 1, and a poor χ^2/ν value (1000-2000
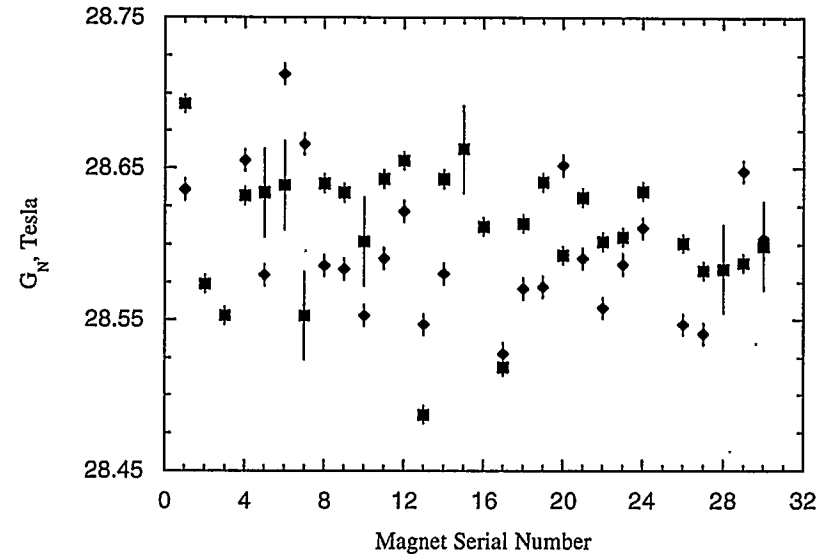
Figure B.9: Comparison of the stretched-wire G_N measurements (diamond) with the rotating-coil values (square).

per 43 degrees of freedom). The cause of this difficulty was never isolated, and there was some concern that the cause was an actual problem with the magnets, especially since five of the six magnets were consecutive in serial number (specifically magnets 2, 4, 5, 6, and 7; magnet number 3 was never measured). However, magnet number 23 was originally found to be anomalous, but was remeasured and found to be normal. This lends weight to the possibility of an undetermined measurement error. Magnet 29 was also anomalous, but no remeasurement was made. Since 25 magnets were measured, a total of 19 yielded non-anomalous results.

Figure B.9 shows the measured G_N values for the FFTB standard quads with 165 amperes excitation current, from both the stretched-wire (diamond) and rotating-coil (square) measurements. The rotating-coil measurements have been scaled from the precise current at which they were performed to the precise current at which the stretched-wire measurements were made for direct comparison. In several cases, the
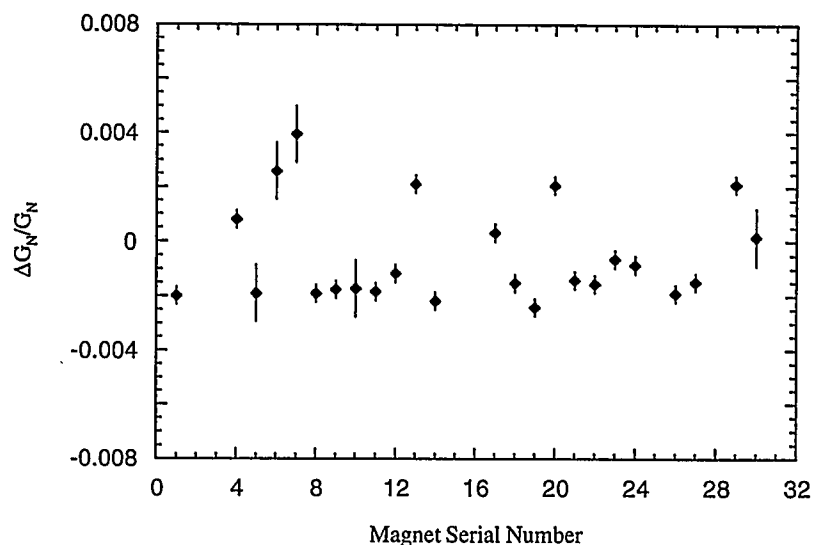
Figure B.10: Fractional deviation between the rotating-coil and stretched-wire measurements of G_N for FFTB standard quads.

stretched-wire currents were not properly recorded, and the average stretched-wire current of 165.845 ± 0.185 amperes was used; this uncertainty in the current results in the large uncertainties on several of the scaled rotating-coil measurements. The resolution of the stretched-wire system is 0.007 T/m, or 0.025%.

Figure B.10 shows the relative difference in the two measurements: $[G_N(\text{wire}) - G_N(\text{coil})]/G_N(\text{coil})$. Overall, a scale factor of 0.2% is indicated, with the rotating coil reporting the larger strength number. Some fliers exist, indicating that the systematic discrepancy is not quite so simple. In addition to the 0.3% systematic error from the scale factor, other possible sources of error include: the transductor used to measure the magnet current; an overall scale factor to the wire-mover stages; and hysteresis-related errors in scaling the rotating-coil measurements to the stretched-wire currents. Nonetheless, the $\pm 0.2\%$ agreement indicated in Figure B.10 indicates that no unacceptable damage occurred to the quads during BPM installation, and
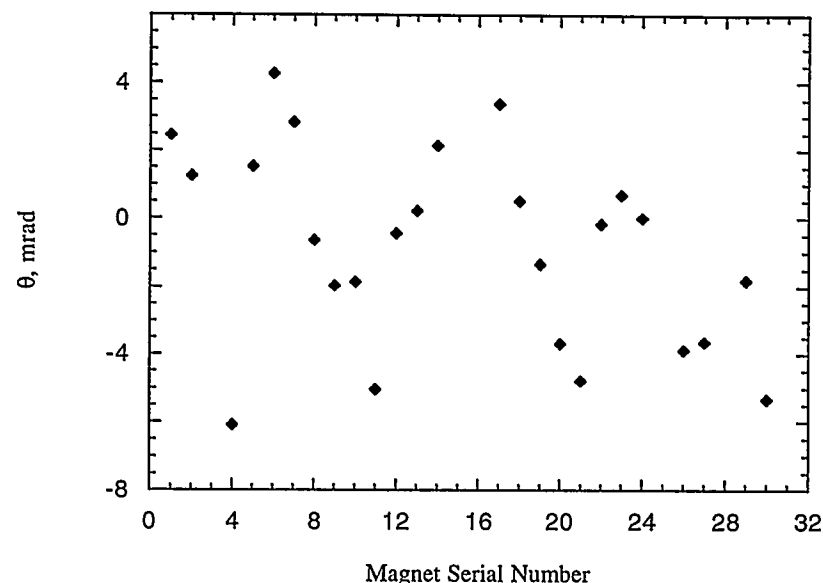
Figure B.11: Measured quadrupole roll angles, in milliradians. The distribution is more indicative of the precision with which the magnet can be placed on the test stand than of any characteristics of the magnet itself.

that the stretched-wire technique used here is an adequate one, with potential for improvement through vigorous elimination of systematic errors.

Figure B.11 shows the fitted quadrupole roll angles relative to the coordinate system of the wire stages. The resolution of this measurement is close to 80 μrad. It was initially hoped that the fitted roll angles could yield some insight into the orientation of the quadrupole field with respect to the mechanical split-planes of the magnet; however, the few magnets which were repeatedly measured did not yield repeatable angle measurements. The roll angle is therefore primarily a measure of the accuracy with which the magnets can be installed on the test stand. The extremely high resolution of this technique, however, makes it a potentially useful one in developing a planarized NLC final focus at the milliradian level, as discussed in Chapter 7.

Finally, Figure B.12 shows the overall sextupole content of each of the 19 magnets
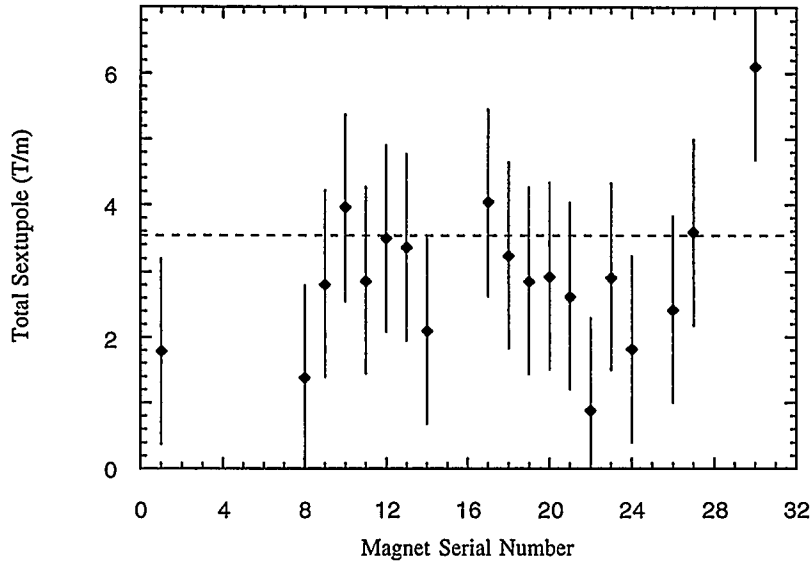
Figure B.12: Total sextupole content of the FFTB standard quads, as measured by the stretched-wire technique. Dashed line represents the construction tolerance described in the text.

with good convergence: $S_{TOT} = \sqrt{S_N^2 + S_S^2}$. The resolution of each of the normal and skew components was approximately 1.0 T/m, for a resolution of S_{TOT} of 1.4 T/m. The dotted line represents the tolerance value, 3.5 T/m, corresponding to a magnetic field from sextupole component equal to 0.1% of the field from quad component at a radius which is 70% of the aperture. All but 1 of the magnets lie at or below the tolerance, to within the measurement error. Note, however, that the construction tolerance is approximately twice the optical tolerance established in Chapter 2. Because the optical tolerance was set for a 2% increase in spot size, the logic of quadrature arithmetic indicates that the spot size increase from sextupole aberrations at the construction tolerance is close to 8%. It is worthwhile to note that the normal and skew sextupole components did reproduce on magnets which were removed from the test stand and later reinstalled.
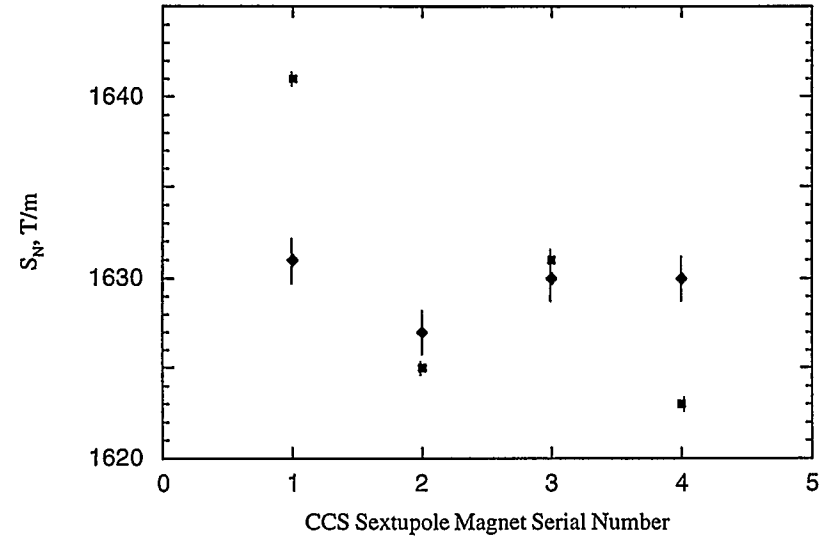
Figure B.13: Measurements of the CCS sextupole magnet S_N values from stretched-wire technique (diamonds) and from rotating-coil technique (squares). The discrepancies are larger than the discrepancies for the quads; however, in the absence of precise excitation current data for the stretched- wire measurements, the discrepancy is equivocal.

CCS Sextupole Measurements

The four CCS sextupoles were measured using the stretched-wire technique, with a sextupole excitation current of 215 amperes. The actual currents were not recorded, and therefore no direct comparison with the rotating-coil are possible. However, figure B.13 shows the stretched-wire measurements (diamonds) and the rotating-coil measurements scaled to 215.0 amperes (squares). The agreement is good at the level of 0.5%. The resolution of the sextupole component is about 1.2 T/m; at the excitation level of 1630 T/m, this corresponds to a resolution of 0.07%. The value of χ^2/ν fell between 0.8 and 1.2 for the four measurements with 3 μm positioning errors assumed. The roll angles were all within 2.5 milliradians of zero.

The symmetries of a sextupole magnet do not require that the dipole field and quadrupole field vanish at the same point in the magnet. In the fit of the sextupole

magnets, the fit was configured to adjust the origin of coordinates such that the quadrupole terms were nulled, and to fit horizontal and vertical dipole kicks at the origin. The largest such field observed was $6.7 \times 10^{-4} \mathrm{T} \cdot \mathrm{m}$, corresponding to a kick to the 46.6 GeV electron beam of 4.3 μrad. The resolution of the dipole field fit was $1.0 \times 10^{-4} \mathrm{T} \cdot \mathrm{m}$.

Appendix C

Beam Position Monitor Calibration

The FFTB Beam Position Monitor (BPM) system is a crucial element of numerous tuning and stabilization algorithms, including the launch and energy feedbacks, beam-based alignment, dispersion, lattice diagnostics, wire scanner jitter correction, and CCS stabilization feedbacks. In future linear colliders the BPMs are expected to be even more important, as they can monitor the behavior of every pulse without impacting the luminosity.

Because of its importance to the FFTB and the future linear collider, a procedure for bench-testing the FFTB BPMs was developed and implemented as part of the overall magnet fiducialization described in Appendix B. In order to understand the bench-testing, it is also necessary to summarize the theory and practice of stripline beam position monitors.
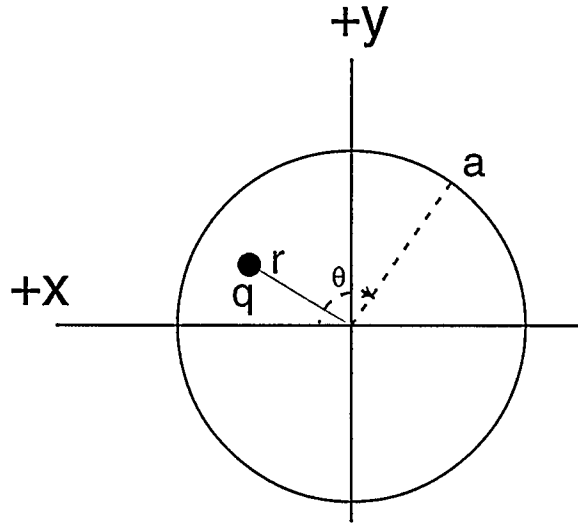
Figure C.1: Infinite line charge in a metal pipe of radius a. The induced surface charge at locations (a, θ) around the pipe is a function of the charge q and its position (r, ϕ).

C.1 Principle of Operation of Stripline Beam Position Monitors

Consider the situation depicted in Figure C.1: an infinite line charge passes through a metal pipe of radius a; its position in the pipe's coordinate system given by (r, ϕ). The surface charge-density σ_Q at all points (a, θ) on the pipe can be derived in a straightforward manner from Poisson's equation [87]:

$$\sigma_Q(a, \theta) = \frac{q}{2\pi a} \frac{a^2 - r^2}{a^2 + r^2 - 2ar\cos(\theta - \phi)}, \tag{C.1}$$

where q is the charge per unit length of the source term. Equation C.1 can be rewritten as an expansion in powers of $\frac{r}{a}$ by performing a Fourier cosine transform:

$$\sigma_Q(a, \theta) = \frac{q}{2\pi a} \left[1 + 2 \sum_{n=1}^{\infty} \left(\frac{r}{a}\right)^n \cos n(\theta - \phi) \right]. \tag{C.2}$$

Equation C.2 can be further transformed by noting that, in the coordinate system in use in Figure C.1, $x = r\cos\phi$, $y = r\sin\phi$, where x and y are the cartesian coordinates of the line charge in the pipe:

$$
\begin{aligned}
\sigma_Q(a, \theta) = {} & \frac{q}{2\pi a} \Bigg\{ 1 + 2 \left[\frac{x}{a}\cos\theta + \frac{y}{a}\sin\theta \right] \\
& + 2 \left[\frac{x^2 - y^2}{a^2}\cos 2\theta + 2\frac{xy}{a^2}\sin 2\theta \right] \\
& + 2 \left[\frac{x^3 - 3xy^2}{a^3}\cos 3\theta + \frac{3yx^2 - y^3}{a^3}\sin 3\theta \right] + \ldots \Bigg\}.
\end{aligned}
\tag{C.3}
$$

Let us now consider only the charge density induced at four points around the pipe wall, specifically those points which are the furthest North, South, up and down (denoted henceforth by N, S, T, B). These directions correspond to $\theta = 0, \pi, \frac{\pi}{2}, \frac{3\pi}{2}$, respectively. Table C.1 shows the contributions to the charge density at each of these points from various orders of the expansion in Equation C.3, normalized to $\frac{q}{2\pi a}$. Table C.1 is instructive in that it indicates how knowledge of the surface charge distribution can allow estimation of the location of the source charge in the pipe. The first-order expressions for the beam horizontal and vertical positions are given by:

$$x_1 = \frac{a(\text{N} - \text{S})}{2(\text{N} + \text{S})}, \qquad y_1 = \frac{a(\text{T} - \text{B})}{2(\text{T} + \text{B})}. \tag{C.4}$$

Furthermore, the sum N+S+T+B gives an excellent estimate for the charge q, in that all x and y dependence cancels out of this summation.

Let us examine the approximations in Equation C.4 more carefully. The expression for x_1 yields:

$$
\begin{aligned}
x_1 &= x\frac{1 + \frac{x^2 - 3y^2}{a^2} + \ldots}{1 + 2\frac{x^2 - y^2}{a^2} + \ldots} \\
&= x\left[1 - \frac{x^2 + y^2}{a^2} + \ldots \right].
\end{aligned}
\tag{C.5}
$$

A similar expression holds true for y_1:

$$y_1 = y\left[1 - \frac{x^2 + y^2}{a^2} + \ldots \right]. \tag{C.6}$$

Order	North	South	Top	Bottom
0	1	1	1	1
1	$\frac{2}{a}x$	$-\frac{2}{a}x$	$\frac{2}{a}y$	$-\frac{2}{a}y$
2	$\frac{2}{a^2}(x^2-y^2)$	$\frac{2}{a^2}(x^2-y^2)$	$-\frac{2}{a^2}(x^2-y^2)$	$-\frac{2}{a^2}(x^2-y^2)$
3	$\frac{2}{a^3}x(x^2-3y^2)$	$-\frac{2}{a^3}x(x^2-3y^2)$	$\frac{2}{a^3}y(3x^2-y^2)$	$-\frac{2}{a^3}y(3x^2-y^2)$

Table C.1: Contributions to charge density at North, South, Top, Bottom locations in units of $\frac{q}{2\pi a}$.

At a given location (x,y), therefore, the approximation of the source-charge location in Equation C.4 is good up to an overall scale factor; and the scale factor is identical for the two cartesian axes. The error induced by this method of position determination, therefore, is in its radial distance from the center of the pipe; the angle within the pipe is determined correctly. Furthermore, a second approximation of the form:

$$x_2 = x_1\left(1 + \frac{x_1^2 + y_1^2}{a^2}\right), \qquad y_2 = y_1\left(1 + \frac{x_1^2 + y_1^2}{a^2}\right) \qquad \text{(C.7)}$$

will result in an overall radius error which is on the order of $\left(\frac{r}{a}\right)^4$. Note that while the sum N+S+T+B contains no dependence on the location of the charge, using this sum in the denominator of Equation C.4 would result in a more complicated form for the nonlinearity correction, including an error in the angle determination.

While the computation of position estimates from surface charge densities is of academic interest, surface-charge meters are not a standard component of a scientist's toolbox. In order to construct an actual BPM, it is necessary to take advantage of the pulsed nature of the beam. When the beam is not present in the pipe in Figure C.1, the surface charge density on the pipe is isotropic; when beam is present, the charge is redistributed as shown in Equation C.1. Therefore, the arrival of the beam must cause the charge to flow between the two distribution states, and this flow of charge can itself be monitored.
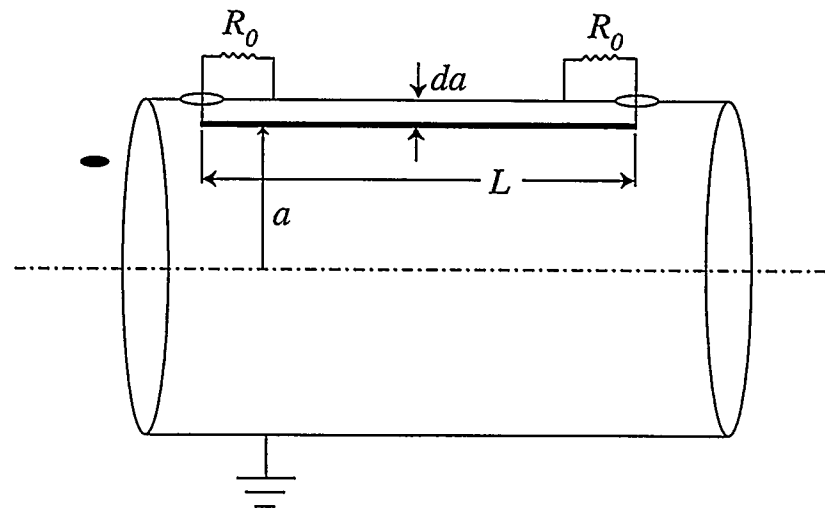
Figure C.2: Schematic diagram of a stripline BPM. A strip of length L is inserted into a metal pipe a distance a from the centerline. The gap between the strip and the wall, da is determined to give an impedance Z_0 equal to R_0. For simplicity only one strip is shown.

Consider a situation as shown in Figure C.2: a stripline is inserted into the beam pipe a distance a from the center of the pipe; the distance between the stripline and the pipe wall, da, is adjusted to give an overall impedance Z_0 to the stripline-pipe system. The strip is terminated at either end with resistances $R_0 = Z_0$.

In the absence of the beam, the stripline is free of surface charge. When the beam arrives, a surface charge is attracted onto the stripline and the pipe in order to maintain overall electrical neutrality; in the case of the stripline, the charge is transported through the resistor. The current flowing through the resistor causes a voltage to appear across the resistor, which can be monitored by a voltage-measuring device. The time-structure of the voltage surge is close to a delta function, and will be so treated in the subsequent formalism. In the case of an electron bunch, the stripline must be at a voltage below ground to attract positive charge, and so the surge is negative in polarity.
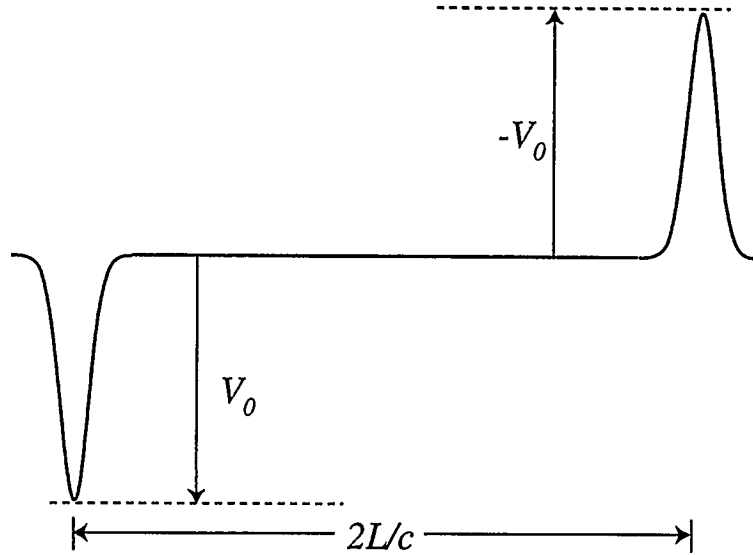
Figure C.3: The voltage detected at the upstream end of a BPM stripline consists of a delta-function pulse of amplitude V_0, followed by a pulse of amplitude $-V_0$ a time $2L/c$ later.

The beam propagates down the pipe at the speed of light, and the surface charge on the stripline (ideally) does so as well. At the downstream end, the surface charge is returned to the ground through the downstream termination, causing a voltage delta function which is equal and opposite to the initial pulse. This voltage propagates back up the stripline to the upstream resistor. As a result, the voltage-measuring device detects a negative voltage pulse and a positive voltage pulse, separated in time by $\frac{2L}{c}$, as shown in Figure C.3. The relative amplitude of the surges as a function of stripline position and beam position are as shown in Equation C.1. While this description of a BPM stripline has assumed a matched termination at the downstream end, it can be shown that the upstream signal is the same regardless of the resistance of the downstream termination [88].

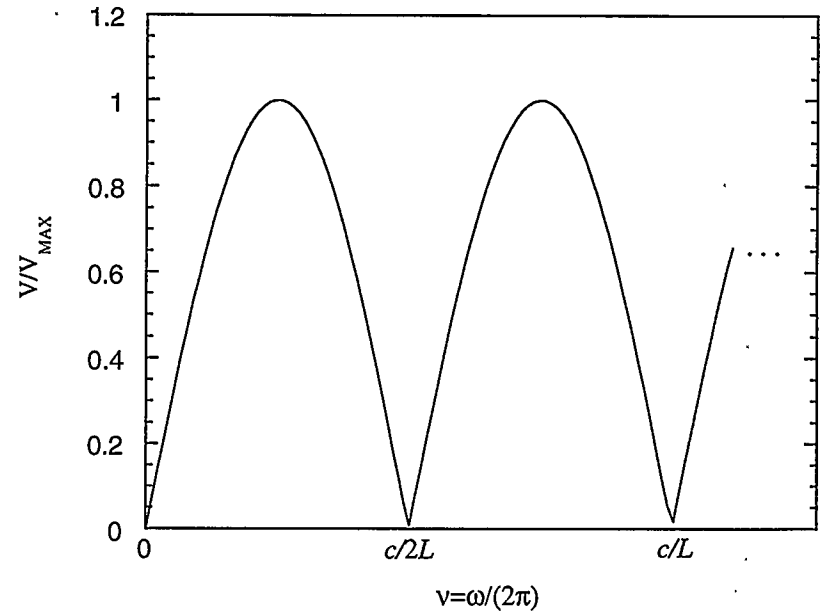The processing electronics which is used to detect the voltage surge shown in

Figure C.4: $|\tilde{V}(\omega)|$, normalized to the maximum voltage of $|V_0|\sqrt{2}/\sqrt{\pi}$.

Figure C.3 has a finite bandwidth, while the surge itself has components up to infinite frequency. A voltage $V(t) = V_0[\delta(t) - \delta(t - 2L/c)]$ can be Fourier-transformed:

$$\tilde{V}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} V(t) e^{i\omega t} dt \qquad (C.8)$$

$$= \frac{V_0}{\sqrt{2\pi}} [1 - \exp(2i\omega L/c)].$$

The absolute magnitude of the voltage surge in the frequency domain, $|\tilde{V}(\omega)|$, is given by:

$$|\tilde{V}(\omega)| = \frac{|V_0|\sqrt{2}}{\sqrt{\pi}} |\sin \frac{\omega L}{c}|. \qquad (C.9)$$

The value of $|\tilde{V}(\omega)|$ is plotted in Figure C.4. Note that the magnitude of the signal achieves a maximum at a frequency $\nu = \frac{c}{4L}$, and zeroes at $\nu = \frac{c}{2L}$.

The FFTB standard BPM has a stripline length of 457 mm [89], corresponding to a maximum voltage at $\nu = 164$ MHz. The bandwidth of the electronics used for

processing is on the order of 30 MHz [88]. In this low-frequency limit, $|\tilde{V}(\omega)| \approx \frac{|V_0|\sqrt{2}}{\sqrt{\pi}}|\frac{\omega L}{c}|$; thus the signal strength increases approximately linearly with the stripline length. The frequency response of the cable used to transport the signal from the BPM to the electronics will also tend to attenuate higher frequencies more than lower frequencies, leading to enhanced low-frequency components. The processing electronics must amplify the signal from the BPM striplines, detect the peak, and measure the peak pulse height. This pulse height is digitized and the resulting digital words from the four striplines are used to compute the beam position and charge.

C.2   Calibration of Stripline Beam Position Monitors

In the previous section, we saw that the signal at the digitizer output is a function of the BPM stripline unit, the processing electronics, and the cables between the striplines and the electronics. The system can be subject to offset-type errors which result in offsets in the measured positions relative to the actual beam position, and to gain-type errors which result in scaling factors and nonlinearities in the BPM response. Each segment of the system is therefore subject to vigorous calibration in an effort to eliminate these errors.

The cables used to carry the four stripline signals out of the tunnel are matched to within 100 picoseconds, or roughly 1 part in 1000, to optimally match the signal attenuations within a given BPM [90]. Electronics calibration is accomplished by introducing the signal from an external pulser to the two channels (N/S or T/B), and comparing the digital words which are generated [91]. While the cable matching is done at installation only, the pulsers are a component of the BPM processing hardware and can be used to calibrate the electronics at any time.

The BPM calibration was accomplished on the same test stand used to measure magnetic fields and magnetic centers, described in Appendix B. In this case, the
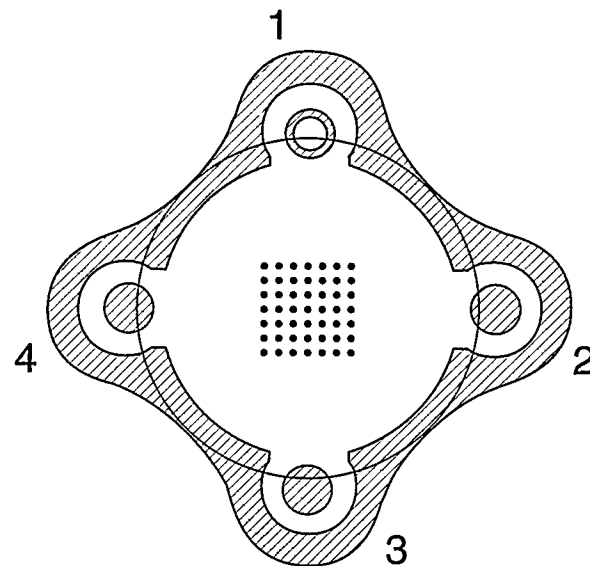
Figure C.5: Wire grid positions used for BPM calibration.

stretched-wire was connected to a pulse generator which emits 20 nanosecond pulses; these pulses excited the striplines in the same way that the electron beam would. A single channel of BPM electronics was then used to read out each of the four strip signals in turn, and these numbers were combined to reconstruct a BPM position measurement, $(x1, y1)$, as in Equation C.4.

As in the magnetic measurements, the stretched wire was positioned at different locations about the BPM aperture via the moving stages; the BPM striplines were read out at each location. In this way a map of $(x1, y1)$ as a function of (x, y) can be constructed. Finally, the map between estimated BPM positions and known wire positions was fitted, with an overall scale factor and horizontal and vertical offsets parameters to be determined. In this way, the contribution to scale factor and offset which is purely a function of BPM geometry and installation can be determined.
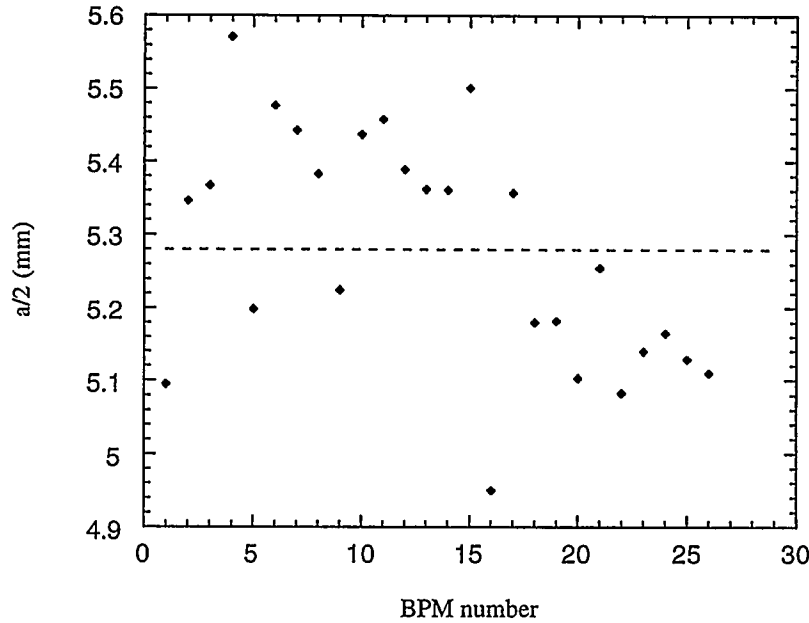
Figure C.6: The effective value of $a/2$ determined by bench tests with stretched wire and pulser. The dashed line represents the average value of 5.28 mm.

C.2.1 Details of the Procedure

The grid of wire positions used is shown schematically in Figure C.5. The furthest position from the BPM center is (± 3 mm, ± 3 mm), with a total distance from the origin of 4.2 mm. This distance is approximately 35% of the total radius of the BPM, which results in a 12% discrepancy between the lowest-order position estimate and the actual position. Consequently, the fitting algorithm includes the $(\frac{r}{a})^2$ correction term in Equation C.5.

While the electronics used with actual beam is capable of measuring with a resolution of 1 μm, the electronics used for this application is somewhat more coarse in its resolution. Furthermore, the pulse used on the wire is much longer and much lower in amplitude than the beam, resulting in extremely weak signals. As an *ad hoc*
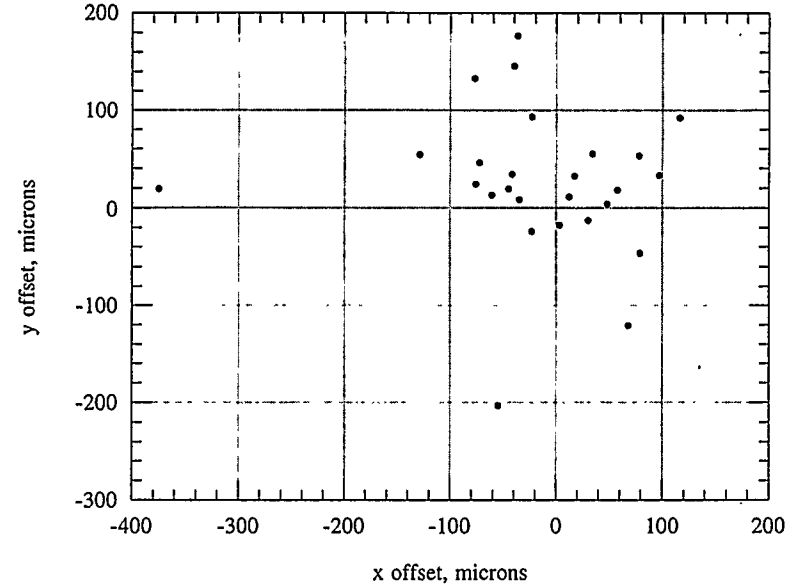
Figure C.7: Results of bench-test measurements of distance between BPM electrical centers and quadrupole magnetic centers.

assumption, the total wire-positioning resolution was assumed to be 20 μm in x and 20 μm in y. This error includes the possible scale factors and positioning errors in the wire, as well as the electronics readback resolution. A more detailed study of the repeatability of measurements was not performed.

C.2.2 Results of the Procedure

From purely geometric considerations, scale factor between (N-S)/(N+S) and the beam horizontal position was expected to be in the neighborhood of 5.75 mm, while POISSON simulation reported a scale factor closer to 5.6 [88]. As shown in Figure C.6, the actual measured scale factor averages 5.28 mm, with an RMS fluctuation over the ensemble of approximately 4%. The expected resolution of the fitted value

is approximately 20 μm, or 0.4%. The fitted scale factor disagrees with the simple geometric calculation by 8.2%, and with the POISSON calculation by 5.9%. Beam-based measurements described in Chapter 4 indicate that the 5.28 mm value may still be 12% too high. The cause for this discrepancy is not known.

The location of the electrical centers of the BPMs with respect to the magnetic centers (determined via vibrating the wire, see Appendix B) of their respective quads is shown in Figure C.7. With the exception of a one flier in x and one in y, the horizontal separation averages to -3 μm with an RMS of 64 μm, and the vertical separation averages to 34 μm with an RMS of 62 μm. The expected fit resolution from the bench-test procedure is approximately 40 μm in each plane.

Appendix D

The FFTB Collaboration

V. Balakin, V.A. Alexandrov, and A. Mikhailichenko
Budker Institute of Nuclear Physics, Novosibirsk and Protvino, Russia

K. Flöttmann, F. Peters, and G.-A. Voss
Deutsches Elektronen Synchrotron, Hamburg, Germany

V. Bharadwaj[1], M. Halling, and J. Holt
Fermi National Accelerator Laboratory, Batavia, Illinois, USA

J. Buon, J. Jeanjean, F. Le Diberder, V. Lepeltier, and P. Puzo [2]
Laboratoire de l'Accélérateur Linéaire, Orsay, France

G. Heimlinger, R. Settles, and U. Stierlin
Max-Planck-Institut für Physik, Werner-Heisenberg Institute, Föhringer Ring 6, D-80805, Munich, Germany

H. Hayano, N. Ishihara, H. Nakayama, K. Oide, T. Shintake, Y. Takeuchi, and N. Yamamoto
National Laboratory for High Energy Physics, KEK, Tsukuba, Japan

F. Bulos, D. Burke, R. Field, S. Hartman[3], R. Helm, J. Irwin, R. Iverson, S. Rokni, G. Roy[4], W. Spence, P. Tenenbaum, S.R. Wagner, D. Walz, and S. Williams
Stanford Linear Accelerator Center, Stanford, California, USA

[1]Currrent Address: Stanford Linear Accelerator Center, Stanford, California, USA.
[2]Current Address: European Laboratory for Particle Physics, CERN, Geneva, Switzerland.
[3]Current Address: Varian Associates, Nuclear Magnetic Resonance Division, Palo Alto, California, USA.
[4]Current Address: European Laboratory for Particle Physics, CERN, Geneva, Switzerland.

Bibliography

[1] See for example *Proceedings of the Workshop on Physics and Experiments with Linear Colliders*, (World Scientific, New Jersey, 1991) ; and the *Proceedings of the Second International Workshop on Physics and Experiments with Linear e^+e^- Colliders*, (World Scientific, New Jersey, 1993).

[2] D. Rice for the CESR Operations Group, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, Piscataway, 1993), p. 1979.

[3] B. Richter, Nuclear Instruments and Methods **136**, 47 (1976).

[4] R. Bailey *et al.*, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, PIscataway, 1993), p. 2001.

[5] K.L. Brown and R.V. Servranckx. "First- and Second-Order Charged Particle Optics," SLAC-PUB-3382. Presented at the 3rd Summer School on High Energy Particle Accelerators, Brookhaven National Laboratory, Upton New York (1983).

[6] G. J. Roy. *Analysis of the Optics of the Final Focus Test Beam Using Lie Algebra Based Techniques*, Ph.D. thesis, Stanford University, 1992. Also available as SLAC-Report-397.

[7] K.L. Brown, SLAC-Report-75 (1982).

[8] *SLC Design Handbook*, 1984.

[9] N.J. Walker *et al.*, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, Piscataway, 1993), p. 92.

[10] F. Zimmerman *et al.* "Performance of the 1994/95 SLC Final Focus System." To be published in *Proceedings of the 1995 IEEE Particle Accelerator Conference*.

[11] Roy, 24-27.

[12] K. Oide, in *Proceedings of the 1989 IEEE Particle Accelerator Conference*, edited by F. Bennett and J. Kopta (IEEE, Piscataway, 1989) p. 1319.

[13] D. Burke, private communication.

[14] H. Nakayama *et al.* Particle Accelerators 46 225 (1994).

[15] Roy, 92.

[16] M. Berndt and D. Macnair. "DC Magnet Power Supplies for the FFTB at SLAC." Unpublished.

[17] Roy, 85.

[18] S. Hartman, private communication.

[19] Roy, 96.

[20] N. Ishihara *et al.*, in *Proceedings of the 1992 International Conference on High Energy Accelerators*, edited by H. Henke, H. Homeyer, and C. Petit-Jean-Jenez (Editions Frontieres, Gif-Sur-Yvette, 1992) p. 945.

[21] Roy, 95.

[22] Roy, 99.

[23] Roy, 91.

[24] H. Hayano et al. Nuclear Instruments and Methods A320 47 (1992).

[25] F. Bulos, private communication.

[26] C. Field. Nuclear Instruments and Methods 360 467 (1995).

[27] C. Field, private communication.

[28] Field, 469.

[29] *SLC Design Handbook*, 7-23 – 7-24.

[30] G. Heimlinger. *Untersuchung von Prazisionsmechanik fur zukunftige 500 GeV e^+e^- Linear-Collider*, Ph.D. Thesis, München, 1993. Also available as MPI-PhE 93-13.

[31] G. Bowden et al. "Precision Magnet Movers for the Final Focus Test Beam." Nuclear Instruments and Methods A (to be published).

[32] T. Shintake. Nuclear Instruments and Methods A311 453 (1992).

[33] J. Buon et al. Nuclear Instruments and Methods A306 93 (1991).

[34] K. D. Jacobs et al., in *Proceedings of the 1989 Particle Accelerator Conference*, edited by F. Bennett and J. Kopta (IEEE, Piscataway, 1989) p. 1526.

[35] W. L. Spence, private communication.

[36] F. Bulos, in Proceedings of the Final Focus Test Beam Collaboration Meeting, 1992, edited by D. Burke and J. Rogers (unpublished).

[37] M.D. Hildreth, *Studies of the Strong and Electroweak Interactions at the Z^0 Pole*, Ph.D. Thesis, Stanford University, 1995, p. 301. Also Available as SLAC-Report-95-458.

[38] M.C. Ross et al., in *Proceedings of the 1987 IEEE Particle Accelerator Conference*, edited by E.R. Lindstron and L.S. Taylor (IEEE, Piscataway, 1987) p. 725.

[39] W. L Spence, private communication.

[40] P. Raimondi et al., in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, Piscataway, 1993) p. 98.

[41] CERN Application Software Group. *Minuit Reference Manual*, 1992 (not published) .

[42] N.J. Walker, J. Irwin, and M. Woodley, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, Piscataway, 1993) p. 95.

[43] C. Field, private communication.

[44] S. Smith, private communication.

[45] M. Halling, private communication.

[46] C.E. Adolphsen et al., in *Proceedings of the 1989 IEEE Particle Accelerator Conference*, edited by F. Bennett and J. Kopta (IEEE, Piscataway, 1989) p. 977.

[47] P. Emma, in *Proceedings of the 1992 European Particle Accelerator Conference*, edited by H. Henke, H. Homeyer, and C. Petit-Jean-Jenaz (Editions Frontieres, Gif-Sur-Yvette, 1992) p. 1625.

[48] D. Plouffe, private communication.

[49] C.W. Hawkes and P.S. Bambade. Nuclear Instruments and Methods A274 27 (1989).

[50] M.J. Lee *et al.*, in *Proceedings of the 1987 IEEE Particle Accelerator Conference*, edited by E.R. Lindstron and L.S. Taylor (IEEE, Piscataway, 1987) p. 1334.

[51] M. Halling, private communication.

[52] V. Bharadwaj *et al.* "Fermilab Contributions to the FFTB." To be published in *Proceedings of the 1995 IEEE Particle Accelerator Conference*.

[53] Paul Emma *et al.*, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by S.T. Corneliussen (IEEE, Piscataway, 1993) p. 116.

[54] The FFTB Sextupole alignment technique was originally suggested by K. Flöttmann, K. Oide, and P. Raimondi.

[55] D. Burke, private communication.

[56] K. Oide, private communication.

[57] P. S. Bambade. *Optical Tuning in the Arcs and Final Focus Sections of the Stanford Linear Collider*, Ph.D. thesis, Stanford University, 1989, p. 140. Also available as SLAC-Report-340.

[58] C. Field, private communication.

[59] D. Burke, private communication.

[60] T. Shintake, private communication.

[61] K. Oide, private communication.

[62] V. Lepeltier, in Proceedings of the Final Focus Test Beam Collaboration Meeting, 1995 (unpublished) 125.

[63] Buon, 102.

[64] Lepeltier, 141.

[65] V.A. Alexandrof *et al.* "Results of Final Focus Test Beam." To be published in *Proceedings of the 1995 IEEE Particle Accelerator Conference*.

[66] Lepeltier, 142-146.

[67] P. Emma, private communication.

[68] J. Irwin, private communication.

[69] Roy, 99.

[70] V.E. Bressler, R.E. Ruland, and D. Plouffe, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, Edited by S.T. Corneliussen (IEEE, Piscataway, 1993) p. 2736.

[71] S.C. Hartman. "Nanometer Resolution BPM Using Damped Slot Resonator." To be published in *Proceedings of the 1995 IEEE Particle Accelerator Conference*.

[72] F. Zimmermann *et al.* "A Final Focus System for the Next Linear Collider." To be published in *Proceedings of the 1995 IEEE Particle Accelerator Conference*.

[73] M. Ross. Proceedings of the 1994 AIP Workshop on Advanced Accelerator Concepts (1994) XXX.

[74] G.S. Abrams *et al.*, in *Proceedings of the 1987 IEEE Particle Accelerator Conference*, edited by E.R. Lindstrom and L.S. Taylor (IEEE, Piscataway, 1987) p. 1258.

[75] H. Wiedemann, *Particle Accelerator Physics* (Springer-Verlag, New York, 1993).

[76] D.C. Carey, *The Optics of Charged Particle Beams* (Harwood, Philadelphia, 1986)

[77] Carey, 34.

[78] J. Holt, private communication.

[79] J. Irwin, private communication.

[80] J. Irwin. Nuclear Instruments and Methods **A298** 460 (1990).

[81] Wiedemann, 109.

[82] Wiedemann, 259.

[83] J.K. Cobb and R. Cole, in *Proceedings of the International Symposium on Magnet Technology*, edited by H. Brechna and H.S. Gordon (Stanford, Stanford, 1965) p. 431.

[84] J.K. Cobb and D. Horelick, in *Proceedings of the International Conference on Magnet Technology*, edited by G. Söhngen, I. Schütz, and M. Tödt (DESY, Hamburg, 1970) p. 1439.

[85] J. Cobb, private communication.

[86] P. Tenenbaum *et al.*, in *Proceedings of the 1993 IEEE Particle Accelerator Conference*, edited by Steven T. Corneliussen (IEEE, Piscataway, 1993) p. 2838.

[87] R.T. Avery *et al.* IEEE Transactions on Nuclear Science **NS-18** 920 (1971).

[88] S. Smith, private communication.

[89] Hayano, 47.

[90] Hayano, 48.

[91] Hayano, 48.