SLAC-402
UC-414
(A)

A HIGH-BRIGHTNESS THERMIONIC MICROWAVE ELECTRON GUN*

Michael Borland

Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

February 1991

Prepared for the Department of Energy
under contract number DE-AC03-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161

*Ph.D. Thesis

SLAC—402

DE92 018587

Abstract

In a collaborative effort by SSRL, AET Associates, and Varian Associates, a high-brightness microwave electron gun using a thermionic cathode has been designed, built, tested, and installed for use with the SSRL 150 MeV linear accelerator. This thesis discusses the physics behind the design and operation of the gun and associated systems, presenting predictions and experimental tests of the gun's performance.

The microwave gun concept is of increasing interest due to its promise of providing higher-current, lower-emittance electron beams than possible from conventional, DC gun technology. In a DC guns, accelerating gradients are less than 8 MV/m, while those in a microwave gun can exceed 100 MV/m, providing much more rapid initial acceleration, thereby reducing the deleterious effects of space-charge. Microwave guns produce higher momentum beams than DC guns, thus lessening space-charge effects during subsequent beam transport. Typical DC guns produce kinetic energies of 80-400 KeV, compared to 2-3 MeV for the SSRL microwave gun.

"State-of-the-art" microwave gun designs employ laser-driven photocathodes, providing excellent performance but with greater complexity and monetary costs. A thermionic microwave gun with a magnetic bunching system is comparable in cost and complexity to a conventional system, but provides performance that is orders of magnitude better.

Simulations of the SSRL microwave gun predict a normalized RMS emittance at the gun exit of $< 10\ \pi \cdot \mathrm{m_e c} \cdot \mu\mathrm{m}$ for a beam consisting of approximately 50% of the particles emitted from the gun, and having a momentum spread of ± 10 %. These emittances are for up to $5 \times 10^9 e^-$ per bunch. Chromatic aberrations in the transport line between the gun and linear accelerator (GTL) increase this to typically $< 30\ \pi \cdot \mathrm{m_e} \cdot \mu\mathrm{m}$.

The SSRL microwave gun was designed to have a longitudinal phase-space suited to magnetic bunch compression. Simulations predict that peak currents of several hundred amperes are achievable.

Acknowledgements

As with any project of the complexity of the SSRL microwave gun and its associated systems, the task of going from the conceptual stage to the point where the system is turned over for operational use involves a great many people. It is my pleasure to acknowledge the contributions of those who participated in the project that I am reporting on in this thesis. With so many people deserving of thanks, there is the danger that I have overlooked someone or failed to fully recognize his contributions. I have done my best to avoid this, and apologize in advance if I fail in this regard.

First and foremost, I wish to thank my thesis advisor Helmut Wiedemann for introducing me to the exciting field of accelerator physics, for providing me with guidance and instruction as I learned about the field, and for giving me the rare opportunity to take on the microwave gun project. I am particularly grateful for Helmut's patience, support, and confidence in me when others thought that more conventional technology and a more experienced person were needed. Virtually all of the work that I report on here has benefited from Helmut's suggestions, ideas, scrutiny, and criticism.

Jim Weaver of SSRL participated in the design of the gun and many components of the gun-to-linac (GTL) system, in addition to designing the RF system for the gun and preinjector. I am convinced that without Jim's tireless attention to detail and his sound design ideas, the gun, the alpha-magnet chamber, and the chopper would not have been as successful as they were. I am also grateful to Jim for converting inches to millimeters when speaking to me and for laughing at most of my jokes.

Michael Baltay of SSRL did most of the mechanical engineering for the GTL and directed the assembly and alignment as well. All of the GTL magnets were designed in collaboration between Mike and myself (with many helpful ideas from Helmut, John

Voss, and others). Mike also helped me to assemble the necessary hardware for the magnetic measurements and helped with alignment of the measurement equipment.

Other SSRL and Injector people made significant contributions to the project: John Voss was the project manager for the Injector, and as such helped to organize the effort and push us toward completion. Clarence Chavis designed and directed the building of the modulators, and he and Jim Haydon frequently came to the rescue when something broke. Harry Morales oversaw the vacuum system work in the GTL and contributed to the design of many components. Paul Golceff's meticulousness when it came to the gun vacuum system and his prominent participation in the assembly of the GTL line, alpha-magnet chamber, and chopper are very much appreciated. Defa Wang assembled much of the equipment used for running the RF system, did cold test measurements on RF system components, and calibrated the gun RF diodes; he also proved tireless in the RF processing of the gun and linac, and in the commissioning of the chopper and linac. Jim Sebek did the TDR of the chopper, tested and refined the toroid design, stayed all night several times during the commissioning of the chopper, and, with Bob Hettel, helped to track down noise problems in many of the diagnostic signals. Ray Ortiz found and packaged the coaxial switches that proved so useful during the commissioning and during experiments, and also assembled the necessary circuits for running the gun filament. Bill Lavender's "test" control program for the GTL magnets ended up being the principle means of controlling the GTL during the commissioning and experiments. Brad Youngman, Tom Sanchez, and Ken Ruble participated in the design of the alpha-magnet chamber, chopper, and other GTL components.

Several of Helmut's other graduate students participated in the gun and GTL commissioning as well. Louis Emery in particular was indispensible during the crucial early gun experiments at SSRL, when many remained to be convinced that the gun could function reliably. In addition, it was very valuable to be able to refer to the LaTeX input for his thesis in preparing mine. James Safranek was responsible for the phosphor screens, TV cameras, and related hardware, and assumed his share of those seemingly endless RF processing shifts.

Substantial credit for the success of the microwave gun goes to Eiji Tanabe of AET Associates and Luther Nelson of Varian Associates. While I performed all of

the simulations, evaluated the predicted performance of the gun, and modified gun parameters in order to achieve what SSRL needed, Eiji provided the practical knowledge of what could be done and how it could be done, along with suggestions of what to try when a given idea didn't work out or wasn't easily realizable. Luther came up with the idea of using a toroidal tungsten spring and a metal-coated ceramic annulus around the cathode, eliminating the need for a RF choke and greatly simplifying the design; he also provided technical support in the assembly and testing of the gun.

Other Varian people contributed to the gun project as well. Mike Green supplied the cathodes for the prototype and consulted on many occasions on issues relevant to the cathode. Steve Skellenger did the initial RF processing of the gun and played a major part in many of the early experiments at Varian. Ralph Tol machined the gun cavity and many related components. Bill Leong tuned the gun and performed the bead-drop measurements. Ray McIntrye backed the Varian effort and generously provided time for initial testing in one of Varian's test cells.

Roger Miller of SLAC served as a reviewer and advisor on many aspects of the gun and GTL project, including the gun itself, the instrumentation, the magnets, and the chopper. I am grateful for his advice and for many stimulating conversations.

Radiation safety concerns for the project were addressed by Ed Benson, Nisy Ipe, and Ralph Nelson of SLAC.

On a more personal note, I wish to thank my Mother and Father for their support and understanding throughout my long education, and for being so proud of me.

Finally, thanks to Kim Adams, for everything we share.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The history of accelerator physics is inseparable from the history of the application of particle beams. The production of high-brightness electron beams is one of increasing interest among accelerator physicists precisely because of increasing interest in the applications of such beams. These applications provide the motivation for the research reported on in this thesis.

In this introductory chapter, I will indicate the nature of the these applications to give the reader some appreciation of the motivation for research into microwave electron guns. Prior to this, I review fundamental concepts—such as phase-space, Liouville's theorem, and emittance—that are necessary to the understanding of high brightness. I also review issues relevant to the production of high-brightness photon beams from synchrotron radiation emitted by high-energy electrons, including a discussion of coherent radiation and free electron lasers.

Microwave, or "RF", electron guns[1, 2] are a relatively recent development in the production of high-brightness beams. Their general characteristics and brief history are reviewed in the third section of this chapter.

I end this chapter with an overview of the main body of the thesis.

1.1 Review of Fundamental Concepts

In order to understand the meaning of "high-brightness", it is necessary to first understand several prior concepts. The first of these is the concept of emittance. The emittance of a beam is related to the phase-space volume occupied by the beam, or some fraction of it. To properly define the emittance requires a discussion of phase-space and Liouville's theorem.

1.1.1 Phase-Space and Liouville's Theorem

Phase-space refers to the six-dimensional space in which the spatial position and dynamical properties of any particle are defined. For example, the coordinates of phase-space may be taken to be the Cartesian coordinates (q_1, q_2, q_3) and the corresponding momenta (p_1, p_2, p_3). In classical mechanics, these six coordinates uniquely and completely describe the state of any particle. (More will be said about the choice of coordinates below.)

Suppose that at some time, t=0 say, an arbitrary closed surface S in phase space is chosen, and that it bounds a volume V. Allow S to evolve in time as if it were anchored to imaginary particles that lie on S at t=0. That is, consider S(t) to be defined as the surface occupied by those particles. Liouville's theorem[3] states that the phase-space volume, V(t), bounded by S(t) is constant, provided that only conservative forces act on the particles.

Proof of Liouville's Theorem

Perhaps the simplest and most intuitively appealing proof of Liouville's theorem is that given by Weiss[4]. For simplicity in notation, consider a two-dimensional phase-space, with coordinates (q, p), so that V(t) is the area bounded by a closed curve S(t).

In order to calculate the total time derivative of V(t), one need only look at the motion of the boundary, which again may be thought of as determined by the motion of hypothetical boundary particles that are defined by the choice of S(t=0). Let df represent the outward vector for an infinitesimal segment of the boundary. A motion

of this boundary element by $(\Delta q, \Delta p)$ will increase the area bounded by S by

$$df_q \Delta q - df_p \Delta p. \tag{1.1}$$

Dividing by Δt and integrating over S, one sees that

$$\frac{dV}{dt} = \int_S (\dot{q}, \dot{p}) \cdot df. \tag{1.2}$$

Use of the Divergence Theorem allows one to convert the surface integral into an integral over the region of phase-space bounded by S:

$$\frac{dV}{dt} = \int_V \nabla \cdot (\dot{q}, \dot{p})\, dq\, dp = \int_V (\partial_q \dot{q} - \partial_p \dot{p})\, dq\, dp. \tag{1.3}$$

For a conservative system [3], Hamilton's equations,

$$\dot{q} = \partial_p H \tag{1.4}$$

$$\dot{p} = -\partial_q H, \tag{1.5}$$

are applicable. Since the boundary S moves as if anchored to particles, these equations for particle motion specify the motion of the boundary, and may meaningfully be used in equation (1.3), yielding

$$\frac{dV}{dt} = \int_V (\partial_q \partial_p H - \partial_p \partial_q H)\, dq\, dp = 0. \tag{1.6}$$

Hence, the phase-space area bounded by $S(t)$ is constant as the system evolves.

Implications of Liouville's Theorem

Most statements of Liouville's theorem [5, 6] make use of the particle distribution function, $\Psi(\mathbf{q}, \mathbf{p}, t)$, which gives the density of particles in phase space. Since the paths of particles cannot cross in phase-space, any particle inside S at $t=0$ will remain inside S. Thus, if one accepts that the volume bounded by any surface $S(t)$ is constant, then it follows that the integral of Ψ over that volume is also constant, since this integral gives the number of particles in the volume and since particles cannot cross S. Since the volume bounded by $S(t)$ and the number of particles inside $S(t)$ are both constant, the average density of particles inside $S(t)$ must also be constant.

To deduce a further result, assume that $S(t)$ bounds an *infinitesimal* volume $dV = dq\,dp$, so that one can ignore the variation of Ψ over the volume. Then the statement that the average of Ψ over $V(t)$ is constant implies that the value of Ψ at the center of the volume (or for any other point inside it) is constant. That is, if one chooses any point $(\mathbf{q}, \mathbf{p})$ in phase space at $t=0$, then as one travels with a particle starting at that point and moving under the influence of conservative forces, the value of Ψ evaluated at the position of the particle is constant. Writing this in mathematical form, one obtains

$$\frac{d\Psi}{dt} = \frac{\partial \Psi}{\partial t} + \frac{d\mathbf{q}}{dt} \cdot \frac{\partial \Psi}{\partial \mathbf{q}} + \frac{d\mathbf{p}}{dt} \cdot \frac{\partial \Psi}{\partial \mathbf{p}} = 0, \tag{1.7}$$

which is the mathematical result commonly referred to as Liouville's theorem [5, 6].

Liouville's Theorem and Real Beams

In accelerator physics, one deals with the properties of ensembles of large numbers of particles, referred to as "beams" or "bunches". The real phase-space distribution of a bunch is the sum of many delta-functions:

$$\Psi_c(\mathbf{q}, \mathbf{p}, t) = \sum_{i=1}^{N} \delta(\mathbf{q} - \mathbf{q}^{(i)}(t)) \delta(\mathbf{p} - \mathbf{p}^{(i)}(t)), \tag{1.8}$$

where N is the number of particles, and $(\mathbf{q}^{(i)}, \mathbf{p}^{(i)})$ are the phase-space coordinates of the i^{th} particle.

What does Liouville's theorem tell us about the evolution of such a bunch? Contrary to the impression given by some discussions [5], there is nothing in the derivation [6] of equation (1.7) that invalidates it for a distribution of this type. In my discussion, I have taken care to refer to "imaginary" particles in defining boundary surfaces, in order to emphasize that Liouville's theorem is not dependent for its validity on having an infinite number of particles or a smooth continuous distribution of particles. Liouville's theorem stated in terms of the constancy of the volume inside a closed, evolving surface in phase-space is clearly a result that is valid regardless of what sort of actual beam distribution one has. This applies just as well to Liouville's theorem, as stated in equation (1.7).

One caveat needs to be issue in this regard [7]. The derivation of Liouville's theorem implicitly assumed that the Hamiltonian was macroscopic in nature, and in

particular that particle-particle forces could be included in the Hamiltonian in a way that did not refer to the individual coordinates of particles. If particle-particle collisions are considered, then the real Hamiltonian is a function of 6N variables, where N is the number of particles. In this case, Liouville's theorem is strictly valid only in 6N dimensional phase-space. In 6 dimensional phase-space, particle-particle collisions will increase the phase-volume occupied by the beam and *appear* to violate Liouville's theorem. For practical purposes, the phase-volume in 6 dimensional phase-space is what matters, so this can be an important issue. For sparsely populated beams, the issue becomes even more relevant, since then the particle fields cannot be smoothed into a macroscopic field.

Even given this conclusion, a distribution as defined by equation (1.8) is unwieldy and contains more information than is needed or useful. In the limit of a very large number of particles, Ψ_o can generally (though not always) be approximated by a smooth, continuous function of $\mathbf{q}$ and $\mathbf{p}$. The applicability of a smooth distribution depends on practical considerations of how well one wants the smooth distribution to match the actual distribution. Since real bunches are always confined to a limited volume in phase-space, a practical way to gauge whether a bunch is well-approximated by a smoothed distribution is to ask whether an arbitrary phase-space volume inside the bunch that is small compared to the total phase-space volume contains a number of particles that is much larger than 1.

1.1.2 Beam Emittance

I mentioned above that the emittance is related to the volume occupied by a bunch, or some part of it, in phase-space. The intervening discussion of Liouville's theorem indicates why the emittance is an important concept for accelerator physics. Ignoring dissipative effects such as synchrotron radiation[8], all of the forces in an accelerator are conservative. Hence, Liouville's theorem is applicable, and potentially provides a means of describing a bunch in terms of a conserved property that applies to the whole bunch, rather than in terms of the coordinates of the individual particles. Indeed, the standard analysis of beam evolution in terms of the Twiss parameters[8, 9] makes use of the emittance in order to simplify the computation of bunch properties along

an accelerator.

Commonly-Employed Simplifications

In this analysis, certain simplifications are commonly employed. First, instead of dealing with six-dimensional phase-space, it is common to deal with the 2-dimensional projections of the phase-space coordinates into the (q_1, p_1), (q_2, p_2), and (q_3, p_3) planes. In the absence of coupling between these planes, Liouville's theorem is valid for each of the 2-dimensional phase-spaces.

In accelerator physics, it is common to have a bunch that travels largely in one direction, i.e., as a well-collimated beam. Hence, instead of a Cartesian coordinate system, one employs a curvilinear coordinate system which follows the path of an ideal, central particle[10]. At each point along this path, one defines a locally Cartesian system in such a way that one axis (z) lies along the direction of motion of the central trajectory, while the other two axes (x and y), are perpendicular to the central trajectory, so that (x, y, z) is a right-handed system. Corresponding to each of these spatial coordinates is a momentum, thus completing the six-dimensional phase-space.

Discussion of the emittance per se is usually confined to the transverse planes, (x, p_x) and (y, p_y). For the longitudinal plane, (z, p_z), it is more common to speak of the bunch length and momentum spread without defining an emittance. Hence, I will confine myself to the transverse planes in what follows. I will also write the equations only for the (x, p_x) plane, though they are equally valid for the (y, p_y) plane.

Another simplification commonly used pertains to the method used for computing the area occupied by the bunch in each phase-plane. A seemingly straight-forward definition of the emittance would be: the smallest contiguous phase-space area containing a specified fraction, say 99%, of the particles. While this would give an accurate measure of the phase-space area occupied by the bunch, it is difficult to use in practice, and does not lend itself to analytical treatments. (In addition, it may not accurately characterize the effective phase-space area occupied by a bunch, due to filamentation and non-linear correlations in phase-space.)

The RMS Emittances

For this reason, and for others to be seen presently, the most commonly used means of measuring the phase-space area occupied by a bunch is the "normalized RMS emittance", defined as

$$\varepsilon_n = \pi m_e c \sqrt{\langle x^2 \rangle \langle p_x^2 \rangle - \langle p_x x \rangle^2}, \tag{1.9}$$

where angle brackets indicate averages of the bracketed quantities over the entire bunch. The factor of $m_e c$ indicates that the momentum is dimensionless, i.e., $p_x = \beta_x \gamma$. The reason for the factor of π will be seen presently. Both of these factors are absorbed in the units of ε_n, which in the present work are $\pi \cdot m_e c \cdot \mu$rad.

ε_n is referred to as the normalized emittance to distinguish it from the "geometric RMS emittance", defined as

$$\varepsilon = \pi \sqrt{\langle x^2 \rangle \langle x'^2 \rangle - \langle x' x \rangle^2}, \tag{1.10}$$

where $x' = p_x / p_z = \beta_x / \beta_z$ is the slope of a particle's trajectory. The units of ε are $\pi \cdot \mu$m or $\pi \cdot$ mm $\cdot$ mrad. If p_z is nearly the same for all particles, then

$$\varepsilon_n \approx \langle p_z \rangle m_e c \varepsilon. \tag{1.11}$$

In most applications, the geometric emittance is a more important quantity, since the divergence is what is relevant to the optical properties of the bunch or any radiation produced by it. (I will return to this issue later in this chapter.) Note, however, that if the bunch is accelerated, so that $\langle p_z \rangle$ is increased, the normalized emittance will be unchanged while the geometric emittance will decrease. Hence, the "phase-space" formed by (x, x') is really not a phase-space in the strict sense (it is sometimes referred to as "trace-space", instead [11]). p_x is the momentum conjugate to x, while x' is a ratio of two momenta. Because of these considerations, the normalized emittance is most relevant to the comparison of different accelerators that produce bunches of different longitudinal momenta.

Some authors [12, 13, 11, 14] prefer to define the RMS emittance with an additional factor of 4, in order to obtain a measure of the phase-space area occupied by a larger (though not necessarily well-defined) fraction of the beam. This definition is

also put forward because the emittance so defined is equal to the total phase-space area for a uniformly populated ellipse in (x, p_x) space (a projection of the "K-V" distribution [7, 14]). I have used the definition given in equation (1.9) principly because it is the definition used in the electron storage ring community.

For electron storage rings [8], a gaussian phase-space distribution is found to approximate the actual phase-space to a high degree. Such a phase-space distribution may be specified as

$$\Psi(x, p_x) = \frac{m_e c}{2\pi\varepsilon_n} \exp\left(-\frac{\pi^2 m_e^2 c^2 \langle x^2 \rangle \langle p_x^2 \rangle}{2\varepsilon_n^2} \left\{ \frac{x^2}{\langle x^2 \rangle} - 2\frac{\langle x p_x \rangle}{\langle x^2 \rangle \langle p_x^2 \rangle} x p_x + \frac{p_x^2}{\langle p_x^2 \rangle} \right\} \right) \tag{1.12}$$

where the normalization is such that

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp_x \Psi(x, p_x) = 1. \tag{1.13}$$

The interested reader may verify that if this distribution used to compute the right-hand side of equation (1.9), then the parameter ε_n in equation (1.12) is indeed the normalized RMS emittance.

It is now possible to explore the connection between the RMS emittance and the area occupied by the bunch in phase space. To simplify the analysis, let $\langle x p_x \rangle = 0$, so that $\varepsilon_n = \pi m_e c \sqrt{\langle x^2 \rangle \langle p_x^2 \rangle}$. (This simple equation for the normalized RMS emittance for an uncoupled phase-space distribution is one of the appeals of this definition of the emittance.) The distribution in equation (1.12) becomes

$$\Psi(x, p_x) = \frac{1}{2\pi \sqrt{\langle x^2 \rangle \langle p_x^2 \rangle}} \exp\left(-\frac{1}{2} \left\{ \frac{x^2}{\langle x^2 \rangle} + \frac{p_x^2}{\langle p_x^2 \rangle} \right\} \right). \tag{1.14}$$

Consider an ellipse in (x, p_x) space defined by

$$\frac{x^2}{\langle x^2 \rangle} + \frac{p_x^2}{\langle p_x^2 \rangle} = K^2, \tag{1.15}$$

where K is a dimensionless constant. The area of this ellipse is

$$A(K) = \pi m_e c K^2 \sqrt{\langle x^2 \rangle \langle p_x^2 \rangle}, \tag{1.16}$$

where I have included the "units" ($m_e c$) of p_x. For K=1, the area is equal to the normalized RMS emittance, which is the motivation for including the factor of π in

the definition of the emittance (equations (1.9) and (1.10)). The fraction of the bunch within this ellipse is readily computed:

$$F(K) = \int_0^K k\, dk \exp\left(-\frac{k^2}{2}\right) = 1 - \exp\left(-\frac{K^2}{2}\right). \qquad (1.17)$$

For K=1, the ellipse has an area equal to the normalized RMS emittance, and contains 39.35% of the particles. The maximum x coordinate of the ellipse is the RMS value of x, while the maximum p_x coordinate of the ellipse is the RMS value of p_x. For K=2, the ellipse has an area equal to four times the normalized RMS emittance, and contains 86.47% of the particles.

Virtues and Pitfalls of RMS Emittances

One virtue of the geometric and normalized RMS emittances is that both are constant for propagation through a system with no acceleration and where all transverse forces are linear in x (for the normalized RMS emittance, this is true only if the longitudinal momentum spread is zero[11]). For a bunch with no longitudinal momentum spread, the normalized RMS emittance is constant for a linear system, even with acceleration. Because of this, the degradation of the RMS emittances in a beamline is an indication of the severity of non-linear effects and cross-plane coupling in the beamline. This is discussed in more detail in Chapter 5.

In Chapter 5, I also discuss how the geometric RMS emittance can be measured experimentally. In the absence of noise and longitudinal momentum spread, measurements of RMS beam-sizes can be used to *exactly* measure the RMS geometric emittance, which is yet another of the appeals of the using the RMS definition of the emittance.

One problem with the RMS emittances is that they do not measure the actual phase-space area occupied by a fixed fraction of the bunch for an arbitrary phase-space distribution. In fact, it is easy to construct phase-space distributions occupying zero area while having non-zero RMS emittances. One solution to this problem is to use higher-moment descriptions of the beam phase-space, in order to correct the area estimate for higher-order correlations, thus producing an emittance that corresponds more closely to the actual phase-space area occupied by the beam[15]. One can also

use such analysis to evaluate and potentially eliminate higher-order phase-space correlations (which are caused by non-linear forces), thus decreasing the RMS emittance.

In the final analysis, one must realize that for non-gaussian beams the RMS emittances are simply convenient but potentially crude estimates of the actual phase-space area occupied by a bunch. In fact, there is simply no way to accurately describe a complicated phase-space distribution with a single number (like the RMS emittance) that will be appropriate to all cases and relevant in all applications. Whenever a precise description of the bunch phase-space is required, it is necessary to specify the distribution itself, either in some functional form or in terms of a representative sample of particles from the distribution. From the standpoint of gun and accelerator simulations, the later method is indeed the method used to specify and compute the evolution of the bunch. That is, in gun and accelerator simulations, one frequently simulates the bunch by a number of "macro-particles" which are representative of the actual distribution. This technique is used extensively in the present work.

If an accurate description of a complicated beam phase-space is to be provided by the accelerator physicist to those interested in applications of the beam, then that description may need to go beyond the RMS emittances and deal directly with the distribution. In many cases, particularly for guns and linear accelerators, RMS and other averaged beam properties must be recognized as approximate characterizations of the phase-space, suitable for approximate calculations only.

1.1.3 Beam Brightness

Bunch Length and Peak Current

The transverse emittances of a bunch, or its transverse phase-space distribution, do not fully characterize the bunch and its usefulness for applications. Missing is information about the number of particles in the bunch and their distribution in longitudinal phase-space. As noted above, the longitudinal phase-space of a bunch is most often characterized in terms of the momentum spread and the bunch-length. Bunch length can be specified either as the literal, spatial length of the bunch in the longitudinal dimension, or as the time-length of bunch, i.e., the time it takes for the particles in the bunch to pass by a fixed location along the beamline.

Some authors[14] define the bunch length as an RMS value about the centroid. For present purposes, I prefer to specify the bunch length as the time-interval occupied by the central 90% of the particles in the bunch. The reason is that the RMS bunch length and the 100% bunch length are sensitive to straggling particles that comprise a small fraction of the bunch and hence are unimportant. As a result, these are unreliable and potentially misleading measures of the bunch length.

Having defined the bunch length, one can go on to define the average current during the bunch, referred to as the "peak current"

$$I_{peak} = \frac{Q}{\delta t}. \tag{1.18}$$

where δt is the 90% bunch length and Q is the amount of charge in the interval δt. More specifically, $\delta t = t_2 - t_1$, where

$$\int_{t_m - t_1}^{t_m} \Psi(t)dt = \int_{t_m}^{t_m + t_2} \Psi(t)dt = 0.45. \tag{1.19}$$

t_m is the "median" time, that is, the time at which half the particles in the bunch, specified by the temporal distribution $\Psi(t)$, have passed by some specified point in the beamline. (Note that $\Psi(t)$ is normalized to unit area.) While this is not the only way one might define the 90% bunch length, it has the advantage of being readily computed and robust (in the sense of being insensitive to outlying particles or noisy distributions).

In contrast to the peak current, the "average current" is

$$I_{ave} = \frac{Q_b}{T_b} \tag{1.20}$$

where Q_b is the total charge per bunch and T_b is the time between bunches (equal to the RF frequency for a RF gun). In general, the average current is much less than the peak current, since the distance between bunches is much greater than the bunch length.

Brightness

The normalized brightness is the proportional to average current density in phase-space during the bunch, defined[16] in terms of the peak current and emittance as

$$B_n = \frac{2I_{peak}}{\varepsilon_{x,n}\varepsilon_{y,n}(\pi m_e c)^2}, \tag{1.21}$$

where $\varepsilon_{q,n}$ is the normalized RMS emittance for the q-plane. The units of the normalized brightness that I will use are $A/mm^2/mrad^2$.

The motivation for this definition[7] is that the volume of an uncoupled ellipse in four-dimensional phase space is $V_4 = \pi^2/2abcd$, where a, b, c, and d are the semi-major axes of the ellipse in each of the dimensions. Hence, for a uniformly filled ellipse (i.e., a K-V distribution), $V_4 = \pi^2/2x_{max}p_{x,max}y_{max}p_{y,max}$, so that the current density in phase-space is

$$\frac{I}{V_4} = \frac{2I}{\pi^2 x_{max}p_{x,max}y_{max}p_{y,max}}, \tag{1.22}$$

which corresponds to equation (1.21 except for the factor of π. (In retrospect, it would have been preferable to leave the factor of π^2 out of equation (1.21).)

To increase the brightness, one needs to increase the peak current, decrease the emittance, or both. To increase the peak current, a process known as "bunching" or "bunch compression" is often employed, which involves increasing the peak current by compressing the bunch into a shorter time-length. (This is discussed in Chapter 4.) Increasing the emittance is often a matter of mitigating emittance-diluting effects, such as non-linear fields. More will be said about about these issues in Section 1.3.

Having reviewed these concepts, many of which will be used throughout this work, I now proceed with a discussion of applications of high-brightness electron beams.

1.2 Applications of High-Brightness Beams

While the problem of producing high-brightness electron beams is a challenging one for the accelerator physicist, these beams are not pursued as ends in themselves. Among the applications that require such beams are linear colliders[17], wake-field and other two-beam acceleration schemes[18, 19], and radiation-producing devices, such as Free Electron Lasers (FELs)[20, 21]. Of these applications, I will discuss only those related to the production of radiation.

1.2.1 Synchrotron Radiation

Synchrotron radiation is the radiation emitted when a charged particle (usually an electron or positron) is subjected to accelerating forces imposed by external magnetic or electric fields. In this section, I will review the properties of synchrotron radiation, taking results from some of the many excellent discussions that appear in the literature[22, 23, 24, 25]

Bending-Magnet Radiation

The simplest way to produce synchrotron radiation is to send a high-energy electron beam through a uniform magnetic field transverse to the direction of travel of the beam. In such a circumstance, the beam undergoes acceleration perpendicular to its direction of motion, which bends its path into an arc of a circle of some fixed radius, ρ. This is commonly referred to as "bending-magnet radiation", since it is produced by the magnets used to "bend" the central trajectory of an electron beam (often into a closed path, as in an electron storage ring[8]). This radiation is characterized by the "critical frequency",

$$\omega_c = 2\pi f_c = \frac{3c\gamma^3}{2\rho}, \tag{1.23}$$

where γ is the electron energy in units of the rest mass. The spectral distribution of bending magnet radiation is such that half the radiated power is below the critical frequency and half above. In addition, the power spectrum of the radiation observed in the bend plane is peaked at $\omega \approx 0.83\omega_c$. The wavelength corresponding to the

critical frequency is

$$\lambda_c = \frac{4\pi\rho}{\gamma^3}. \tag{1.24}$$

It is well known that the radiation from a relativistic particle undergoing instantaneously circular motion is emitted primarily in the forward direction. If one defines the horizontal plane to be the plane in which the magnet bends the beam trajectory, then the radiation emitted from a bending magnet is spread over a horizontal angle essentially equal to the bend angle, assuming that the bend angle is large compared to $1/\gamma$. In the vertical plane, however, the "opening angle" of the radiation is much narrower. For $\omega = \omega_c$, the RMS vertical opening angle of the radiation power distribution is $\psi \approx 0.57/\gamma$.

The instantaneous total synchrotron radiation power for a single electron is

$$P_{se} = 4.611 \times 10^{-20}\text{Watt} \cdot \text{meter}^2 \times \frac{\beta^4\gamma^4}{\rho^2}, \tag{1.25}$$

Hence, if a beam consisting of bunches of N_e electrons with a repetition rate of f, passes through a bending magnet with bending radius ρ that bends through an angle θ, the average radiation power is

$$\langle P \rangle = P_{se}\frac{\theta\rho}{\beta c}N_e f \tag{1.26}$$

$$= 1.538 \times 10^{-25}\text{Joule} \cdot \text{meter} \times \frac{\beta^3\gamma^4}{\rho}\theta N_e f \tag{1.27}$$

As an example, the SSRL pre-injector linac[26] delivers a 120 MeV beam ($\gamma \approx 235$), which is subsequently deflected by 0.72 rad by a bending magnet with $\rho \approx 0.6$m, so that $\lambda_c \approx 0.6\mu$m, which is in the visible part of the spectrum. Typically 2×10^9 electrons are accelerated per pulse, with 10 pulses per second, so that the average radiation power is a mere 11 nW. However, as will be seen in Chapter 4, each electron bunch has a length of order 1 ps, so that the peak radiation power is of order 1 kW. The "natural" RMS opening angle of the radiation (i.e., ignoring beam emittance effects, which are discussed below) is 2.4 milli-radians.

For a 120 MeV beam with 2×10^9 electrons per bunch in approximately 1 ps bunches, the peak instantaneous beam power is of order 40 GW. This is seven orders of magnitude greater than the peak instantaneous radiation power produced in

the example. Clearly, a more efficient way of converting electron beam power into radiation power is desirable. From equation (1.27), it would appear that the most straight-forward ways to do this are to increase the field strength (i.e., decreasing ρ) and increase the path length, $l = \rho\theta$, in the bending magnet (these are not, of course, independent quantities).

Undulators and Wigglers

The power from a bending magnet is fanned out in the bend plane over an angle equal to the bending angle. Hence, the radiation power density per unit solid angle does not increase if one has a longer bending magnet with the same bending radius (which has the same fields but a greater bending angle). However, by using a series of bending magnets of equal but opposite bending angles, one *can* increase the radiation power per unit solid angle. Such a device is called a "wiggler" or "undulator" magnet [25]. By reducing the fanning-out of the radiation caused by the deflection of the beam path, these devices not only allow one to extract more power from a beam, but also to concentrate that power into a narrower solid angle, thus increasing the brightness of the radiation.

The dominant field component in such magnets is a transverse field described by

$$B_y = B_o \cos(2\pi z / \lambda_w), \tag{1.28}$$

where λ is the periodic length of the magnet. (As in the last section, z is the longitudinal coordinate, while x and y are the horizontal and vertical coordinates, respectively.) If B_o is not too large, the particle trajectory in the y=0 plane is sinusoidal, i.e.,

$$x = a \cdot \cos(2\pi z / \lambda_w), \tag{1.29}$$

where

$$a = \frac{eB_o\lambda_w^2}{4\pi^2 m_e c\gamma}. \tag{1.30}$$

The maximum slope of x(z) is thus

$$x'_{max} = a\frac{2\pi}{\lambda_w} \equiv \frac{K}{\gamma}, \tag{1.31}$$

where K is the usual undulator strength parameter. The distinction between wigglers and undulators is that for wigglers, $K \gg 1$, whereas for undulators K is less than or of order 1. That is, an undulator is a wiggler that causes only very slight motion of the beam, so slight that the natural opening angle of the radiation is not greatly degraded in either plane due to the fanning-out effect that occurs in ordinary bending magnets.

From equation (1.25), one sees that the instantaneous radiation power is proportional to $1/\rho^2$. Since $1/\rho \sim B_y$, the highest instantaneous power is radiated when the electron is at the crests of its sinusoidal trajectory, where the acceleration is greatest. For $K \gg 1$, x'_{max} is much greater than the natural opening angle of the radiation at the crest, which means that a distant, on-axis observer principly sees radiation emitted from near the crest of the oscillations (where the electron is traveling with $x' < \gamma$), which enhances the dominance of the radiation emitted at the crest. Hence, for $K \gg 1$, it is plausible that the radiation seen by a distant, on-axis observer will take the form of a series of pulses, one emitted from each crest in the electron's oscillations. The frequency spectrum of this radiation is dominated by the instantaneous spectrum at the crests, and one can show that the spectrum from a wiggler is indeed very much like that from a bending magnet with field B_o.

If $K \approx 1$, however, the same observer will receive radiation from a significant portion of the electron's oscillation, and one finds the the spectrum of this radiation is related to the frequency of the electron's oscillatory motion. In the average rest-frame of the electron, the electron executes transverse oscillations with frequency

$$f_e \approx \frac{\gamma\beta c}{\lambda_u}, \tag{1.32}$$

where the factor of γ is due to the Lorentz-contraction of the undulator period as seen in the moving frame. In its rest frame, then, the electron emits dipole radiation at this frequency. In the laboratory frame, this frequency is Doppler shifted, so that a spectrum of frequencies is produced:

$$f(\theta) = \frac{c}{\lambda_u}\frac{2\gamma^2}{1 + K^2/2 + \gamma^2\theta^2}, \tag{1.33}$$

where θ is the angle in the x-z plane relative to the axis of the undulator. For

$\theta < 1/\gamma$, i.e., near the axis, the wavelength of the radiation is simply

$$\lambda_1 = \lambda_u \frac{1 + K^2/2}{2\gamma^2}, \tag{1.34}$$

so that for a high-energy electron (i.e., $\gamma^2 \gg 1$), the wavelength of the radiation will be much shorter than the undulator period. For very weak undulators, the radiation spectrum is dominated by this "first-harmonic" radiation. For $K \geq 1$, the electron motion in the rest frame contains significant frequency components other than the fundamental oscillation, which generate additional frequencies in the rest frame at even and odd harmonics of the fundamental. The odd harmonics are a result of transverse oscillations, while the even harmonics are a result of longitudinal oscillations relative to the average rest frame. The odd harmonics of the motion produce spectral peaks at frequencies that are, of course, odd multiples of f_1. The odd harmonics are of the greatest interest, since the radiation in this case is confined to a narrow forward cone with an RMS divergence of

$$\sigma'_{r,k} = \frac{1}{\gamma}\sqrt{\frac{1 + K^2/2}{2nN}} = \sqrt{\frac{\lambda_1}{nL}} = \sqrt{\frac{\lambda_n}{L}} \tag{1.35}$$

where N is the number of undulator periods, L is the total undulator length, n is the (odd) harmonic number, and $\lambda_n = \lambda_1/n$ is the wavelength of the n^{th} harmonic. The subscript "r" is used to emphasize that this angular divergence is an intrinsic property of the radiation, separate from the electron beam divergence.

This result assumes that the radiation from the undulator may be approximated as coming from a source at the longitudinal center of the undulator and that the light is observed from a distance that is large compared to the length of the undulator[27]. If this assumption is made, then for a zero-emittance electron beam (e.g., a single electron), the radiation produced in an undulator has an angular divergence given by equation (1.35), as well as an apparent source size (due to the length of the undulator), given by

$$\sigma_{r,k} = \frac{1}{4\pi}\sqrt{\lambda_n L} \tag{1.36}$$

Hence, the geometric RMS emittance of the photon beam emitted by a single electron passing through an undulator is

$$\varepsilon_{r,k} = \pi \sigma_{r,k} \sigma'_{r,k} = \frac{\lambda_n}{4}. \tag{1.37}$$

For an electron beam with a gaussian transverse distribution, characterized by a spatial parameter σ_e and an angular parameter σ'_e, the effective photon beam parameters are obtained by adding in quadrature with the single-electron parameters[23, 27, 28] :

$$\begin{aligned} \sigma_s^2 &= \sigma_e^2 + \sigma_r^2 \\ \sigma_s'^2 &= \sigma_e'^2 + \sigma_r'^2 \end{aligned}$$

Thus, the emittance of weak undulator radiation for a non-zero emittance electron beam is

$$\varepsilon_s = \pi\sqrt{\sigma_e^2 + \sigma_r^2}\sqrt{\sigma_e'^2 + \sigma_r'^2}. \tag{1.38}$$

From this, one concludes that the emittance of the photon beam is maintained at its minimum value if

$$\sigma_e^2 \ll \sigma_r^2 \tag{1.39}$$

and

$$\sigma_e'^2 \ll \sigma_r'^2. \tag{1.40}$$

This implies that

$$\epsilon_e \ll \epsilon_r. \tag{1.41}$$

(Note that the latter condition is necessary but not sufficient to fulfill equations (1.39) and (1.40).)

This result two important implications. First, if one desires a low-emittance photon beam, then the best one can do is to supply an electron beam with emittance significantly less than the wavelength being produced, and with $\sigma_e/\sigma'_e = \sigma_r/\sigma'_r = 4\pi L$ (where L is, again, the length of the undulator). Second, as will be seen presently, if one maintains the intrinsic photon beam emittance, then (if other conditions are also satisfied) the radiation will be spatially coherent. This is path to Free Electron Lasers.

1.2.2 Coherent Radiation

In order to discuss the possibility that undulator radiation might be coherent, it is helpful to review the meaning of coherence[29, 30]. There are two varieties of

coherence, spatial coherence and temporal (or longitudinal) coherence.

Spatial Coherence of Transversely Extended Sources

Spatial coherence refers to the constancy of the phase across a wave-front of light. In particular, spatially coherent light will form interference fringes when used in a Young's two-slit experiment. Figure 1.1 shows an experiment in which an on-axis point source is used to illuminate two slits, spaced by 2A, a distance D from the source. Suppose further that the light from these slits falls on a screen at a distance S from the slits. If the width of the individual slits is small compared to D, L, and A, and if A is small compared to D and L, then interference fringes are formed on the screen, spaced by

$$\Delta x_f = \frac{\lambda S}{2A}. \tag{1.42}$$

If the point source is moved transversely (i.e., to an off-axis position), then the center of the fringe pattern moves as well, to

$$x_c = -\frac{x_s S}{D}, \tag{1.43}$$

where x_s is the distance of the source from the axis.

If one had two equal-strength point sources that individually produced fringe patterns offset by $\Delta x_f/2$ relative to one another, then the combined irradiance would be flat—i.e., no fringes would be seen. (One must, of course, add the electric fields and not the irradiances to see this.) This occurs when the two point sources are off axis at

$$x = \pm\frac{\lambda D}{8A}. \tag{1.44}$$

Hence, even though these two sources might be radiating in phase, the combined source does not produce interference fringes and is said to be spatially incoherent. As the point sources are moved toward one another, the fringes gradually reappear. To obtain a high degree of spatial coherence for two point sources offset by $\pm x_s$, one wants the centers of the two fringe patterns to be spaced by, say, less than $\Delta x_f/4$:

$$\frac{2x_s S}{D} < \frac{1}{4}\Delta x_f = \frac{\lambda S}{8A}. \tag{1.45}$$

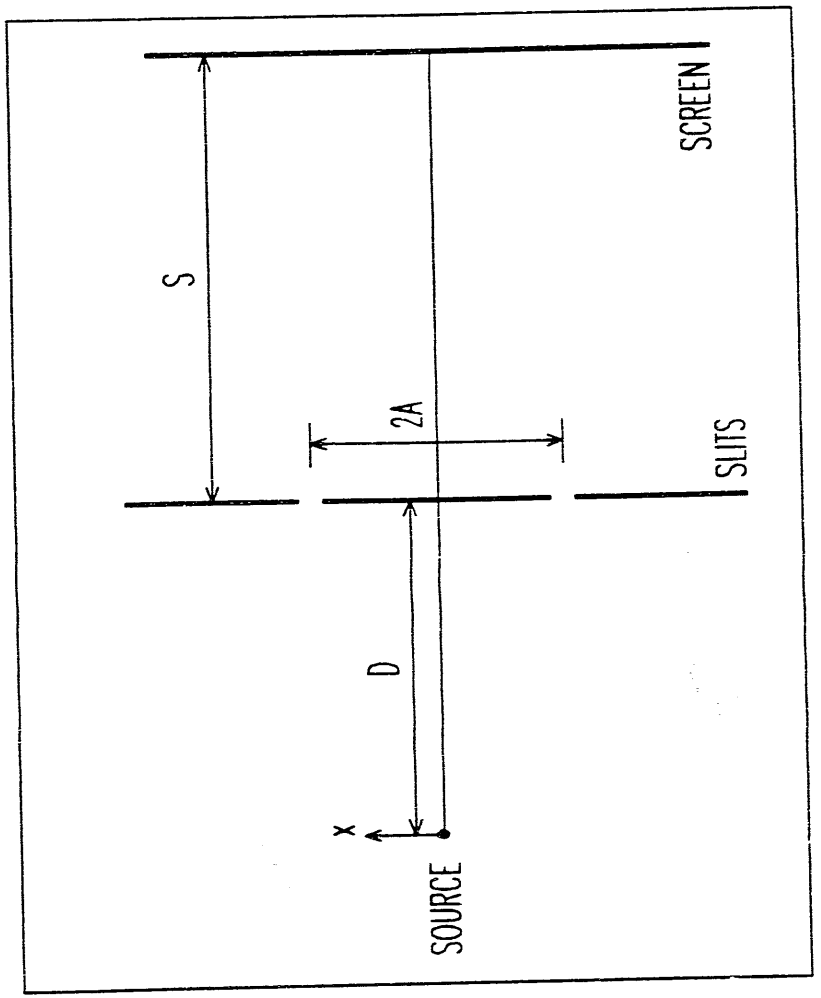

Figure 1.1: Young's Two-Slit Experiment

This can be rewritten as

$$x_s \frac{A}{D} = x_s \theta_s < \frac{\lambda}{16}, \tag{1.46}$$

where I have used θ_s to denote the angle of a ray that starts on-axis and passes through one of the slits.

If one now lets x_s represent the spatial size of an extended source and θ_s the characteristic angular divergence of the source, then one concludes that radiation with emittance $\varepsilon_r = \pi x_s \theta_s$ exhibits spatial coherence at certain wavelengths, namely if

$$\lambda > 16\varepsilon_r/\pi. \tag{1.47}$$

A more rigorous analysis shows[24] that for a gaussian beam, complete spatially coherence (i.e., no washing-out of fringes) is obtained if

$$\lambda \geq 4\varepsilon_r = 4\pi\sigma_r\sigma_r'. \tag{1.48}$$

While I have discussed coherence in the context of Young's experiment, and will continue to do so below, this is only for clarity. What the appearance of fringes in Young's experiment attests to is the coherence of the radiation from the source at a certain distance from the source (namely, the position of the slits) and within a certain region (namely, between the slits).

If one refers back to equation (1.37) for the emittance of single-electron undulator radiation, one sees that this condition for spatial coherence is in fact satisfied, so that single-electron radiation from an undulator is spatially coherent. In addition, one sees that if the electron-beam emittance is significantly less than $\lambda/4$, then the radiation is spatially coherent even for a beam of many electrons. One conclusion that can be drawn from this is that for a *zero-length* electron bunch that satisfies equation (1.48), the on-axis intensity is the coherent superposition of the radiation from each electron, so that the flux on-axis will be N^2 times the flux for a single electron.

Spatial Coherence of Longitudinally Extended Sources

Until now, I have considered the sources to lie in a plane of constant z (i.e., in a plane parallel to the planes of the slits and the screen). However, when an electron bunch

acts as a source of synchrotron radiation, it has a longitudinal extent as well as a transverse extent. Imagine then that one has two equal-strength, on-axis point sources, one at z=0 and another at $z = \Delta z$, with $\Delta z \ll D$. Since both sources are on-axis, each individually illuminates both slits with the same phase. Hence, spatial coherence is maintained. However, depending on Δz, the coherent flux passing through the slits may be diminished or increased relative to that for a single source, due to interference between the two sources.

The optical path length difference for light from the two sources to the slits is simply Δz, to first order. Hence, if $\Delta z \ll \lambda$ and the sources emit in phase, then constructive interference will occur at the slits. Clearly, for N equal-strength, in-phase sources, confined to a longitudinal interval $\Delta z \ll \lambda$, the peak on-axis intensity is increased by N^2. Thus, if an electron-bunch is short compared to the wavelength and if the transverse coherence condition (equation (1.48)) is also satisfied, then the on-axis flux will be N^2 times the single-electron flux, preserving the result obtained for a zero-length electron bunch. This provides a dramatic increase in the radiation power (in regions of constructive interference), and is therefore a highly desirable result.

Consider also that if the two sources are out of phase by 180° and spaced by $\lambda/2$, then they will constructively interfere. Hence, if two electron bunches, each short compared to λ, are spaced by $(n + 1/2)\lambda$ (where n is an integer) and experiencing opposite acceleration, then they will radiate in phase at the slits (or some distant on-axis point) and hence their intensities will interfere constructively. In this way, a train of mutually-coherent sources can be created.

A typical undulator period, λ_u, is 2.5 cm. For a 120 MeV beam such as that produced by the SSRL pre-injector, the first-harmonic wavelength (assuming $K \ll 1$) is $\lambda_1 = 0.2\mu$m. A relativistic electron bunch with a bunch-length of 0.2μm would have $\delta t = 0.6 \times 10^{-3}$ps. While bunch-lengths of order 1 ps are possible (see Chapter 4), a sub-femto-second bunch length is not within the realm of current accelerator technology. Hence, it would seem that the promise of coherent undulator radiation is out of reach. The solution to this problem is called a Free Electron Laser (FEL), and I will discuss it in the next section.

Temporal Coherence

First, I wish to complete the discussion of coherence by explaining temporal coherence. Temporal coherence refers essentially to the monochromaticity of light, and can be understood by returning to Young's two-slit experiment with a single on-axis source. If one modifies the experiment by placing refractive material over one of the slits, then the optical path length for light going through that slit is increased, resulting in a phase-shift at the screen relative to the light going through the other slit. Suppose that the light going through the refractive material is delayed by time Δt. For a perfectly monochromatic source, this phase shift is irrelevant, and merely results in a shifting of the fringes. A perfectly monochromatic source is said to be completely temporally coherent, or to have $\Delta t_{coh} = \infty$.

However, for a realistic source, the spread in frequencies means that for too long of a delay, $\Delta t \geq \Delta t_{coh}$, the interference fringes will no longer appear. Instead, one will see the uniform illumination characteristic of incoherent light. To see this, imagine that a source has a spread $\Delta\omega$ of frequencies about ω_o. If these frequencies are in phase at the source at time t=0, then at $t = \pi/\Delta\omega$, the phase spread for light emerging from either of the slits will be 2π. If light from such a source is used in a modified Young's two-slit experiment with a refractive plate giving $\Delta t \geq \pi/\Delta\omega$, then the interference pattern will be washed out because the phase-shift will shift the fringes for the outer frequencies by one-half of the fringe spacing. Hence, $\Delta t_{coh} = \pi/\Delta\omega$ is the time over which the source is said to maintain coherence with itself. The coherence time can be improved to the desired degree by employing a mono-chromator with a sufficiently narrow band-pass.

The spectral broadness of undulator radiation at harmonic n can be estimated by taking the Fourier transform of a sinusoidal field oscillation of Nn periods, where N is the number of undulator periods. One finds that the FWHM of the spectral power distribution around each harmonic is

$$\Delta\omega_n \approx \frac{\omega_n}{nN}, \tag{1.49}$$

implying coherence times of

$$\Delta t_{coh} = \frac{N\pi}{\omega_1} = \frac{N\lambda_1}{2c} \tag{1.50}$$

and coherence lengths of

$$\Delta l_{coh} = c\Delta t_{coh} = \frac{N\lambda_1}{2}. \tag{1.51}$$

Since typically $N \gg 1$, one sees that undulator radiation is temporally coherent over many wavelengths.

1.2.3 Free Electron Lasers

Energy Transfer Condition

Imagine that an electron beam is sent through an undulator, and that the fundamental radiation wavelength is λ_1. Further imagine that a laser beam of the same wavelength is also made to pass through the undulator, so that it overlaps the electron beam. Electrons oscillate at the same frequency as the laser field (otherwise, the electrons wouldn't radiate at the same frequency).

The transverse motion of the electrons is determined primarily by the undulator fields [21], so that for any electron

$$\begin{aligned} \frac{dv}{dt} &= -\frac{e}{m_e c^2} v E \cos(2\pi z(t)/\lambda_u)\cos(\omega t - kz(t)) \\ z(t) &= \beta^* c(t - t_o), \end{aligned}$$

where $\beta^* c$ is the average longitudinal velocity of an electron which enters the undulator and laser fields at $t = t_o$. v and E are positive quantities, v being the peak transverse electron velocity and E the peak transverse field strength of the laser beam. (This ignores the transverse motion that individual electrons have in addition to the motion imposed by the undulator, consistent with the assumption of a small beam-emittance.)

When the phase of the undulations and of the laser field are such that the cosines have the same sign, then energy is transferred from the electron to the laser beam. The energy transfer is greater when the transverse electron velocity is greater, i.e., when the electron is near the zero crossing of its undulating motion. If the cosines are of opposite sign, then the field does work on the electron and thus energy is extracted from the laser beam. Clearly, to amplify the laser beam, one wants the former condition to hold. However, since the electrons do not travel exactly at the speed of light, it would seem that the condition cannot be maintained, and that the electron will alternately gain and loose energy as it falls behind the wave.

To see how net energy transfer to the wave is possible, note that there will be such transfer if

$$A = \int_0^L \cos(2\pi z/\lambda_u)\cos(kz - \omega t(z))dz > 0. \quad (1.52)$$

where I am now considering t a function of z (the electron position), and where L is the length of the undulator. Rewriting this using equation (1.52) one obtains

$$A = \int_0^L \cos(2\pi z/\lambda_u)\cos(2\pi z\ \lambda_1(1/\beta^* - 1))dz \quad (1.53)$$

To maximize A, one equates the factors multiplying z in the cosines, obtaining

$$\frac{\lambda_1}{\lambda_u} = \frac{1}{\beta^*} - 1, \quad (1.54)$$

which is the "phase-slip" condition[20]. If the phase-slip condition is satisfied, then as the electron falls back, its transverse velocity is falling as well, so that when it has fallen back to where the laser field has changed sign, its velocity has also changed sign. The crests of the electron motion, where the velocity is zero, coincide with zero-crossings of the laser fields. The phase-slip condition, together with the electron beam energy and equation (1.34) for the radiation wavelength, must be self-consistently solved in order to find the conditions on K and γ necessary to achieve energy transfer to the laser beam for a given λ_u. Alternatively, one can find the "synchronous velocity", $v_s = \beta^* c$ that electrons must have in order to give (or receive) net energy from the laser fields.

Micro-Bunching

Most of the elements necessary for an intuitive understanding of FELs have been reviewed in the previous sections. I showed that under certain conditions, undulator radiation could be spatially coherent and that the on-axis flux would scale like N^2. It seemed, however, that to realize this would require an unrealistically short electron bunch. FELs succeed in spite of this because the radiation field interacts with the bunch to produce "micro-bunching", i.e., longitudinal density modulations on the scale of the light wavelength[21, 20].

For an initially longitudinally uniform beam traveling at the synchronous velocity, half the particles will loose energy and half will gain energy. Unless the density of the

beam can be modulated on the scale of λ_1, there will be no amplification of the laser beam. However, since the electron beam energy is being modulated at the wavelength of the laser beam, so is the velocity. This velocity modulation has the same spatial frequency as the laser fields, and is thus just what is needed allow coherent power to be generated. (see Section 1.2.2, page 21).

Note that the derivation of the phase-slip condition ignored the fact that as electrons gain or loose energy, their velocities change. In order to get gain, the beam must initially travel somewhat faster than the synchronous velocity, so that bunching is not symmetric about the null of the laser field[20].)

Electron Beam Requirements for FELs

Having given a brief account of the physics at work behind the generation of coherent radiation in FELs, I now list the general beam-quality requirements for an FEL:

- The electron beam emittance should be less than the natural emittance of the undulator radiation. More precisely, equations (1.39) and (1.40) should be satisfied.
- The undulator parameter K, the undulator wavelength, and the initial beam momentum should be chosen to give an initial beam velocity somewhat greater than the synchronous velocity, so that micro-bunching occurs in the region where the electron beam looses energy to the laser beam. (See [20] for a more precise statement.)
- The bunch length, δt, should be as long as practical in order to provide more micro-bunches, which results in greater gain. Similarly, the peak current should be high, in order to give as much charge per micro-bunch as possible.
- The initial momentum spread of the bunch should be small enough to be within the "buckets" created by the laser field. If the initial momentum spread is too large, the gain is decreased because not all particles have the proper velocity relative to the synchronous velocity. More specifically[20], one wants

$$\frac{\Delta\gamma}{\gamma} \leq \frac{1}{2N}, \quad (1.55)$$

where N is the number of undulator periods.

1.3 RF Guns

A RF or microwave electron gun [1, 2] is an electron source consisting of an electron emitter (the "cathode") immersed in the radio-frequency fields inside a metal-walled cavity. The SSRL RF gun is discussed in detail in Chapter 2. For the present, I will make some remarks about RF guns in general and about their history.

1.3.1 Varieties of RF Guns

There are two principle varieties of RF guns in use today[14]: "thermionic" and "photocathode" RF guns. They are distinguished by the type of cathode used or, more precisely, by the way the cathode is made to emit electrons.

Thermionic RF guns employ cathodes that must be heated in order to obtain emission of electrons. In the simplest design, a pillbox cavity[31] might be used with a thermionic cathode being part of one end-wall, and with a beam exit-hole in the opposite end-wall. When the RF fields in the cavity are in the accelerating phase, electrons are accelerated off of the cathode. With proper design, a large fraction of these electrons exits the cavity before the fields go into the decelerating phase. Those that do not exit the cavity are decelerated and turned around, and may return to impact the cathode (a phenomenon referred to as "back-bombardment"). As long as RF power is supplied to the gun, this cycle is repeated every RF period, resulting in a train of bunches spaced at the RF period.

Commonly-used cathode materials for thermionic guns are LaB_6 and dispenser cathodes (a tungsten matrix with work-function-lowering compounds added). Typical operating temperatures for LaB_6 are 1600°C[32], while the dispenser cathode for the SSRL RF gun is run at 950°C. Typical current densities for both these types of cathodes are in the 10-30 A/cm^2 range, though LaB_6 is capable of up to 200 A/cm^2[2], and advanced dispenser cathodes of up to 140 A/cm^2[33].

Problems with thermionic guns stem from the fact that emission occurs throughout the accelerating phase of the RF, and during every RF period. This results in a beam with a large momentum spread and a large time-spread as well. These issues are discussed in more detail below and in Chapter 2

Photocathode (or "laser-driven") RF guns[34] employ a photoemitting cathode

material that is pulsed by a laser. Commonly-used cathode materials[14] include LaB_6, Cs_3Sb, and CsK_2Sb. Typically, the laser pulse is much shorter than the RF period, so that emission occurs only over a short phase interval. Thus, a photocathode RF gun can produce a much shorter bunch than a thermionic RF gun working at the same RF frequency, since in thermionic guns electron emission occurs continuously during the accelerating phase of the RF.

While it is not my purpose here to discuss the physics of photocathode RF guns in detail (see [13, 35, 36, 37]), I wish to indicate the reasons that such guns out-perform thermionic systems.

The high current densities possible from a photoemitter (as much as 600 A/cm^2[38]) result in significantly higher charge per bunch for photocathode RF guns than for thermionic RF guns. Because of the shorter phase-interval during which electrons are emitted in a photocathode RF gun, RF focusing effects (see Chapter 2) are greatly reduced relative to a thermionic RF gun, resulting in smaller emittances. Since the current density is very high, the pulse from a photocathode RF gun need not be greatly compressed in order have high peak currents; hence, one sees that photocathode based systems generate longer bunches than thermionic based systems, since the later systems *must* compress to very short bunches in order to achieve high peak currents. As noted in Section 1.2.3, a long, high peak-current bunch is preferred for FEL work.

Photocathode RF guns also have the advantage of being free of the cathode back-bombardment problem that can plague thermionic RF guns when long RF pulses or high repetition rates are used[39]. In addition, the use of the laser to trigger emission permits a more flexible bunch pattern, at least in principle. For example, if only N bunches are desired per RF pulse, then one need only fire the laser N times per pulse. The firing can be delayed until the cavity has fully charged, so that the fields are not changing between bunches. In a thermionic RF gun, emission occurs continuously, even while the cavity is filling, giving a train of bunches that vary in momentum until the beam-cavity system has come to equilibrium. This has implications for FEL use, where a small momentum spread is required.

The superior performance of a photocathode RF gun comes at the expense of the greater complexity of a photocathode-based system, which requires a complicated and expensive high-power, RF-synchronized laser and, for the highest-performing

systems using Cesiated cathodes, a mechanism to withdraw and re-Cesiate the cathode at intervals [40]. While a thermionic RF gun with a magnetic bunching system is comparable in complexity and cost to a DC-gun-based system with a gap-and-drift buncher[41], a photocathode RF gun is substantially more complicated and much more expensive. Given that the thermionic RF gun system can out-perform a conventional DC gun system (excluding sophisticated state-of-the-art DC gun systems, employing several sub-harmonic bunchers, such as the SLC injector[41]) by orders of magnitude in brightness and peak-current, there is clearly a role for thermionic RF guns. Data to substantiate this claim is given in Chapter 4.

1.3.2 A Brief History of RF Guns

Kapitza and Melekhin[42] report that in 1948 the first microtron was constructed, and that it used field-induced emission from the gap of the accelerating cavity to generate a beam. In the broadest sense of the term, then, this was the first RF gun. In 1959, Melekhin proposed the use of a hot cathode placed off-axis in the microtron cavity. The primary difference between microtron guns and modern RF guns is that in microtrons, the cavity must not only accelerate the beam off of the cathode, but it must also provide acceleration each time the beam returns to the cavity. Hence, the microtron cavity must have entrance and exit beam holes, which greatly complicates the placement of the cathode and the initial trajectories of emitted electrons. Without reviewing microtron performance in detail, I will simply state that it does not equal that of a modern RF gun.

R.B.Neal would appear to be the first to use a RF gun other than a microtron gun. In a 1953 report[43] on work done at Stanford's Microwave Laboratory, Neal discusses experiments done with a hot cathode inserted in the first cell of a 2856-MHz linear accelerator section. These experiments were done in order to explore certain aspects of electron capture by RF fields, and not for the purpose of developing a new type of gun.

In 1975, Y. Minowa of Japan's Mitsubishi Electric filed for a Japanese patent on a RF gun[1], though it is unclear when an operational gun was actually built. Successful operation and experimental characterization of a multi-cell Mitsubishi RF

gun with a LaB_6 cathode was reported at the 1989 Linear Accelerator conference in Nara, Japan [44].

The RF gun is usually considered to have been invented by G.Westenskow and J.M.J.Madey at Stanford's High-Energy Physics Laboratory (HEPL). Their design work [2] on a gun employing a pillbox-like cavity and a LaB_6 cathode was first reported in the literature in 1984, and the gun was installed and operated in 1985 [39]. The HEPL system was the first to combine a RF gun with an alpha-magnet-based magnetic buncher [32, 45], resulting in a simple, compact system.

In 1985, J.S. Fraser and R.L. Sheffield of Los Alamos proposed use of a laser-pulsed cesium antimonide cathode in a multicell RF gun [34]. Experimental characterization of a prototype one-cell gun with such a cathode was reported in 1987 [40].

The number of proposed RF gun projects increased rapidly after 1985. A 1990 review article by C. Travier [14] gives the number of RF gun projects world-wide as 22. Those readers interested in a listing of the parameters of existing and proposed RF gun projects are referred to this article and to Chapter 4.

1.3.3 Factors That Degrade Electron Beam Brightness

The degradation of electron-beam emittance, and hence brightness, in electron acceleration and bunching is discussed in many places in the literature [46, 47, 16]. Knowledge of these effects is necessary if one is to appreciate the advantages and disadvantages of RF guns. For the present, I will list and briefly discuss some of the effects involved, leaving the detailed discussion for the chapters that follow. Corresponding to each of these effects is the brightness increase to be gained by eliminating its influence.

Effects that Degrade Emittance

- *Thermal velocity spread at the cathode* is almost always a negligible effect, being overwhelmed by other effects. This is the case for the SSRL RF gun.
- *Non-linear space-charge forces* are a particular problem for high-brightness beams, which by definition have high charge density. These effects are mitigated by accelerating as rapidly as possible to relativistic energy, where the beam is

"stiffer". For a beam of constant radial cross-section, the radial acceleration scales like $1/\gamma^3$ [7]. This is part of the motivation for RF guns, as I will discuss presently.

- *Time-dependent focusing forces due to RF fields* are a particular problem for thermionic RF guns, as mentioned above. These also have a non-negligible effect in photocathode RF guns [15] and in DC-based injectors, since time-dependent RF fields and long phase-length bunches are present during initial bunching. This is the dominant emittance-increasing effect in the SSRL RF gun.
- *Non-linear tranverse forces in RF cavities and from DC magnets* are avoidable, in general, though not always in practise, by increasing the apertures of the magnets or RF cavities, or by properly shaping the metal surfaces in the cavity. In the SSRL RF gun, non-linear tranverse fields in the gun are responsible for significant emittance increase and for broadening of the momentum spectrum. Solution of this problem by a simple cathode modification is discussed briefly in Chapter 4.
- *Chromatic aberrations from DC magnets* are a problem for thermionic RF guns, which emit a very broad momentum spectrum, since particles starting from the cathode at different phases are given different energies. In order to increase intensity, it is often desirable to use particles emitted over a substantial phase-interval, which entails dealing with a larger momentum spread. For the SSRL RF gun, using electrons emitted over 40° of phase entails accepting a momentum spread of about $\pm 10\%$.

1.3.4 Advantages and Disadvantages of RF Guns

Advantages

Until now, I have not discussed why RF guns are being adopted in preference to well-established DC gun technology. However, from the foregoing discussion it will have been anticipated that RF guns are capable of delivering very high-brightness beams.

The principle advantage of RF guns stems from the rapid acceleration that can be achieved with the strong electric fields possible in an RF cavity. The maximum

fields possible in a DC gun are about 8 MV m^{-1}, while the peak surface fields in a 2856 MHz RF can be as high as 240 MV m^{-1}. Typical output energies for DC guns are from 0.05 MeV to 0.4 MeV (kinetic energy), with the upper ranges being reached only by the most state-of-the-art designs. In contrast, the SSRL RF gun delivers a 2-3 MeV beam. This higher beam momentum greatly decreases the effect of space-charge forces in subsequent bunching and transport.

Another advantage of thermionic and photocathode RF guns is that bunching (in order to decrease the bunch length and hence increase peak current) can make use of relatively a simple magnetic bunching system (see Chapter 4), rather than the more complicated gap-and-drift prebuncher[41] and sub-velocity-of-light RF "buncher" acceleration section required in conventional injectors.

Disadvantages

The principle disadvantage (or difficulty) for a thermionic RF gun is the time dependence of focusing and energy gain in the gun, due to the time-varying nature of the RF fields. DC guns have no RF focusing in the gun *per se*, but it is inevitably encountered in subsequent bunching and acceleration. Similarly, while a DC gun emits a highly mono-energetic beam, that energy must be modulated (in order to modulate the velocity) for bunching. Hence, the time-dependence of the gun fields is not necessarily a *net* disadvantage relative to DC-based systems, but it does decrease the achievable brightness, and is one of the principle motivations for use of a photocathode.

A more significant problem for thermionic RF guns is that of cathode lifetime and survivability in the presence of back-bombardment. For sufficiently high repetition rates and long RF pulse lengths, back-bombardment can damage the cathode and degrade the gun's performance. In addition, back-bombardment can lead to current variation during the RF pulse, which causes problems for FEL applications[39]. The SSRL system is run with a 2 μs RF pulse and a repetition rate of 10 Hz, so that back-bomdardment is not a serious issue.

Another issue in thermionic RF guns is that of cathode thermal isolation. Since thermionic cathodes have operating temperatures typically in excess of 1000° C, the cathode support stem cannot be in direct contact with the metal surfaces of the cavity.

At the same time, it is necessary that the cathode be in electrical contact with the cavity. This is usually done using an RF choke[32], which can be problematical, leading to distortions of the fields in the cavity and unanticipated power losses. For the SSRL RF gun, a simpler concept was employed, as discussed in Chapter 2.

Conclusion

One cannot conclude from this discussion whether RF gun based preinjectors will in fact provide brighter beams than conventional preinjectors. Such a conclusion requires detailed simulations of the effects discussed in this section and the last.

In Chapter 4, I present comparative data for the SSRL preinjector, other RF-gun-based preinjectors, and DC-gun-based preinjectors. It will be seen that, in terms of beam brightness, a thermionic RF gun system with magnetic bunching can outperform all but the most state-of-the-art DC-gun-based injectors, which employ high-performance DC guns with multiple frequency sub-harmonic bunching. Thermionic RF gun systems can outperform conventional "low-technology" injectors by orders of magnitude in brightness and peak current. That this can be done using low-cost magnetic bunching technology, with a system using a single RF frequency (the gun frequency is the same as the linear accelerator frequency), and with only velocity-of-light accelerator section, is a significant improvement in the simplicity and affordability of high-brightness injectors.

1.4 Overview of Thesis

Figure 1.2 shows a schematic layout of the SSRL RF gun project, including the gun, the gun-to-linac transport line (GTL), and the beginning of the first of three SLAC-type 4$^{\text{m}}$ 2856-MHz accelerating sections. Electrons start at the gun cathode and are accelerated by the fields in the gun cavity, forming a beam that is transported through the first part of the GTL to the bunch-compression alpha magnet. After emerging from the alpha-magnet, the train of bunches is "chopped" by the traveling-wave beam-chopper, which admits 3-5 bunches into the linear accelerator. (A more detailed discussion of the GTL will be found in Chapter 5.) The plan of this thesis in large part follows the path that electrons take—that is, it proceeds from the gun, through the alpha-magnet, to the linear accelerator.

In this section, I present a brief overview of the thesis. To do so, I must anticipate much that will not become completely clear until latter in the thesis. It is also appropriate at this time to be explicit about my individual contributions to the project, having noted the contributions of others in the acknowledgements.

Chapter 2 is in part a computer-aided explanation of the detailed workings of the RF gun. It starts with an overview of the concepts behind the gun and the goals of the design. The capabilities of various relevent simulation programs are discussed, as well as my methodology of applying the codes. The codes are used to explore design alternatives, to understand the physics at work in the gun, and, finally, to predict expected gun performance.

The simulation results presented in Chapter 2 are my own work (though the codes used were in some cases created by others, who are acknowledged at appropriate points). My contribution included evaluation and generation of gun design alternatives and modification of the design to satisfy the project goals. In particular, I determined the size and shape of the focusing noses (necessary to obtain good control of the tranverse beam size in the gun over a wide range of currents) as well as the necessary on-axis field ratio between the cells (necessary to obtain a longitudinal phase-space suited to magnetic compression). I created the code `rfgun` to provide a fast, accurate design tool, and took the leading role in the commissioning of the gun and GTL at SSRL.

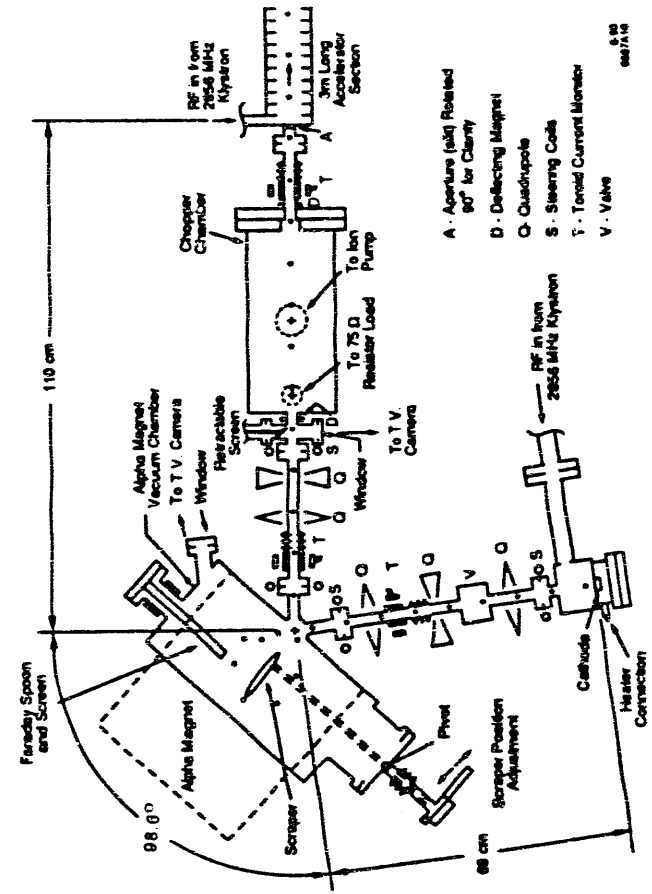

Figure 1.2: Schematic Layout of the Gun-to-Linac Region of the SSRL Preinjector

Design of the gun cannot be separated from the design of the GTL itself, which was also my responsibility. This included determination of the GTL optics, the length of the transport line, and the requirements for the alpha-magnet. While I will not discuss these details here, I did the physics design of all the GTL magnets (quadrupoles, steering magnets, and the alpha-magnet) and the beam-chopper (including the permanent magnet deflectors). My work included magnetic characterization of the magnets using a computer-aided measuring setup that I assembled and wrote the code for.

Chapter 3 goes into the properties of the alpha-magnet[45] in great detail. While the alpha-magnet was invented in 1963 and has been widely used in bunching and other applications, to my knowledge this is the first work in which full second and third order transport matrices are given. These matrices, along with highly accurate zero and first order results, are calculated by a program that I wrote, using a method that I developed and which I explain in the chapter. The method is applicable to finding matrices up to third order for any beam transport element (as is the code, if provided with the appropriate equation of motion). I also report on beam-optics experiments that I performed to test some of these calculations.

Chapter 4 discusses the subject of longitudinal dynamics, in the alpha-magnet-and-drift buncher and in linear accelerators. I present a new way of looking at the problem of matching the injected bunch longitudinal phase space, using contour maps of final phase and momentum as a function of initial phase and momentum. I discuss the well-known general principles of magnetic bunch compression, and employ results from Chapter 3 in order to show how an alpha magnet can be used for magnetic bunching, and under what circumstances. Using these ideas, I employ computer methods to find the optimal bunch-compression parameters for the SSRL RF gun. This is followed by detailed predictions of the performance of the SSRL pre-injector in terms of peak current, emittance, and brightness, as well as comparison with other projects. For this, I used the tracking code `elegant`[49], which I wrote specifically for the SSRL project.

Chapter 5 concentrates on experimental characterization of the gun and the pre-injector. It starts with a detailed walk through the GTL and continues with a discussion of the GTL optics. The remainder of the chapter gives experimental results and

comparable results of *simulated* experiments. These experiments include measurement of the momentum spectrum, the emittance, and the bunch length. I performed all of the experiments and data analysis, as well as the simulations.

Computer simulations are employed extensively in the present work, because the nature of the problems dictates that numerical methods be used for precise solutions. In order to more clearly separate my own contributions from those of others, I have adopted the convention of putting the names of computer programs that I have written in type-writer type face (e.g., "`program`") while putting the names of computer programs written by others in upper-case letters (e.g., "PROGRAM"). (This is not always adhered to in figure labels.) In addition, I have mentioned, where appropriate, what modifications I have made to programs written by others.

Some of my programs (albeit those with the least literal physics content) are described in Appendix A, and others are either described in the text or will be described in forth-coming publications[49, 50]. While these programs represent a substantial amount of original work, I have decided to concentrate on the physics that the codes predict rather than the details of the codes themselves. Appendix A is included in part to advance a new philosophy of code integration, in part as an example of the implementation of that philosophy, and in part as documentation for that implementation.

Chapter 2

Gun Design and Simulations

As discussed above, one of the RF gun's principle advantages, and indeed the essential reason for using an RF gun rather than a DC gun, is the rapid acceleration of electrons, which greatly lessens space-charge induced emittance degradation. Unfortunately, the very fact that electrons emitted from the RF gun cathode are accelerated from thermal velocities to velocities approaching the speed of light means that analytical approaches to the electron dynamics in the gun are unlikely to be fruitful. Additional complications arise from the time-varying nature of the accelerating fields, and from the continuous emission of electrons from the cathode.

Because of these complications, it is necessary to employ computer programs to simulate the detailed operation of the gun and to evaluate alternative designs. In this chapter, I will discuss many aspects of the physics of the gun as explored with various computer codes. In particular, I will concentrate on steady-state simulations of single-bunch evolution in the gun. (Multiple-bunch simulations will be addressed in future publications.) In addition to discussing the physics behind the codes and the methodology of the simulations, I will discuss the design criteria for the gun and how the codes were used to explore alternative designs. Finally, I will present and discuss collated results of simulations of the gun as it was built, giving a picture of many aspects of gun performance. Experimental results are reported in a later chapter.

The primary computer codes that I employed in this context were MASK[51], SUPERFISH[52], and `rfgun`. SUPERFISH is a well-established code that calculates the frequencies and the field distributions for TM modes of resonant cavities. In

the present work, it found application in finding higher-order mode frequencies, in "tuning" simulated cavity shapes to the desired frequency, and in computing field distributions for use by `rfgun`.

`rfgun` is a code that I wrote specifically for the SSRL RF gun project. It uses the longitudinal field profile generated by SUPERFISH and an approximation to the radial electric field and azimuthal magnetic field to calculate beam evolution in the absence of space-charge effects. Like other such codes, `rfgun` integrates the equations of motion for discrete "macro-particles", each of which represents many electrons. Its primary advantages are speed and simplicity of use, which make it a valuable design tool when coupled with the slower, more cumbersome, but also more accurate code MASK. `rfgun` also allows the investigation of the importance of various non-linear field terms by allowing the user to turn such terms on or off at will.

MASK is a "particle-in-cell" code that self-consistently integrates Maxwell's equations for the electromagnetic field and the Lorentz equation for simulated macro-particles, including the effects of space-charge. MASK's advantage over `rfgun` is that it can simulate space-charge and higher-order cavity modes, at the expense of greater complexity and greatly reduced speed. MASK is also more accurate in predicting the effects of non-linear fields near the cathode, which are poorly handled by `rfgun`'s off-axis expansion.

2.1 Gun Design Overview

The SSRL RF gun was designed as part of a larger project, the SSRL 3 GeV Injector for the storage ring SPEAR[26], and hence was required first of all to meet the needs of that project. The primary need of the Injector project was for a reliable high-current electron source that could be matched to the subsequent linear accelerator sections in such a way as to produce a beam with less than 0.5% momentum spread at 120 MeV/c. (This is discussed in more detail in Chapter 4.) The basic goal for the gun was to be able to provide 10^9 usable electrons per gun bunch, which, assuming operation at 10 pulses per second with the equivalent of two bunches accelerated per pulse and a very conservative filling efficiency of 10%, would allow filling of SPEAR to a (quite high) current of 100 mA (5×10^{11} electrons) in under ten minutes. I found in the course of my design studies that this goal was relatively easy to meet, requiring reasonable RF power, cathode current, and bunching.

2.1.1 Design Characteristics

Before examining the design criteria in detail, it is helpful to review the general characteristics of the gun[53, 54]. In doing so, I will necessarily mention many points that I will not discuss in detail until later.

The SSRL RF gun consists of a thermionic cathode mounted in the first cell of a 1-$\frac{1}{2}$ cell side-coupled 2856 MHz velocity-of-light standing wave structure, as illustrated in Figure 2.1. The gun was designed in collaboration with Varian Associates, and the basic cavity design is one used in Varian Medical accelerators. The modifications to the cavity were purposely kept to a minimum in order to reduce the magnitude of the research and development effort. This is not without its costs in terms of beam quality, since the Varian cavity is not optimized for elimination of non-linear RF fields.

Consisting as it does of three coupled resonant cavities, the gun has three possible "structure" modes[55] with frequencies near the fundamental frequency (as distinguished from higher-order modes of the individual cells, which are infinite in number). This subject is discussed in the references, and to be brief I will simply state that the gun is operated in the $\pi/2$ mode, which means that there is a phase-shift of 90°

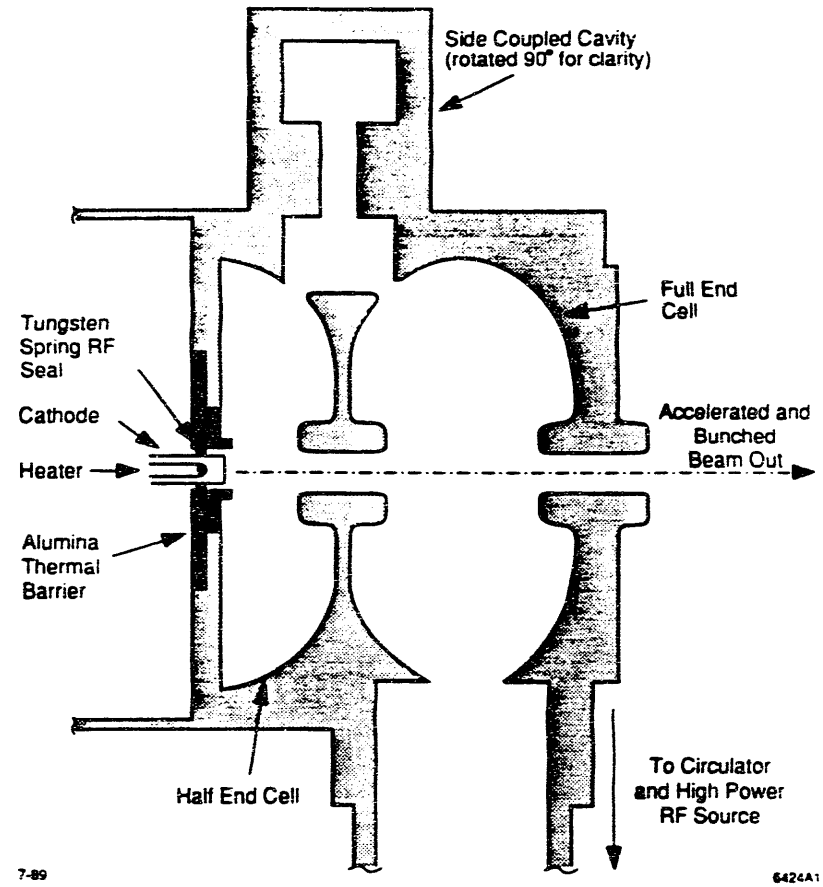

Figure 2.1: Cross Sectional View of the SSRL RF Gun

between the first cell and side-coupling cell, and between the side-coupling cell and second cell. Hence, the fields in the first and second cells are 180° out of phase. RF is fed into the gun through a port in the second cell, via rectangular wave guide, as shown in Figure 2.1, and fills the first cell through the coupling cell.

The design of the Varian side-coupled structure is such that the length of a full cell is one-half the free-space wavelength of 2856 MHz radio waves:

$$L = \frac{\lambda}{2} = \frac{c}{2f}, \tag{2.1}$$

where L represents the periodic length of the structure and $\lambda = 10.497$cm. For an accelerator consisting of a long chain of such cells, a relativistic particle that arrives at the center of one cell at the RF crest is guaranteed to arrive at the center of all subsequent cells at the RF crest in each cell. Hence, it will achieve maximum acceleration from each cell[36]. This is why this structure is referred to as a "velocity-of-light" structure. The RF gun is a very short version of such a structure, with the electrons being emitted from a cathode mounted in the end-wall. (The end wall in the first cell does not change the fields in the remaining part of the gun, since the end wall is placed at a location of symmetry.)

The cathode for the RF gun is a Varian dispenser cathode[33], with a flat, circular emitting surface of 6mm diameter, capable of current densities of up to 140A cm^{-2}. It is mounted in a modified version of Varian's high-voltage isolation mounting, which is used in Varian's DC guns. In order to provide focusing of the electron beam in the first cell, we have put a metalized ceramic annulus around the cathode. This annulus and the cathode must be in RF electrical contact with the metal walls of the RF cavity in order to avoid distortion of the RF fields. One way to achieve this is using an RF choke[32]. For the SSRL gun, a simpler concept was developed.

RF electrical contact between the cathode stem and the annulus is achieved by a toroidal tungsten spring around the cathode stem. This spring fits snugly into a toroidal cavity in the annulus. Electrical contact between the outer diameter of the annulus and the cavity is achieved through a knife-edge on the annulus that bites into a soft metal O-ring that rests in a toroidal channel in the back side the cavity wall. The reason for making the annulus out of ceramic is to provide a heat-barrier to prevent heat from being conducted too readily into the metal walls of the cavity,

which would cool the cathode. If this "heat leak" were too great, the cathode might not achieve the temperature (> 900°C) necessary for emission. The tungsten spring serves a similar purpose, in that it connects the cathode and annulus electrically without providing an easy path for heat flow.

Pyrometric measurements, shown in Figure 2.2, demonstrate that it is an easy matter to achieve cathode temperatures in excess of 1050°C, which is the approximate temperature required to "convert" the cathode (which refers to a chemical change that must occur before good emission is obtained). In operation, the cathode temperature is closer to 950°C. The Figure shows measurements that I took using two different pyrometers, corrected for the emissivity of tungsten. I found that the temperature variation across the cathode was less than 5°C over the entire cathode surface, and less than the measurement resolution of about 2°C out to about 90% of the cathode radius.

2.1.2 Gun Operating Cycle

In the steady-state, as long as RF is supplied to the gun, the gun operating cycle repeats itself every 350 ps (i.e., at the RF frequency). Electrons that start from the cathode during the accelerating phase of the RF in the first cell are initially moving at thermal velocities. The RF fields accelerate these electrons rapidly, as a result of which a beam is injected into the second cell (provided the fields are high enough). By virtue of the time it takes for the beam to get from the cathode to the second cell, the particles arrive in the second cell during the accelerating phase of the fields in that cell. Hence, the second cell continues the acceleration of the beam that began in the first cell. In a matter of about 330–360 ps after being emitted from the cathode (how long depends on the field), the first electrons of the bunch are ejected from the gun. During that same time, electrons that do not make it out of the first cavity are back-accelerated into the cathode, causing additional heating of the cathode surface. In addition, some particles that do not make it out of the second cell (because they entered too late in the accelerating phase of that cell, with too little momentum), will be back-accelerated into the first cell, and contribute to the back-bombardment of the cathode. I will return to these points in the next section, where I will show a series of

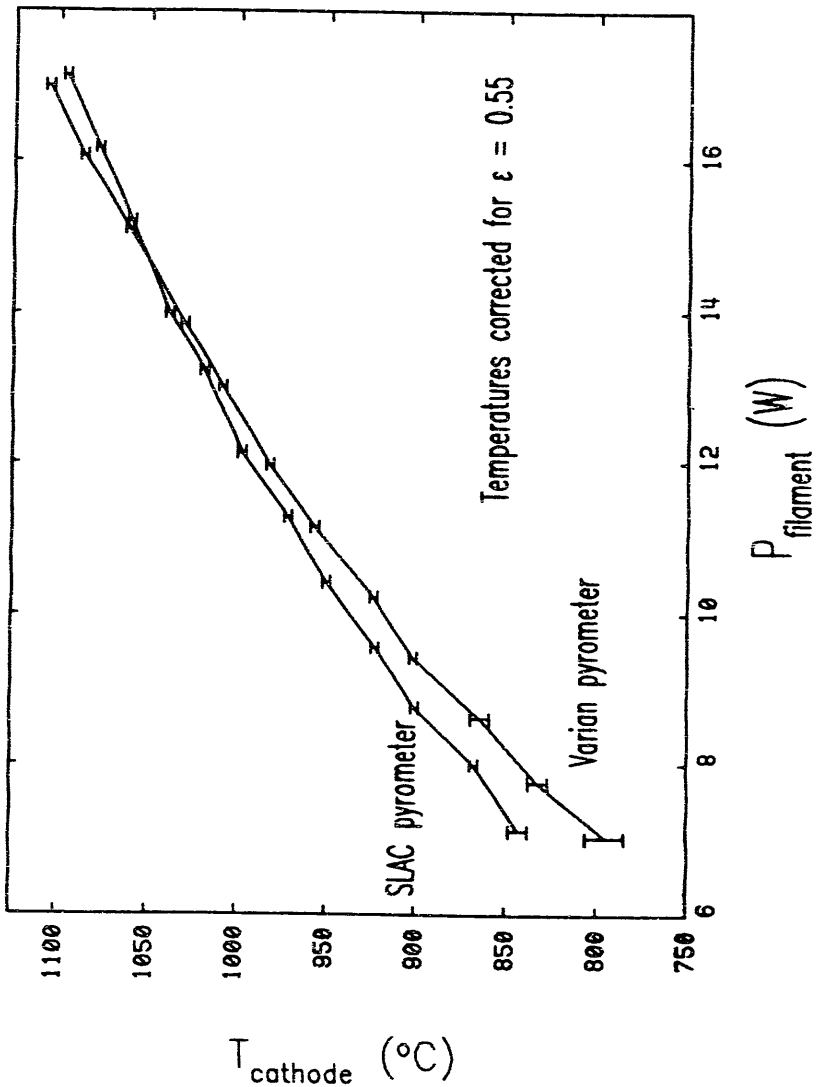

Figure 2.2: Pyrometric Measurements of the Gun Cathode Temperature

"beam snapshots" (from MASK simulations) at various points in the operating cycle.

Those particles that exit the gun comprise the "gun current" or "the beam". Because of the RF nature of the accelerating fields, the gun current for each cycle has a large spread in momentum, from some maximum down to essentially zero. However, we shall see that about half the particles in the beam have at least 80% of the peak momentum. Similarly, there is a wide spread in exit times, but about half the particles are within 25 ps or so of the particle with the peak momentum.

2.1.3 Matching to the RF Source

There are two principle limitations on the gun current. One is the current density that is available from the cathode. The other is the amount of RF power that is available to accelerate that current. Extraction of current from the gun requires a certain minimum electric field level in the cavity, otherwise current emitted from the cathode will not be accelerated sufficiently rapidly to make it out of the gun before the RF fields reverse sign. In order to maintain the electric fields in the cavity and accelerate electrons, one must supply sufficient RF power to compensate the power that goes into the beam, as well as the power that must be dissipated in the cavity walls to maintain the electric fields.

Early design studies on the gun indicated that 1–1.5 A of current at an average kinetic energy of 2 MeV was feasible. In addition, it was anticipated that 5 MW of RF power could be supplied to the gun. The beam power is simply the product of beam energy and current, from which one concludes that 2-3 MW of RF power must be supplied "for the beam" in this case. An additional power loss of about 1 MW occurs in the walls of the RF cavity, as a result of creating the electric fields that provide acceleration to 2 MeV. This makes a total of 4 MW, which is conservatively below what we anticipated would be available. The cavity must be "matched" to the RF source in order to make best use of the available RF power in the presence of beam. The cavity was thus made to be "over-coupled", with a normalized load impedance of $\beta = P_{beam}/P_{wall} + 1 \approx 4.15$. (This is topic is discussed in the references[57, 58].)

2.1.4 On-Axis Field Profiles

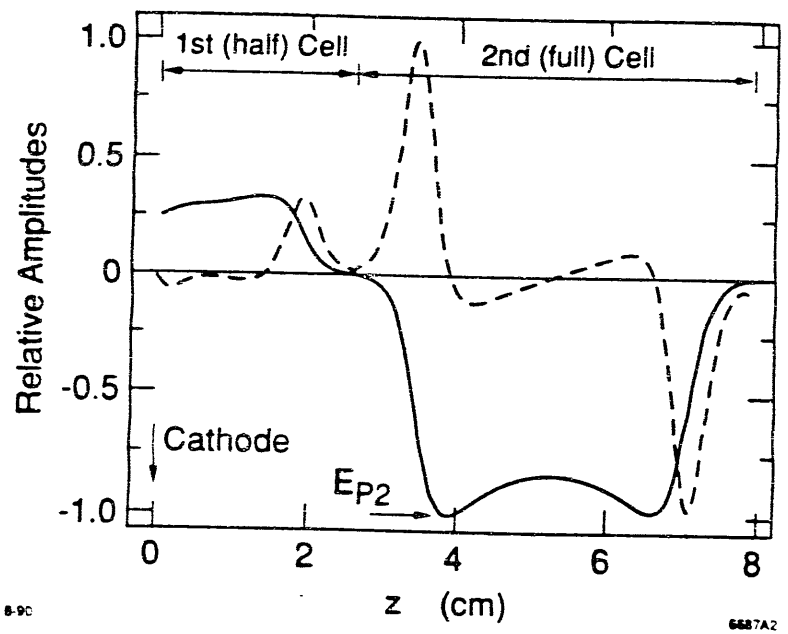

Figure 2.3: Gun On-Axis Longitudinal and Radial Electric Field Profiles

Computer studies also showed that in order to optimize the bunch longitudinal phase-space for magnetic compression, it would be necessary to delay the arrival of electrons in the second cell. The most straight-forward way to do this was by accelerating the electrons less rapidly in the first cell. Hence, the field amplitude in the first cell is approximately one-third of that in the second cell, as shown in Figure 2.3 (calculated with SUPERFISH). This was accomplished by modifying the coupling slots that connect the on-axis cells and the coupling cell[44].

The on-axis fields in the gun are related to those shown in Figure 2.3 (which is for the first RF gun, as built) by the multiplicative factor $E_{p2}\cos(\omega t)$, where E_{p2} is the peak, on-axis field in the second cell. E_{p2} provides a convenient measure of the excitation level of the RF fields, and I shall use it for this purpose throughout my discussion. Another important quantity related to the excitation level of the cavity is the peak surface field, E_{ps}. This field is of course proportional to E_{p2}, and since there are breakdown limitations on how high E_{ps} may be, there are limits on how high E_{p2} may be. For the RF gun as it was built, SUPERFISH gives $E_{ps} \approx 2 \cdot E_{p2}$. For 2856 MHz RF in a copper cavity with a surface finish of the quality that is achieved in the RF gun, $E_{ps} \leq 240$MV/m is feasible[44]. Typically, Varian operates its accelerators with $E_{ps} \leq 165$MV/m[44]. We will see below that the optimum operation of the RF gun is comfortably below the breakdown limit.

2.1.5 Design Goals

Having given an overview of the gun and the concepts involved, I am now in a position to list and briefly discuss the design goals [53, 54]:

1. There should be at least 10^9 "usable" electrons per S-Band bunch, for cathode current densities of less than 100A/cm^2, i.e., at least 10^9 electrons per bunch with momentum greater than 80% of the maximum momentum. This momentum window was established because it seemed feasible to transport a $\pm 10\%$ momentum spread from the gun to the linear accelerator without excessive losses, and because this momentum range turns out to contain about half the current that exits the gun.

2. The average momentum in the useful beam should be 2 to 3 MeV/c. for surface fields below the breakdown limit. The primary reason for choosing this momentum range was to reduce the influence of space charge after the gun. Operating in this momentum range also results in more efficient extraction of current from the gun than does operating at lower momenta.

3. The longitudinal phase-space should be suitable for magnetic compression. As I will discuss in detail in Chapter 4, this implies that the momentum-vs-time characteristic of the gun longitudinal phase-space should be near-linear and monotonically decreasing with time.

4. The focusing structure around the cathode should provide a gently-converging beam for a wide range of current densities. Since it was initially not known what current densities would be achievable in normal operation, it was desirable that low current beams should not be over-focused and that high current beams should show some convergence. In the absence of focusing noses, the beam would fill the exit tube in the first cell or even hit the cell noses (for high current density).

5. The normalized RMS emittance for the useful beam, defined by equation (1.9), should be less than $15\pi \cdot m_e c \cdot \mu m$ over the entire range of current densities, where the averages are taken over the useful electrons in the beam. This corresponds to a geometric emittance of less than about $3\pi \cdot$ mm $\cdot$ mrad, and was selected based on what seemed feasible from initial studies.

6. The average beam power returning to hit the cathode should be manageable, i.e., not greater than the filament power (about 11W) used to heat the cathode, and preferably below 5 W, in order to ensure stable operation and long cathode lifetime.

2.2 Simulation Codes and Methodology

As I indicated in the introduction, the nature of particle motion in the RF gun is such that analytical methods are of little use if one wants detailed, accurate predictions. (This is less so for photocathode RF guns, where the beam is emitted in a short pulse, triggered by a short laser pulse[37, 12, 13].) Electrons go from thermal velocities at the cathode to relativistic velocities in the space of a few centimeters, meaning that neither non-relativistic nor highly-relativistic approximations are adequate. What is more, the rate at which an electron is accelerated depends on the phase at which it is emitted from the cathode. Depending on when it is emitted and what the fields are in the gun, an electron may exit the gun with $\beta = 0.98$ or $\beta = 0.01$, or it may not exit the gun at all, returning rather to hit the cathode (again, with a wide range of possible velocities). Some electrons even oscillate between the first and second cells one or more times before finally exiting the gun or hitting the cathode. When one adds to this complexity the additional complexity of space-charge effects, the problem is even more clearly out of the realm of analytical solution. Rather than attempt to find approximate analytical tools, then, I have employed numerical methods exclusively.

I will not attempt to explain the detailed workings of the codes that I have used. This is treated in the references[51, 52, 50]. While the version of MASK that I use is non-standard, the modifications I have made are primarily to the user interface. In addition, I have added a number of capabilities that were necessary for simulation of the RF gun. My version of MASK has the capability to simulate a cathode with emission limited at some specified uniform current density. It also allows one to easily inject simulation particles from one MASK run into another MASK run, which proved necessary because the two cells of the gun had to be simulated separately.

2.2.1 Tuning and Boundary Conditions

Before simulating the gun with MASK or `rfgun`, one must first check the cavity profile using SUPERFISH, in order to determine that the computed resonant frequency for the fundamental mode is as expected. Because of the off-axis coupling cell, it is not possible to run SUPERFISH for the entire gun, since SUPERFISH is for cylindrically symmetric cavities only. Hence, I have run SUPERFISH for the first and second cells

separately. Simulation of the entire gun cavity could be done with a three-dimensional code like MAFIA[59], but I have not done this.

If the cavity had a uniform $\pi/2$-mode, then the longitudinal field at the junction between the cells (i.e., at $z = \lambda/4$, measuring from the end wall of the first cell) would necessarily be zero for the $\pi/2$ mode, since the fields in the two cells are 180° out of phase. In this case, one would use Dirichlet boundary conditions at $z = \lambda/4$ for both cells. The boundary condition at $z = 3\lambda/4$ is less clear, since there is no following cell to provide symmetry—the boundary condition is a combination of Neumann and Dirichlet. Dirichlet boundary conditions would be appropriate at $z = 3\lambda/4$ only if there were following cells to provide the necessary symmetry. What is more, since the RF gun has a non-uniform $\pi/2$-mode, one cannot conclude that Dirichlet boundary conditions are appropriate at the $z = \lambda/4$ boundary either. Because one cannot decide exactly what the boundary conditions should be without first simulating both cells (which SUPERFISH cannot do) and the structure following the second cell, the problem is in fact only solved by implicitly giving up the approximation of independent cells. Rather than do this, I elected to use Dirichlet boundary conditions at $z = \lambda/4$ and $z = 3\lambda/4$, since these are most likely to be closest to the actual boundary conditions.

To see that this is a justifiable approximation, consider that the cutoff frequency of the beam tube (which has a radius of $R_t = 3.8$mm) is[31]

$$f_{cutoff} = \frac{2.405c}{2\pi R_t}. \tag{2.2}$$

or approximately 30 GHz, compared to f = 2.856GHz for the fundamental mode of the gun. Hence, the fields in the beam tube should fall off rapidly in moving into the beam tube from either cell. The 1/e distance is given by

$$d = \frac{c}{2\pi\sqrt{f_{cutoff}^2 - f^2}}, \tag{2.3}$$

which comes out to 1.6mm for the present case, compared to about 7mm for the distance from the beginning of any cell nose to the nearest boundary plane. Hence, one expects that the fields in the cells and the resonant frequency of the cells will be insensitive to the boundary conditions, and this is indeed what I have found to be the

case. I find that the resonant frequencies of both cells change by less than 400kHz in changing from Dirichlet to Neumann boundary conditions in SUPERFISH. The on-axis longitudinal fields differ visibly only near the boundaries, and then only by a small fraction of the peak field.

The actual dimensions of the gun cavity are different from the original design. In particular, the re-entrant noses (not the focusing noses, but those belonging to the Varian cell design) were of different lengths than specified in the design. This was a result of cavity tuning during machining. All of the data I present will be for the first RF gun as it was actually built (which I refer to as "the gun as built"), unless otherwise noted.

I found that when the cavity shapes for the gun as built were put into SUPERFISH, the frequencies of the cells were different from each other and from measurements, with the second cell calculated frequency being about 2838 MHz. This discrepancy is most probably a result of imprecise knowledge of the exact actual cavity dimensions and the effect of the coupling slots and the coupling cell. Upon modeling the second cell in MASK, I found that I obtained very nearly the same frequency as in SUPERFISH. I decided to tune the first cell to 2838 MHz also, as this frequency is within 0.5% of the goal of 2856 MHz (I will show below that the error introduced by this is small). Even after tuning the first cell to within 1 MHz of 2838 in SUPERFISH (which was done by slight alterations of the upper radius of the cell nose), I found that the frequency given by MASK was 30 MHz low. This is probably attributable to the coarseness of the mesh in MASK, implying that the agreement obtained for the second cell was fortuitous. I found it necessary to insert an artificial tuning plug into the MASK simulation of the first cell (see below). This plug is far from the beam and makes no significant change in the on-axis fields.

Figure 2.4 shows the SUPERFISH-generated field lines for the first and second cell. Table 2.1 lists SUPERFISH-generated parameters for both cells, some of which are self-explanatory, others of which are explained below.

While SUPERFISH calculates cavity resonant frequencies explicitly, MASK calculates only the time-dependent field evolution. In order to find the resonant frequency of a cavity simulated in MASK, I "hit" the cavity with a relatively broad-band signal and looked at the frequency of the ringing. How this was done is discussed below. For

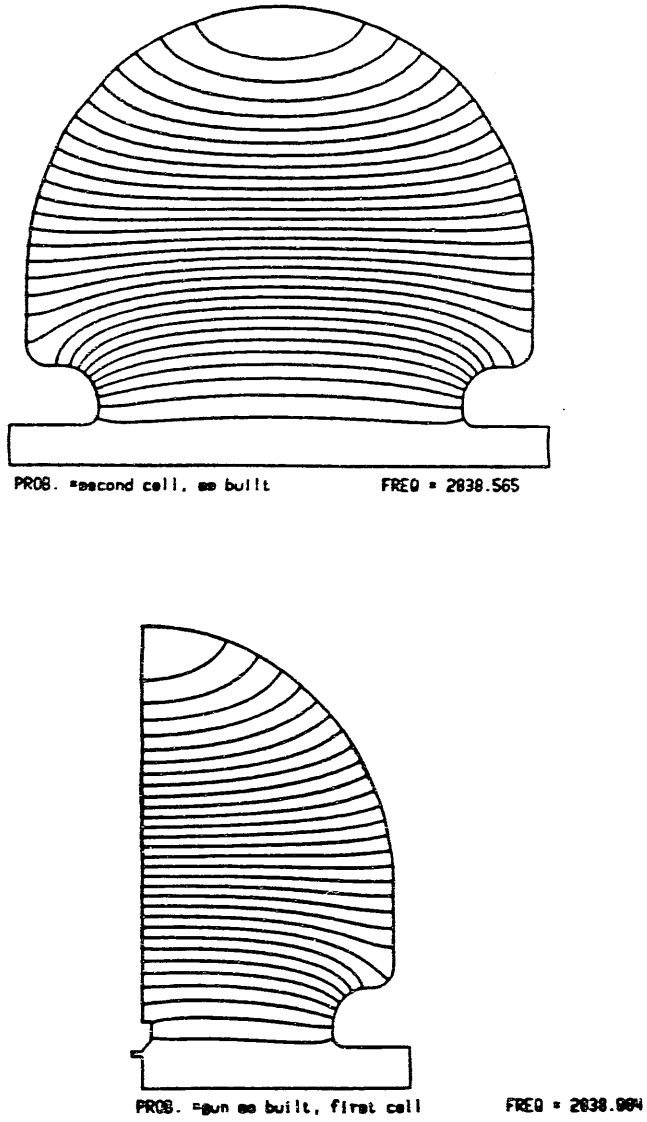

Figure 2.4: SUPERFISH Field Line Plots for RF Gun Cells

Table 2.1: **SUPERFISH-generated Cell Parameters**

quantity	Cell 1	Cell 2	unit
length	2.624	5.248	cm
Q	11777	18671	
shunt impedance	2.44 100	8.45 161	MΩ MΩ/m
K_i	0.0831	0.1473	$mJ/(MV/m)^2$
E_{ps}/E_{pi}	1.79	1.92	
transit time factor	0.766	0.787	
effective length V/E_{peak}	1.82	3.45	cm

now, suffice it to say that the simulated frequencies of the first and second cells were 2834.5 MHz and 2837.8 MHz, respectively, where I determine the frequency from the time between subsequent zero-crossings of the electric field at some fixed point in each cell. This is a valid procedure provided there is only one mode that is appreciably excited, which I took pains to ensure was the case, as I discuss below. Fourier analysis would have required simulating the cells for 1 μsec in order to determine the frequencies to within 1 MHz. This would have taken about 20 days of dedicated computation by SSRL's VAX 8810. While these frequencies are the same to within 0.1%, the difference is not negligible and must be compensated for.

The final cavity profiles used in MASK are shown in Figures 2.5 and 2.6. The solid lines show the desired profile (which is the actual profile, except for the tuning artifice in the first cell), while dots show the grid points that were filled with "metal" in MASK in order to achieve that profile. The choice of boundary conditions in MASK is even more complicated than for SUPERFISH, since the beam-induced fields have no definite symmetry and are not constrained to frequencies below cutoff. I chose to use Neumann boundary conditions in all of my MASK simulations, because the copy of the code that I have does not implement Dirichlet boundary conditions. I have verified that the influence of the boundaries on the beam is negligible by simulating the first cell with a long exit-tube, and comparing the results to a simulation which ends 0.43 mm (one longitudinal grid spacing) after $z = \lambda/4$.

Figure 2.7 shows on-axis longitudinal field amplitudes from SUPERFISH and

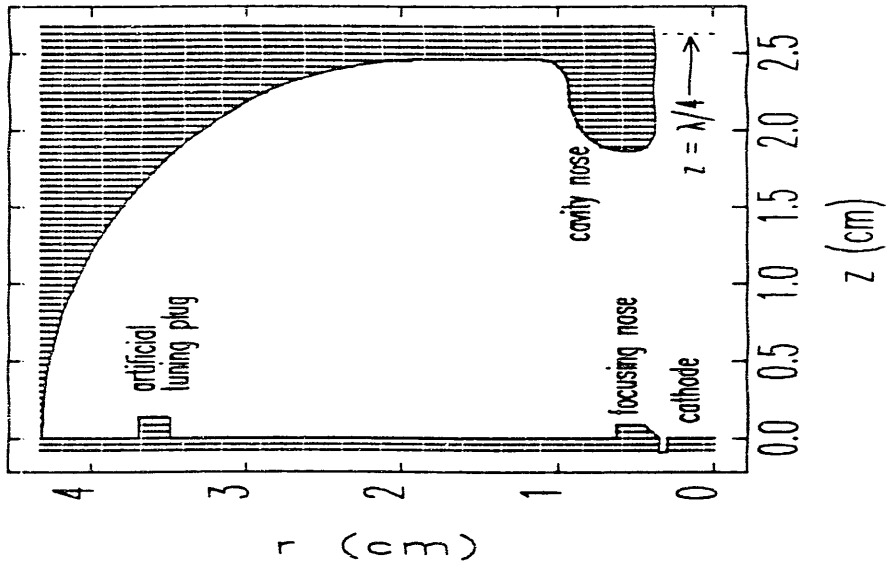

Figure 2.5: Profile Used in MASK for the First Cell

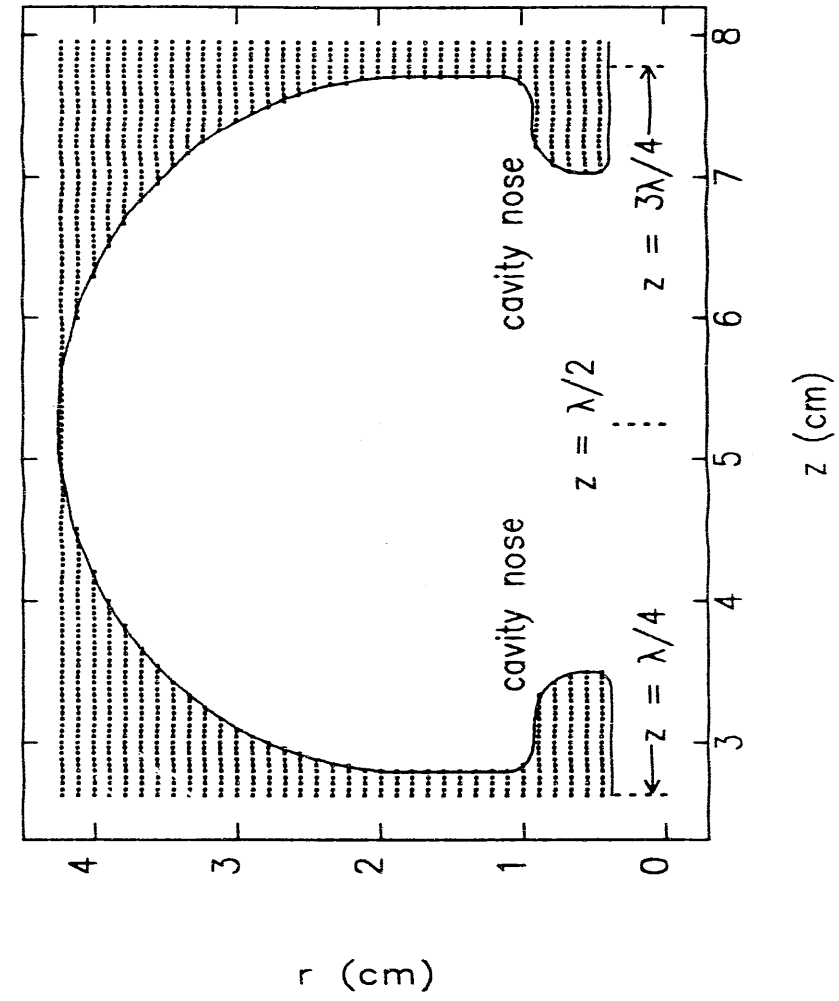

Figure 2.6: Profile Used in MASK for the Second Cell

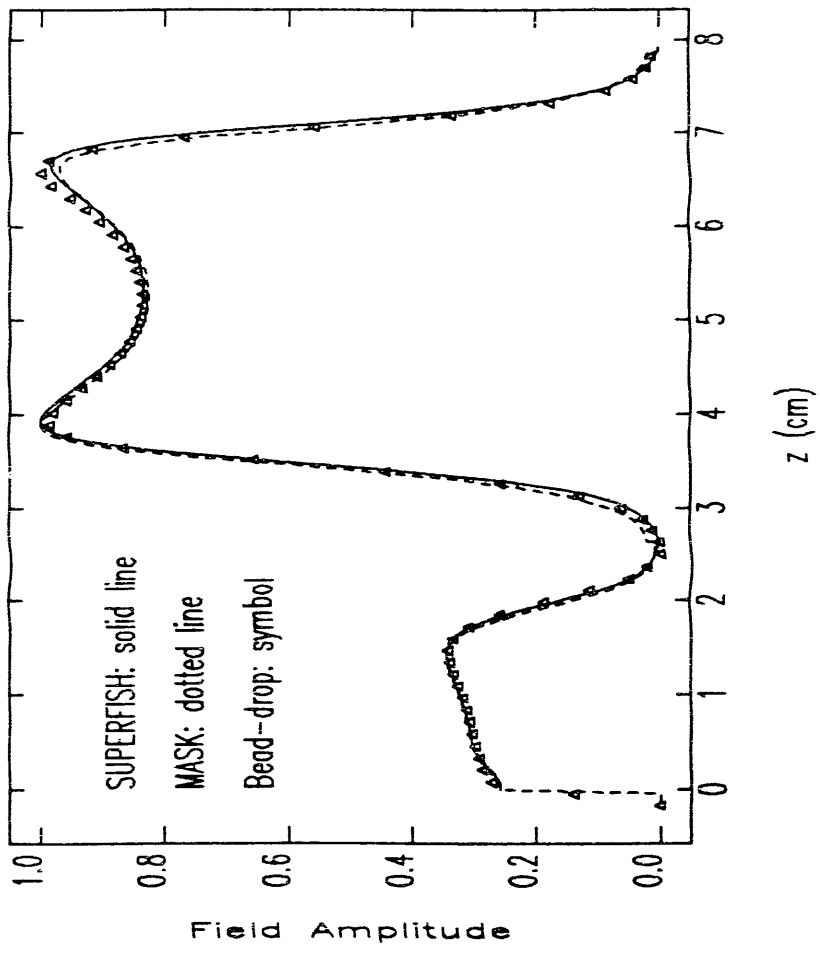

Figure 2.7: Longitudinal Field Profiles from SUPERFISH, MASK, and Bead-Drop Measurement

MASK, along with the results of a bead-drop (or "bead-pull") measurement[60] taken at Varian. Each profile is normalized so that the maximum value is 1. Note the small differences between SUPERFISH and MASK at the cell boundary. These are a result of the boundary conditions used in MASK, as just mentioned. There are also some differences near the electric-field peaks, due to inaccuracy in simulating the cell noses and beam-pipe radius in MASK (a result of the coarseness of the grid).

2.2.2 Gun Cavity Parameters

While SUPERFISH gives parameters for the individual cells directly, it does not give results for the gun as a whole. I will digress briefly to show to obtain such results. These will prove useful in analysis of experimental results in Chapter 5. The definition of the Q for either cell is[56]

$$Q_i = \frac{\omega U_i}{P_i}, \tag{2.4}$$

where U_i is the stored energy and P_i is the power dissipated in the cavity walls, for the i^{th} cell, where i is 1 or 2 for the first or second on-axis cell (I ignore the coupling cell, since there is no field stored in it in the $\pi/2$ mode, as discussed earlier in this chapter). The stored energy may be expressed in terms of the peak, on-axis electric field as

$$U_i = K_i \cdot E_{pi}^2, \tag{2.5}$$

where the K_i are constants that can be deduced from SUPERFISH output (see Table 2.1). Since $E_{p1} = E_{p2}/\alpha$, where $\alpha \approx 3$ is a constant, the total stored energy is

$$U = U_1 + U_2 = U_2(1 + \frac{K_1}{K_2\alpha^2}). \tag{2.6}$$

This form is convenient, since if K_1 and K_2 are of the same order, as one expects, then U is dominated by U_2. Similarly, the total power lost in the cavity walls is

$$P = P_1 + P_2 = P_2(1 + \frac{1}{\alpha^2}\frac{K_1 Q_2}{K_2 Q_1}) \tag{2.7}$$

Hence, the predicted Q of the cavity in the $\pi/2$ mode is

$$Q = \omega\frac{U_2}{P_2}\frac{1 + \frac{K_1}{K_2\alpha^2}}{1 + \frac{1}{\alpha^2}\frac{K_1 Q_2}{K_2 Q_1}} \tag{2.8}$$

$$= Q_2 \frac{1 - \frac{K_1}{K_2 \alpha^2}}{1 + \frac{1}{\alpha^2}\frac{K_1 Q_1}{K_2 Q_2}},$$

from which one sees that for K_1 and K_2 of the same order and $\alpha^2 \gg 1$, then $Q \approx Q_2$. Table 2.2 lists the results of cold-test measurements performed at Varian and along with inferred properties for the gun, based on this analysis. For some of the inferred properties, I have used the values of K_1, K_2, and Q_1, Q_2 calculated with SUPERFISH. I have also listed certain "desired" values, along with predictions based on SUPERFISH results, with $\alpha = 2.9$ (the measured value) used where necessary.

I have calculated the peak electric field in the second cell as a function of total wall power using

$$E_{p2} = \sqrt{\frac{U_2}{K_2}} = \sqrt{\frac{Q_2}{\omega K_2}}\sqrt{P_2}, \tag{2.9}$$

or (using 2.7)

$$E_{p2} = \sqrt{\frac{Q_2}{\omega K_2 (1 - \frac{1}{\alpha^2}\frac{K_1 Q_1}{K_2 Q_2})}}\sqrt{P}. \tag{2.10}$$

Table 2.2: **Measured and Desired Cavity Parameters for the RF Gun**

quantity	desired (predicted)	measured (inferred)	unit
Q for π 2-mode	(15008)	14000	
frequency at 20° C, air			
0-mode	—	2922.975	MHz
π 2-mode	2855.8	2855.835	MHz
π-mode	—	2802.960	MHz
$\alpha = E_{p2}/E_{p1}$	2-3	2.9	
β	4	4.15	
$E_{peak,2}/\sqrt{P_1 - P_2}$	(79.9)	(70.5)	MV m MW$^{\frac{1}{2}}$

2.2.3 Methodology of MASK Simulations

I return now to the discussion of simulations, and in particular MASK simulations. My methodology in using MASK was heavily influenced by the need to economize

computer time. A single cycle of the RF gun with beam takes approximately 30 hours of CPU on SSRL's VAX 8810. Part of the reason for this is the need to use a large number of simulation macro-electrons. The emission algorithm in MASK is such that charge is emitted at each step in order to obtain the desired current density. Hence, if the macro-electron weight (the number of electrons that each macro-electron represents) is made too large in an attempt to decrease the number of macro-electrons, the simulation may end up emitting no macro-electrons at all, because one macro-electron per time-step may exceed the allowed current density. Thus, having many simulation particles is a result in part of having a small time-step. The time-step, Δt, is chosen under the constraint of the Courant stability condition [61] (for integration of Maxwell's equations), which requires that

$$\Delta t \leq \frac{\text{minimum}(\Delta r, \Delta z)}{c\sqrt{2}}, \tag{2.11}$$

where Δr and Δz are the grid spacings in r and z, respectively. In order to accurately simulate the fields in the vicinity of the cathode, I chose $\Delta r = 0.25$mm, which gives 12 grid points across the cathode and 3 grid points spanning the recess between the cathode and focusing annulus. $\Delta z = 0.43$mm was also chosen, based on the need to have 64 grid points between the recess around the cathode and the end of the first cell (the power of 2 is required by MASK.) Hence, the time step would need to be less than about 0.6 ps. I found that a smaller time-step was needed in order to get stability (perhaps a result of the fact that MASK uses single-precision), and chose a conservative value of 0.171 ps, which is convenient in that it gives 2048 time-steps per 2856 MHz RF period.

As I will discuss below, the longitudinal mesh spacing was belatedly discovered to be somewhat larger than needed to accurately fulfill the boundary conditions on the slope of E_z at the cathode. However, this has no significant effect on the results. Test runs with smaller radial and longitudinal mesh sizes were found to give virtually identical results to those with the mesh sizes listed in the last paragraph.

I then chose the ratio of the macro-electron weight to the current density in order to ensure that several macro-electrons were emitted per time step. The number of

macro-electrons emitted per time step is

$$N_s = \frac{J\pi R_c^2}{eN_m}\Delta t, \tag{2.12}$$

where N_m is the number of electrons represented by each macro-electron. In order to obtain good statistics, it is desirable that there be several thousand macro-electrons in the simulated useful beam. For examining the results of filtering small momentum spreads from the beam, it is necessary to have even more macro-electrons. Roughly speaking, N_u useful-beam macro-electrons requires $2N_u$ macro-electrons total exiting the second cell, which requires roughly $4N_u$ macro-electrons emitted from the cathode during the accelerating phase in the first cell (the reasons for these factors will be discussed below). This analysis would lead one to conclude that, for 2048 time-steps per cycle, $N_s = N_u/256$. I chose $N_s = 15$, which gives about 4000 useful-beam macro particles and gives good statistics even for analysis of small momentum intervals. In retrospect, this is probably higher that it needs to be, but I have used in throughout in order to avoid changing simulation parameters which would confuse comparison of different MASK runs.

Because of the long running times, I decided to simulate only the steady-state behavior of the gun, that is, the behavior after the RF fields and beam current have come to stability. This happens in the last half of the 2μs RF pulse used at SSRL. Even taking $N_s = 1$ would not help to decrease the running time sufficiently to allow simulation of the entire operating cycle of the gun in a reasonable time. There are a number of assumptions upon which the validity of this procedure rests. It assumes that it is sufficient to simulate only the $\pi/2$ structure mode, which allowed simulation of the cells separately. If, for example, the beam drives significant power into the zero or π modes, this approximation would be invalid. Since MASK cannot simulate the coupling cell, there was little choice about this.

This methodology also assumes that each bunch sees the cavity with only the fundamental mode excited—i.e., that the higher-order-mode fields excited by previous bunches have no significant influence on any particular bunch. This is equivalent to saying that the only significant effect of previous bunches is to remove power from the fundamental cell mode. In the steady-state, this power is assumed to be replaced by the RF power source, so that each bunch sees that same fields; hence, if the

assumptions are valid, only one cycle needs to be simulated with particles. Similarly, this methodology assumes that electrons left in the gun from the previous operating cycle have no significant influence on newly emitted particles. These assumptions can be checked by simulating the gun for several operating cycles; results of such a simulation are present in the next section. Normally, however, I simulate only a single operating cycle with beam, in addition to simulating several RF cycles for the build-up of the cell fields.

Since MASK is a time-dependent code, the cavity fields must be built up by means of some suitable simulated RF power source. The program allows the simulation of both RF ports and antennae. I chose to use an antenna because of the greater simplicity. Since I wished to simulate only the steady-state behavior of the gun, it made no difference how the cavity was driven. (I was not attempting to simulate the evolution of the beam during the charging of the cavity.) In driving an RF current in the antenna, MASK creates fields in the vicinity of the antenna that propagate throughout the cavity via Maxwell's equations. The sinusoidal RF current was modulated by the envelope shown in Figure 2.8, in order to avoid excitation of higher-order modes (the particular shape is composed of two cubic splines, using a standard feature of MASK). In this way, the Fourier amplitudes excited in the first three higher-order modes were kept below 10^{-3} of the fundamental. In addition, the antenna was placed at the intersection of nodes of the first two higher-order modes, to reduce the excitation of these modes even further (shaping the current envelope is by far the most important consideration).

In general, a simulation of the gun with MASK consisted of the following steps:

1. The fields in the first cell were excited to the desired level, or else the fields saved on disk from a previous run were read in (and optionally scaled). Once the driving current envelope has fallen to zero, the fields can be saved for use in subsequent runs, or used immediately in the next stage.

2. Particle emission from the cathode, limited to some fixed current density and by space-charge forces, was then allowed for one RF period, beginning during the accelerating phase of the fifth RF period since the excitation began. During this period, MASK "pushes" the macro-electrons through the first cell, in addition

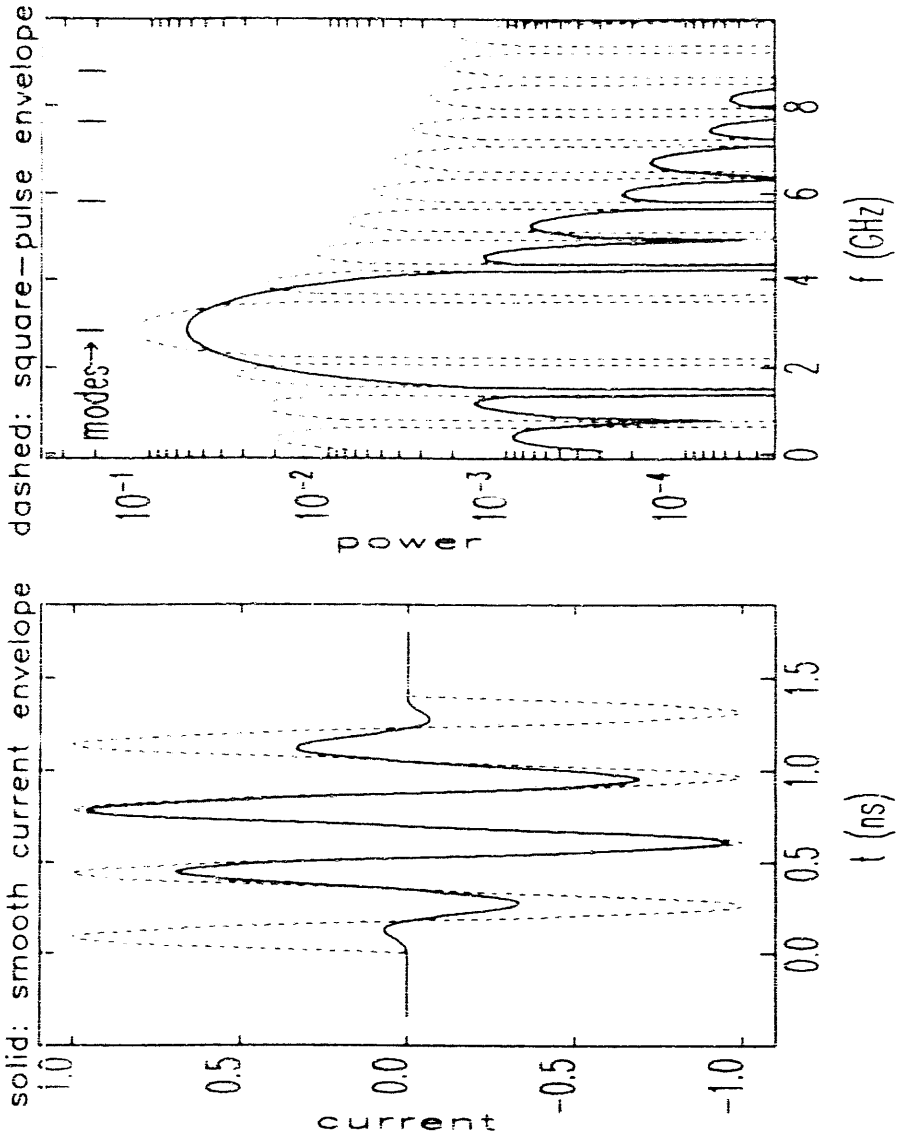

Figure 2.8: RF Current Waveform for Exciting Cells in MASK

to simultaneously integrating Maxwell's equations. As macro-electrons pass $z = \lambda/4$, the end of the first cell, their phase-space coordinates are saved to disk.

3. The fields in the second cell were excited to the desired level or read in from disk, as for the first cell. Of course, these fields should be 180° out of phase with the first-cell fields. (In fact, they were *driven* 178° out of phase, to compensate for the difference in frequency, as discussed below.)

4. Macro-electrons from the simulation of the first cell were injected at $z = \lambda/4$ in the second cell. Since both simulations start at time $t = 0$, each macro-electron is injected into the second simulation when the time counter is the same as it was when the macro-electron left the first simulation.

The measured field ratio in the gun is $\alpha = 2.9$. The MASK simulations reported on here used $\alpha = 3.0$, as a result of a mis-reading of the data that was not noticed until the simulations had already been run. Because the simulations take so long, time does not permit me to repeat them. I will show below that this error has only a small effect on the results. Given the many other approximations being made (e.g., independent bunches, ignoring the three-dimensional nature of the cavity with coupling holes having the simulation cell frequency differ from the actual frequency), this is not a serious error.

2.2.4 Cathode Simulation

Cathode emission was simulated by injecting a total charge of $\pi R_c^2 J_{emitted} \Delta t$ during each time interval, where R_c is the cathode radius, $J_{emitted}$ the emitted current density, and Δt the time interval (i.e., the time-step for the integration). The space-charge limitation on the current density was assumed to follow the Gaussian emission law (a standard feature of MASK), namely,

$$-(J_{emitted}\Delta t - Q_{surf}) \le -\epsilon_o E_{z,cathode} \qquad (2.13)$$

where $E_{z,cathode}$ must be negative for the emission of electrons (negative J), and where $Q_{surf} \le 0$ is the existing "surface charge" (i.e., charge just above the cathode surface)

due to previously emitted macro-electrons. In other words, the emitted charge density cannot be so high that, in combination with previously emitted charge density, the newly emitted charge depresses the field at the cathode so much that the sign of the field is such as to push electrons back into the cathode. In addition I required that

$$-J_{emitted} \leq -J_{limit}. \tag{2.14}$$

where J_{limit} is the user-specified maximum current density available from the cathode, which in actual operation is controlled by controlling cathode temperature. In addition to adding this feature, I modified MASK to emit a bona-fide uniform random distribution from the circular cathode surface (the standard code uses an approximation to such a distribution). Each macro-electron emitted in the simulations has the same charge, in order to make the interpretation of graphs and other data more straight-forward. This implies that in order to have a uniform emission over the surface, fewer macro-electrons should be emitted near $r = 0$ than near $r = R_c$. The reader will see this effect in some of the graphs that follow.

It is interesting to look at the details of this emission process. SUPERFISH calculations give the peak field at the cathode as $E_{pc} = 0.258E_{p2}$ (assuming $\alpha = 2.9$). Hence, for a typical operating value of $E_{p2} = 75$ MV/m, E_{pc} is about 20 MV/m. As the field at the cathode passes into the accelerating phase for electrons, emission begins. Around this time (call it t=0), the field at the cathode may be approximated as $E_c \approx -E_{pc}\omega t$. Hence, if the integration time step is Δt, and if I momentarily assume $J_{limit} \rightarrow -\infty$, then

$$J_{emitted}(t = n\Delta t)\Delta t = -Q_{surf}(t = n\Delta t) - \epsilon_o E_{pc}\omega n\Delta t \tag{2.15}$$

where n is the number of time steps since t=0. Emission can occur as long as $J_{emitted} < 0$.

One expects that, initially, the surface charge will simply be the sum of all charge emitted since t=0, since the fields have not had time to accelerate that charge away from the cathode. It is possible to approximate the distance a macro-electron will move and confirm this expectation. Ignoring the variation of the field with position, the Lorentz equation gives

$$\frac{dp}{dt} \approx \frac{eE_{pc}\omega t}{m_e c}. \tag{2.16}$$

where $p \equiv \beta\gamma$. For a macro-electron that starts from the cathode at t=0,

$$p(t) \approx \frac{eE_{pc}\omega t^2}{2m_e c}. \tag{2.17}$$

For a non-relativistic particle, the velocity is approximately pc, so that the position is

$$z(t) \approx \frac{eE_{pc}\omega t^3}{6m_e}. \tag{2.18}$$

For $E_{pc} = 20$MV/m,

$$z(t) \approx 10^{-6}t^3, \tag{2.19}$$

where z is in meters and t in pico-seconds. The cathode radius, $R_c = 3$mm, provides a natural length scale, and one sees that it will take about 70 ps for the first macro-electron to move by R_c. In other words, the effects of space-charge on emission might be expected to be quite large, since it takes so long for emitted charge to be moved away from the cathode.

Since the charge is moved away from the cathode so slowly, one can assume that the surface charge is just the sum of all previously emitted charge:

$$Q_{surf}(t = n\Delta t) = \sum_{i=1}^{n-1} J_{emitted}(t = i\Delta t)\Delta t, \tag{2.20}$$

so that

$$J_{emitted}(t = n\Delta t) = -\sum_{i=1}^{n-1} J_{emitted}(t = i\Delta t) - \epsilon_o E_{pc}\omega n. \tag{2.21}$$

Subtracting this expression from the same expression with $n \rightarrow n-1$, one obtains

$$J_{emitted}(t = n\Delta t) = -\epsilon_o E_{pc}\omega. \tag{2.22}$$

For $E_{pc} = 20$MV/m, one obtains $J_{emitted} = -318$A/cm^2. Since the cathode is capable of no more than 140 A/cm^2, we see that the space charge limitation will never come into play. I will hereafter use $J = |J_{limit}|$ to refer to the cathode current density assumed in any simulation.

MASK supports the inclusion of a initial thermal velocity distribution for the macro-electrons emitted from the cathode. I found that inclusion of this effect made

no significant difference in the results. To see why, consider that the RMS thermal momentum ($\beta\gamma$) is given by the Maxwell velocity distribution[62] as

$$p_{rms} = \sqrt{\frac{3kT}{m_e c^2}} \tag{2.23}$$

A cathode temperature of 1000°C or 1273°K gives $p_{rms} \approx 10^{-3}$. The RMS radial momentum will be of the same order. During acceleration, this radial momentum be adiabatically damped. Assuming acceleration to a typical value of $p_z = 5$, the slope due to the thermal velocity would be of order p_{rms}, $p_z \approx 0.2$mrad. A typical angular half-width for an RF gun beam is 20 mrad, and hence one sees that thermal effects may safely be ignored. This is confirmed by MASK simulations done with and without thermal effects. While adding thermal effects to MASK entails little cost in computer resources, I have ignored thermal effects in all the MASK simulations reported on here, since this permits cleaner comparison with **rfgun**, which did not originally support an initial thermal velocity distribution.

A number of different types of output are available from MASK. I have upgraded MASK extensively to improve the ease with which one can make use of these outputs; in particular, I have used my self-describing data format (**aws** format) extensively. This is discussed in more detail in Appendix A.

MASK allows one to sample the cavity fields along any line in r or z at intervals in time, as well as to sample the fields at a fixed location as a function of time. The initial coordinates of macro-electrons emitted from the cathode may be saved, as well as the phase-space coordinates (time, positions, and velocities) of every macro-electron as it crosses a number of user-defined planes of constant z coordinate. This later capability is the principle means of getting information about beam energy and emittance, and for keeping track of particles that return to hit the cathode. It is also used for the re-injection of macro-electrons from first-cell simulations into second-cell simulations. MASK keeps a record of particles lost by hitting metal boundaries or exiting the simulation region, so that the total power going into particles can be calculated. This is important for assessing how much power is needed to produce a given beam, since not all the power that goes into electrons goes into the electrons *in the beam*. MASK can also produce output containing "beam-snapshots", which record the positions of all particles at a given time during the simulation. Examples of all of these output

facilities will be seen in the next section.

2.2.5 Compensation of Cell Frequency Mismatch

I mentioned above that the phase-shift between the driving RF currents for the two cells must be different from 180° to compensate for the difference in resonant frequencies between the first and second cells[63]. The first simulated cell resonates at 2834.5 MHz, and the second at 2837.8 MHz. Particles from the simulation of cell 1 are injected into the simulation at cell 2 during the fifth RF cycle since the beginning of the RF excitation. Hence, one wants the fields in the cells to be 180° out of phase during the fifth RF cycle. The phase in each cell is

$$\phi_i = \omega_i t - \Delta\phi_i, \tag{2.24}$$

where $\Delta\phi_1 = 0$ and $\Delta\phi_2 \approx 180°$. Setting $t = 5T_o$, with T_o the average RF period of the cells, one obtains the phase difference at the beginning of the fifth RF period

$$\phi_2 - \phi_1 = 5T_o(\omega_2 - \omega_1) + \Delta\phi_2, \tag{2.25}$$

requiring $\phi_2 - \phi_1 = \pi$, one obtains the necessary value of $\Delta\phi_2$

$$\Delta\phi_2 = \pi - 5T_o(\omega_2 - \omega_1). \tag{2.26}$$

Using the values of the simulated cell frequencies that I gave just above gives $\Delta\phi_2 = 3.105$, or 177.9°. I used this value in of my all MASK simulations of the gun as built. I shall use 2836 (the average frequency) as the nominal frequency of the MASK simulations. In the next section, I shall show what the effect of this frequency error is.

2.2.6 The rfgun Program

Having dealt with MASK at some length, I now turn to **rfgun**. As I have mentioned above, **rfgun** uses the longitudinal field profiles generated by SUPERFISH (or any equivalent program) to calculate particle motion in the gun. **rfgun** works in cartesian coordinates (x, y, z), allowing the imposition of both transverse and solenoidal

external magnetic fields. The cavity fields are assumed to have a perfect sinusoidal variation in time, at some user-specified frequency. (Since the fields for the whole cell are specified at once, there is no need to worry about the individual cell frequencies.) Macro-electrons are emitted from the cathode at discrete phase-intervals and from discrete radial positions on the cathode, with macro-electron charges adjusted to achieve approximately uniform current density. The solution for macro-electron motion proceeds from the Lorentz equation, which is integrated numerically using fourth-order Runge-Kutta[61]. Like MASK, **rfgun** supplies the user with phase-space coordinates of exiting particles as well as those of particles that return to hit the cathode. Unlike MASK, **rfgun** does not include space-charge, nor does consider the positions of metal structures that macro-electrons might hit (a maximum radius may be specified, however, to simulate the collimating effect of the beam tube).

2.2.7 Off-Axis Field Expansion

It is well known[31] that the modes for a cavity like one of the RF gun cells can be separated into independent transverse-electric (TE) and transverse-magnetic (TM) modes. TM modes have $B_z = 0$ and involve only E_z, E_r, and B_ϕ, whereas TE modes have $E_z = 0$ and involve only B_z, B_r, and E_ϕ. While the gun cavity can support TE modes, the fundamental mode is TM, which has a non-zero accelerating field E_z. (Note that a cylindrically symmetric beam will induce only TM modes, since it always produces $E_r \neq 0$ and $B_\phi \neq 0$, but never $E_\phi = 0$ or $B_r = 0$.)

Knowing E_z along the axis allows one to compute series expansions in r for E_z, E_r, and B_ϕ. One starts with Maxwell's equations (in MKSA units):

$$\nabla \cdot \mathbf{E}(z,r,t) = 0 \tag{2.27}$$

$$\nabla \times \mathbf{E}(z,r,t) = -\partial_t \mathbf{B}(z,r,t) \tag{2.28}$$

$$\nabla \times \mathbf{B}(z,r,t) = \mu_o \epsilon_o \partial_t \mathbf{E}(z,r,t). \tag{2.29}$$

Expressing these in cylindrical coordinates (r,ϕ,z) and assuming that the electric fields vary like $\sin(\omega t)$ (and hence that $B_\phi \sim \cos(\omega t)$), one obtains

$$\partial_z(rE_r(z,r)) - r\partial_r E_z(z,r) = 0, \tag{2.30}$$

$$\partial_z E_r(z,r) - \partial_r E_z(z,r) = \omega B_\phi(z,r), \tag{2.31}$$

$$\partial_r(rB_\phi(z,r)) = r\mu_o\epsilon_o\omega E_z(z,r). \tag{2.32}$$

and

$$-\partial_z B_\phi(z,r) = \mu_o\epsilon_o\omega E_r(z,r), \tag{2.33}$$

(2.34)

Because of cylindrical symmetry, $E_r(z,r)$ must be an odd function of r, and hence from equation (2.30) one sees that $E_z(z,r)$ must be an even function of r. Equation (2.33) implies that $B_\phi(z,r)$ has the same symmetry as $E_r(z,r)$. Thus, I can expand these functions as

$$E_z(z,r) = \sum_{i=0}^{\infty} \tilde{E}_{z,2i}(z) r^{2i} \tag{2.35}$$

$$E_r(z,r) = \sum_{i=0}^{\infty} \tilde{E}_{r,2i-1}(z) r^{2i+1} \tag{2.36}$$

$$B_\phi(z,r) = \sum_{i=0}^{\infty} \tilde{B}_{\phi,2i-1}(z) r^{2i-1}, \tag{2.37}$$

where the tildes are used to emphasize that the coefficients do not necessarily have the same units as the fields.

Inserting expressions (2.35) and (2.36) into the divergence equation, (2.30) and equating terms with the same power of r, one obtains

$$\tilde{E}_{r,2i-1}(z) = -\frac{1}{2i-2}\partial_z \tilde{E}_{z,2i}(z) \quad i = 0, 1, \ldots \infty. \tag{2.38}$$

Inserting expressions (2.35) and (2.37) into (2.32), one obtains

$$\tilde{B}_{\phi,2i-1}(z) = \frac{\mu_o\epsilon_o\omega}{2i-2}\tilde{E}_{z,2i}(z) \tag{2.39}$$

Using equations (2.38) and (2.39) in (2.31) one obtains

$$\sum_{i=0}^{\infty}\left[-\frac{r^{2i-1}}{2i-2}\left(\partial_z^2 - k^2\right)\tilde{E}_{z,2i}(z) - 2ir^{2i-1}\tilde{E}_{z,2i}(z)\right] = 0. \tag{2.40}$$

where $\omega = kc$. Equating terms of the same power in r yields

$$\tilde{E}_{z,2i-2}(z) = -\frac{1}{(2i-2)^2}\left(\partial_z^2 - k^2\right)\tilde{E}_{z,2i}(z). \tag{2.41}$$

which completes the solution for the radial expansion coefficients by providing a recursion relation that gives all the $\tilde{E}_{z,2i}$'s, from which all the $\tilde{E}_{r,2i+1}$'s and $\tilde{B}_{\theta,2i}$'s can be obtained via equations (2.38) and (2.39).

It is useful to work out these expressions for the first few terms:

$$\tilde{E}_{r,1}(z) = -\frac{1}{2}\partial_z \tilde{E}_{z,0}(z), \tag{2.42}$$

$$\tilde{E}_{r,3}(z) = \frac{1}{16}\left[\partial_z^3 - k^2\partial_z\right]\tilde{E}_{z,0}(z), \tag{2.43}$$

$$\tilde{E}_{z,2}(z) = -\frac{1}{4}\left[\partial_z^2 - k^2\right]\tilde{E}_{z,0}(z), \tag{2.44}$$

$$\tilde{B}_{\theta,1}(z) = \frac{\omega}{2c^2}\tilde{E}_{z,0}(z), \tag{2.45}$$

and

$$\tilde{B}_{\theta,3}(z) = -\frac{\omega}{16c^2}\left[\partial_z^2 - k^2\right]\tilde{E}_{z,0}(z). \tag{2.46}$$

rfgun includes terms up to third order in r, with the option to use only the linear terms in r. This allows investigation of the importance of higher-order terms. Note that $\tilde{E}_{z,0}(z) = E_z(z, r = 0)$, where values of the latter function are given by SUPERFISH. Since it has the units of an electric field, I'll drop the tilde from $\tilde{E}_{z,0}(z)$ in what follows.

2.2.8 Non-Linear Field Terms

To see how influential the nonlinear terms might be, note that the importance of the nonlinear magnetic fields is related to their magnitude relative to the linear terms, which is characterized by the function

$$T_1(z) = \frac{R_c^2}{8E_{p2}}\left[\partial_z^2 - k^2\right]E_{z,0}(z), \tag{2.47}$$

where R_c is the cathode radius, which gives an upper limit on the radius of any particle that makes it out of the gun (as will be seen in the next section). Note that the average radius of particles emitted from a uniformly-emitting circular cathode is $\frac{2}{3}R_c$. If $T_1(z)$ is small compared to $E_{z,0}(z)/E_{p2}$, then nonlinear magnetic fields are

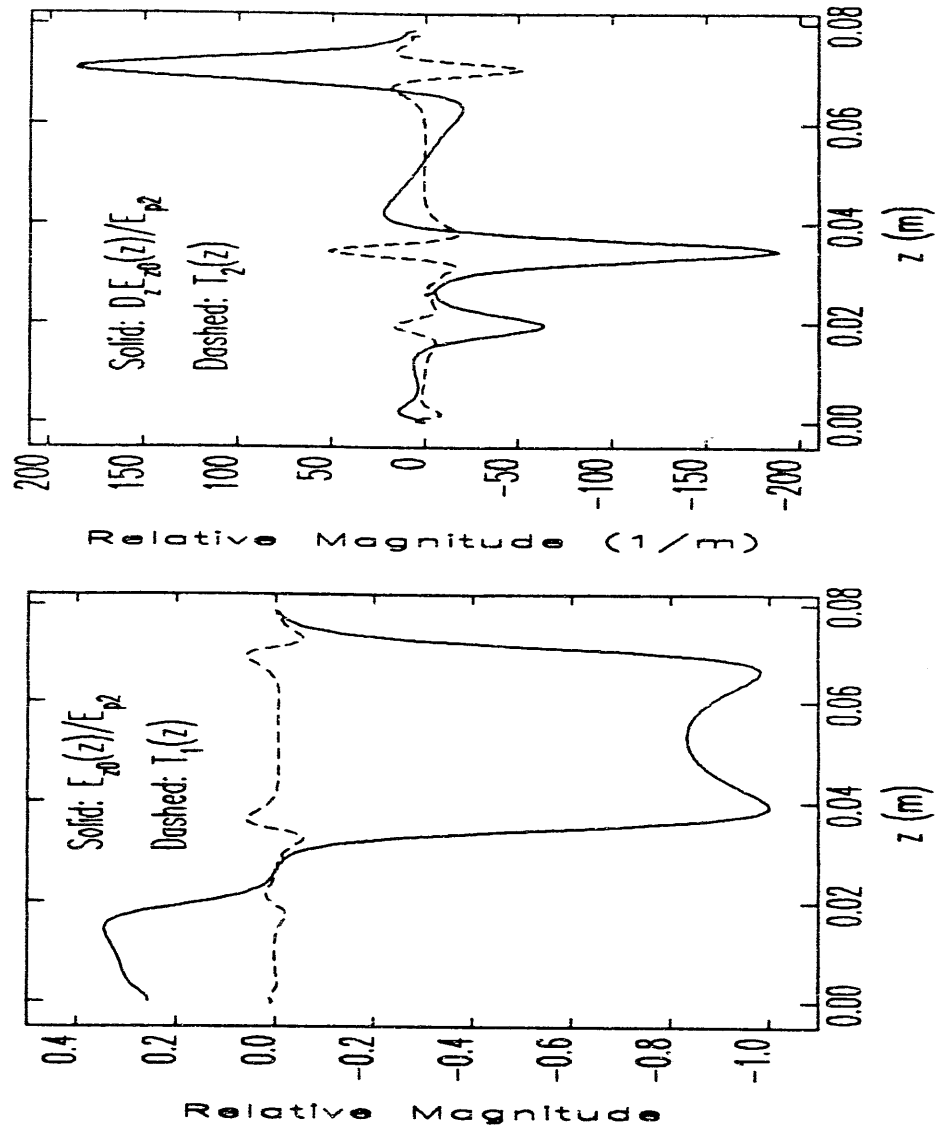

Figure 2.9: Nonlinear Field Terms in the RF Gun

unimportant. $T_1(z)$ is also related to the importance of the nonlinear longitudinal electric fields, though $4T_1(z)$ is a more accurate measure.

Similarly, importance of the nonlinear radial electric fields is characterized by the function

$$T_2(z) = \frac{R_c^2}{8E_{p2}}\left[\partial_z^3 - k^2\partial_z\right] E_{z,0}(z). \tag{2.48}$$

If $T_2(z)$ is small compared to $\frac{\partial}{\partial z}E_{z,0}(z)$, then nonlinear radial electric fields are unimportant.

Figure 2.9 compares $T_1(z)$ and $T_2(z)$ to the relevant lower-order terms. One sees that the non-linear magnetic fields are quite small, and one infers that the non-linear longitudinal electric fields are small as well. The non-linear radial electric fields are more significant, particularly in the vicinity of the cathode and near the cell noses. Of these, the fields near the cell noses are less important, since most particles of interest pass the cell noses well off the RF crest.

Finally, I note that $T_1(z)$ is completely dominated by the $\partial_z^2 E_{z,0}(z)$ term, while $T_2(z)$ is completely dominated by the $\partial_z^3 E_{z,0}(z)$ term. This is as expected, since the fields (and their derivatives) are changing rapidly on the scale of $1/k \approx 1.7$cm. I shall return to the issue of non-linear fields, specifically their effect on the beam, in the next section. For now, I simply point out that if one wants to reduce the effect of non-linearities, one must reduce $\partial_z^2 E_{z,0}(z)$ and $\partial_z^3 E_{z,0}(z)$. In particular, it is important to avoid non-linear terms in z near the cathode, where the beam is particularly sensitive (because of its relatively large radial size and its low momentum). This was not attempted for the SSRL RF gun design.

2.2.9 Boundary Conditions for rfgun

The data shown in Figure 2.9 was assembled from SUPERFISH output for the individual cells. SUPERFISH gives values of E_z along the axis at equi-spaced intervals in z. However, these values are calculated with boundary conditions for the individual cells that do not properly reflect the boundary conditions appropriate to the full cavity, as discussed above. Hence, the two solutions do not join smoothly at $z = \lambda/4$. One is tempted to smooth the solutions in order to eliminate the discontinuities in the fields and its derivatives at this point, but I have found that this can be done only at

the expense of smoothing away real, significant derivative information in the rest of the cavity. Hence, the data in Figure 2.9 as well as that used in RF gun simulations is simply composed from concatenating the solutions for the individual cells, with appropriate scaling to get $a = 2.9$ and a $\pi/2$-mode configuration.

In order to calculate the non-linear field terms, `rfgun` takes z derivatives of $E_{z,0}(z)$, up to the fourth derivative (so that the third derivative can be interpolated in between data points). Each derivative is calculated using the second-order formula[64]

$$\partial_z F(z) = \frac{F(z+\Delta z) - F(z-\Delta z)}{2\Delta z} + \mathcal{O}(\Delta z^2). \tag{2.49}$$

There is no problem in the interior of the cells, where each point has two neighboring points with well-defined field values, but problems do arise at the boundaries. In order to be able to calculate the required derivatives at the beginning and end of the gun, `rfgun` needs data outside these intervals.

I have provided this data for the beginning of the gun by applying an idea gleaned from the above analysis of non-linear fields. Since z=0 is a metallic surface for $r < R_c$, $E_r(z=0) = 0$ for $r < R_c$, and equation (2.38) implies that

$$\left(\partial_z \tilde{E}_{z,2i}(z)\right)_{z=0,r<R_c} = 0. \tag{2.50}$$

Taking the partial of equation (2.41) with respect to z, one then sees that

$$\left(\partial_z^3 \tilde{E}_{z,2i}(z)\right)_{z=0,r<R_c} = 0. \tag{2.51}$$

Taking ∂_z^3 of (2.41), one then sees that

$$\left(\partial_z^3 \tilde{E}_{z,2i+2}(z)\right)_{z=0,r<R_c} = 0 = \left[-\frac{1}{(2i+2)^2}\left(\partial_z^5 + k^2\partial^3\right)\tilde{E}_{z,2i}(z)\right]_{z=0,r<R_c}, \tag{2.52}$$

and hence that

$$\left(\partial_z^5 \tilde{E}_{z,2i}(z)\right)_{z=0,r<R_c} = 0. \tag{2.53}$$

Proceeding in this fashion, it is clear that

$$\left(\partial_z^{2n+1} \tilde{E}_{z,0}(z)\right)_{z=0,r<R_c} = 0 \qquad n = 0, 1, \ldots \infty, \tag{2.54}$$

and hence that $\tilde{E}_{z,2i}(z)$ is an even function of z. Therefore, one can easily supply `rfgun` with the necessary information by taking $E_{z,0}(-z) = E_{z,0}(z)$.

For the end of the gun ($z = 3\lambda/4$), the boundary conditions are ambiguous, as I discussed above. Since the field amplitude is small here and the beam is at full energy by the time it is in this region, what one does in at this boundary is of little importance. This being so, it is reasonable to simply extend the beam tube beyond the nominal end of the gun in order to obtain information necessary for taking derivatives up to the end of the gun. The primary benefit of doing this is aesthetic, and I have not done the equivalent in MASK, since this would force me to use a coarser mesh in the part of the gun that is really important.

2.3 Simulation Results and Predictions

Having outlined the capabilities of the codes and the methodology of their use, I now present results of my simulations of the SSRL RF gun. In addition to giving predictions of how the actual gun is expected to perform, I also report on simulations done to evaluate design alternatives, to check the assumptions of my methods, to compare the predictions of `rfgun` and MASK, and to evaluate the importance of different effects (e.g., space-charge, non-linear fields) on the beam. Experimental results and comparison of these with simulations will be presented in a Chapter 5.

2.3.1 Effects of the Cell Field Ratio

I will discuss `rfgun` results first, and in particular the effect of varying $\alpha = E_{p2}/E_{p1}$. I have done a series of `rfgun` simulations for a range of α values from 1 to 4. For each series, I varied E_{p2} in 20 MV/m steps from 40 to 120 MV/m (the latter being close to the break-down limit in the second cell).

Longitudinal Phase-Space

Figure 2.10 shows longitudinal phase space results for $\alpha = 1, 2, 3$, and 4. For each value of α, a series of curves appears, one for each value of E_{p2}. Each curve represents the momentum ($p = \beta\gamma$) for macro-electrons as a function of the time of exit from the gun (i.e., the time at which the plane $z = 3\lambda/4$ is crossed). $t = 0$ corresponds to the emission of the first particle from the cathode, which occurs just as the field enters the accelerating phase in the first cell. For these simulations, I directed `rfgun` to emit macro-electrons at 2° phase intervals, the individual macro-electrons being represented by the points on the graphs. The macro-electrons where emitted from r=0, and hence no radial motion was involved.

Note that for $\alpha = 1$, $p(t)$ is non-monotonic over most of the range of E_{p2}, exhibiting a distinctly sinusoidal shape for higher values of E_{p2}. This sinusoidal shape results from the arrival of particles in the second cell ahead of the accelerating crest in that cell. Obviously, if the first particles arrive ahead of the crest, then those that follow will arrive nearer to the crest, and thus gain more momentum. (This issue is

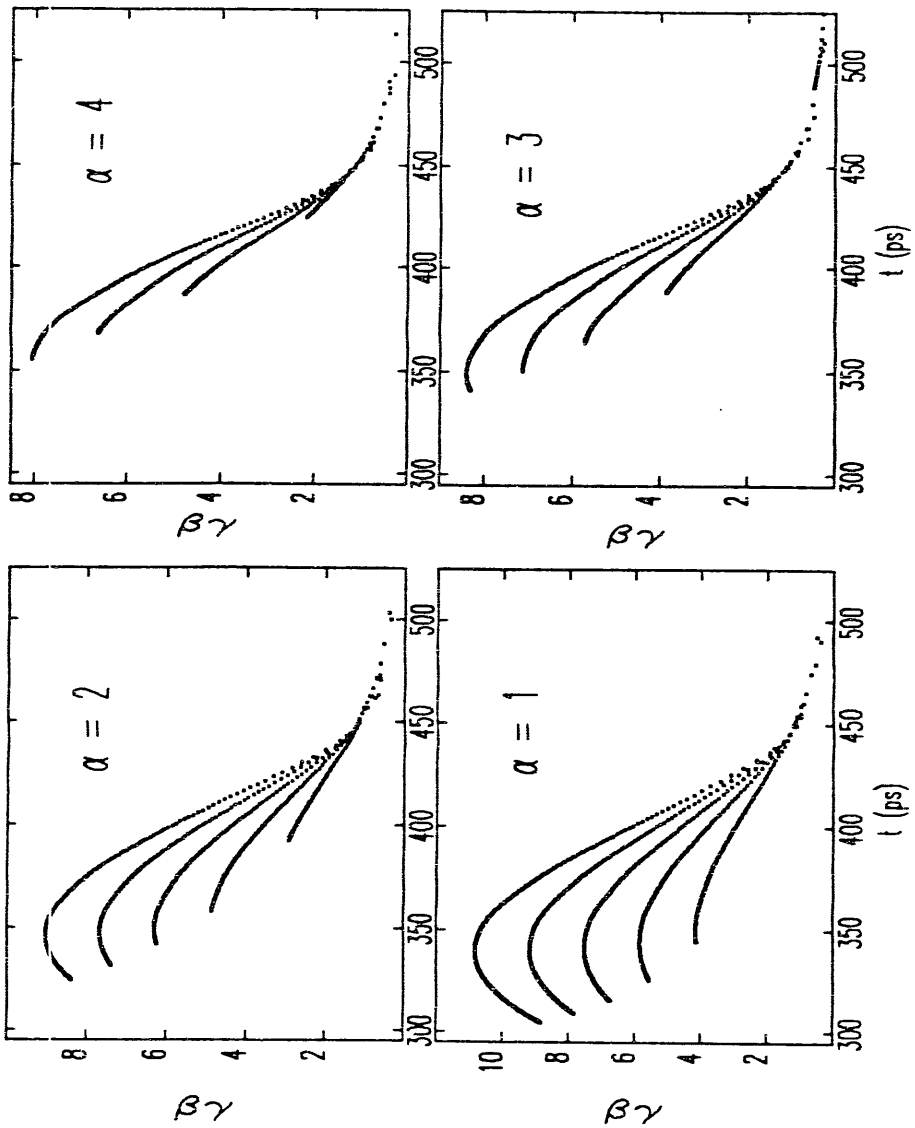

Figure 2.10: Gun Longitudinal Phase Space for Various Values of α

discussed in Chapter 4, as is the need for a monotonic p(t) curve in order to allow the use magnetic of bunch compression.) As α is increased, p(t) becomes monotonic for increasingly higher values of E_{p2}, and also for increasingly high momentum levels. For $\alpha = 4$, p(t) is monotonic for the full range of $E_{p2} > 60$MV/m, but for $E_{p2} = 40$ the fields in the first cell are too weak to deliver any significant beam to the second cell: hence, few particles exit the gun. For $\alpha = 3$, which is approximately what was achieved in the gun as built, p(t) is reasonably monotonic over $60 < E_{p2} < 100$.

Effective Cathode Area

Figure 2.11 shows the dependence on α and E_{p2} of several beam properties. Each set of connected points corresponds to the value of α indicated by the plotting symbol. The effective cathode area A_{eff} is a measure of how efficiently charge is extracted from the gun. It is defined by

$$A_{eff} = \frac{I}{J} = \frac{Q}{T_o J}, \tag{2.55}$$

where Q is the total charge exiting the gun during one RF period (i.e., the total charge "in the beam"), T_o is the RF period, and J is the current density. As one might expect, the effective cathode area increases as E_{p2} is increased and as α is decreased. This is due to the increase in the field in the first cell that accompanies both of these changes, which results in more rapid acceleration of charge from the cathode and hence more efficient extraction of beam from the gun. Recall that the physical area of the cathode is $A_c = 0.28\text{cm}^2$. One expects $A_{eff} < \frac{1}{2}A_c$, since the RF field in the first cell is only in the accelerating phase half the time. A further decrease in A_{eff} results from the fact that not all particles that leave the cathode make it out of the first cell. Many that do not exit the first cell before the RF goes into the decelerating phase will not exit the first cell at all, their momentum being insufficient to overcome the decelerating fields. These particles return to hit the cathode.

Figure 2.11 also shows the normalized beam power and the average momentum in the beam. As expected, both of these increase with increasing E_{p2} and decreasing α (i.e., increasing E_{p1}).

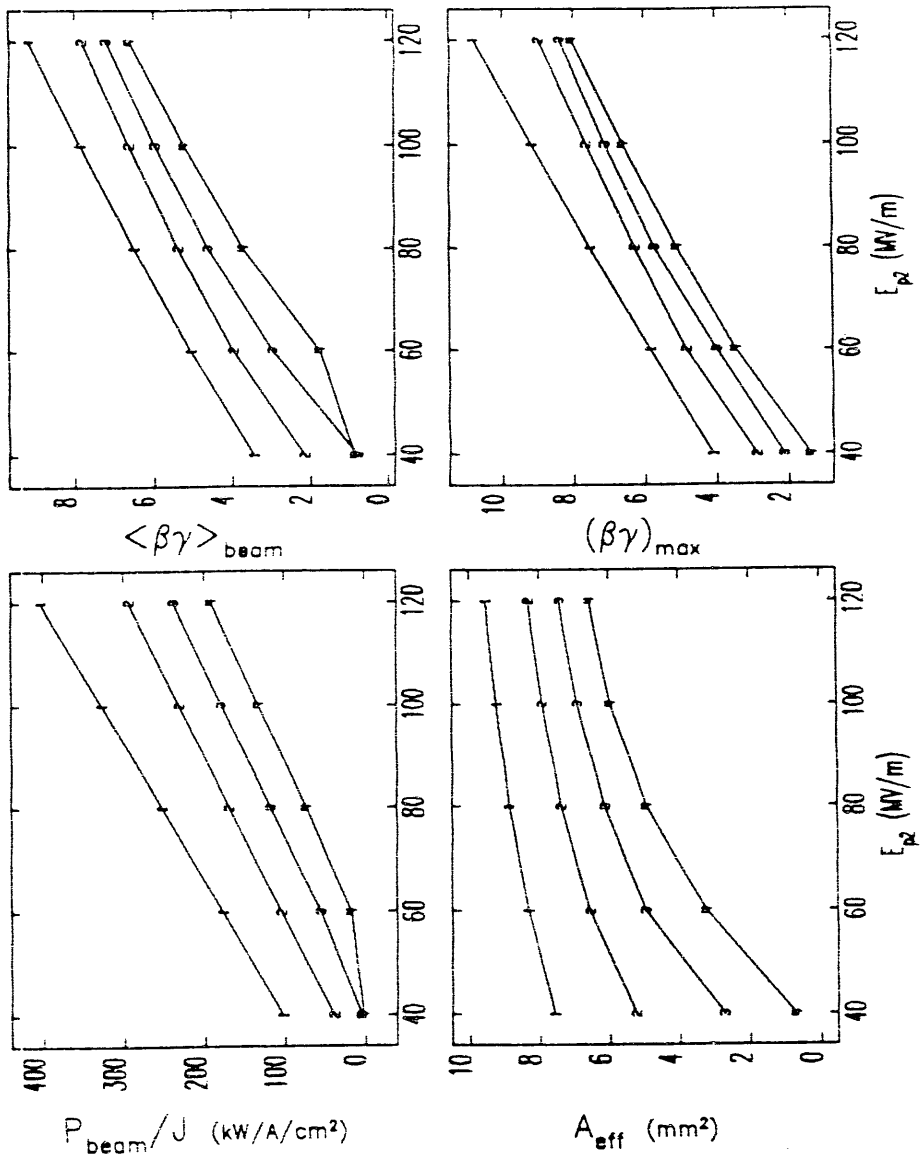

Figure 2.11: Dependence of Beam Properties on α and E_{p2}

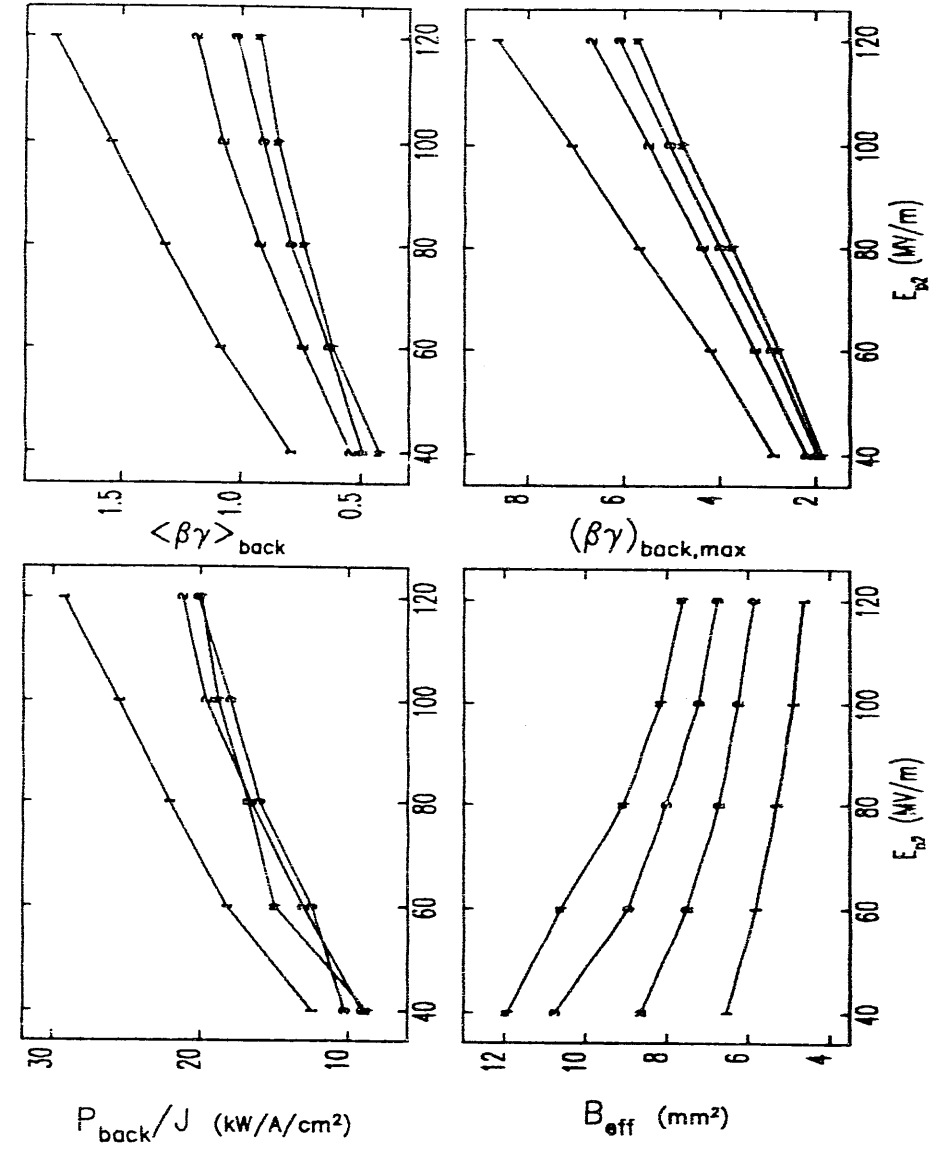

Figure 2.12: Dependence of Back-Bombardment Properties on α and E_{p2}

Back-Bombardment Power

Figure 2.12 shows corresponding quantities for the particles that return to hit the cathode. B_{eff} is the effective back-bombarder area, defined by

$$B_{eff} = \frac{Q_{back}}{T_o J}. \tag{2.56}$$

where Q_{back} is the charge that returns to hit the cathode during one RF period. One expects that $A_{eff} - B_{eff}$ will be constant, and this is found to be the case by comparison with the last figure (slight discrepancies are due to particles that do not exit the gun or hit the cathode within the total time interval of the simulation).

The Figure also shows the back-bombardment power normalized to the cathode current density. Paradoxically, there is a strong dependence on E_{p2}, but no clear trend with α.

It is actually unclear what one should expect to see here, since increasing E_{p2} or decreasing α is expected to increase the efficiency of charge extraction (as the graphs of A_{eff} confirm), thus decreasing the number of back-bombarding electrons (as the graphs of B_{eff} confirm). However, increasing E_{p2} or decreasing α is also expected to increase the momentum of any electrons that do return to hit the cathode. This expectation is confirmed by Figure 2.12. What one sees is that for constant E_{p2} and $\alpha \geq 2$, the increase in the amount of back-bombarding charge is compensated by the decrease in the average kinetic energy carried by each particle. As E_{p2} is changed, a different effect comes into play, namely the change in the back-bombardment power due to highly energetic electrons returning from the second cell. This accounts for the difference between increasing E_{p2} and decreasing α. The Figure shows the maximum momentum of any back-bombarding electron as a function of α and E_{p2}, confirming this analysis.

2.3.2 Effects of the RF Frequency

Next, I investigate the effect of changing the RF frequency while keeping the field profile the same, choosing $\alpha = 2.9$ for this and all subsequent `rfgun` studies. Note that it is sensible to imagine changing the RF frequency while keeping the on-axis fields the same, since this can be done by means of a tuning device far from the axis,

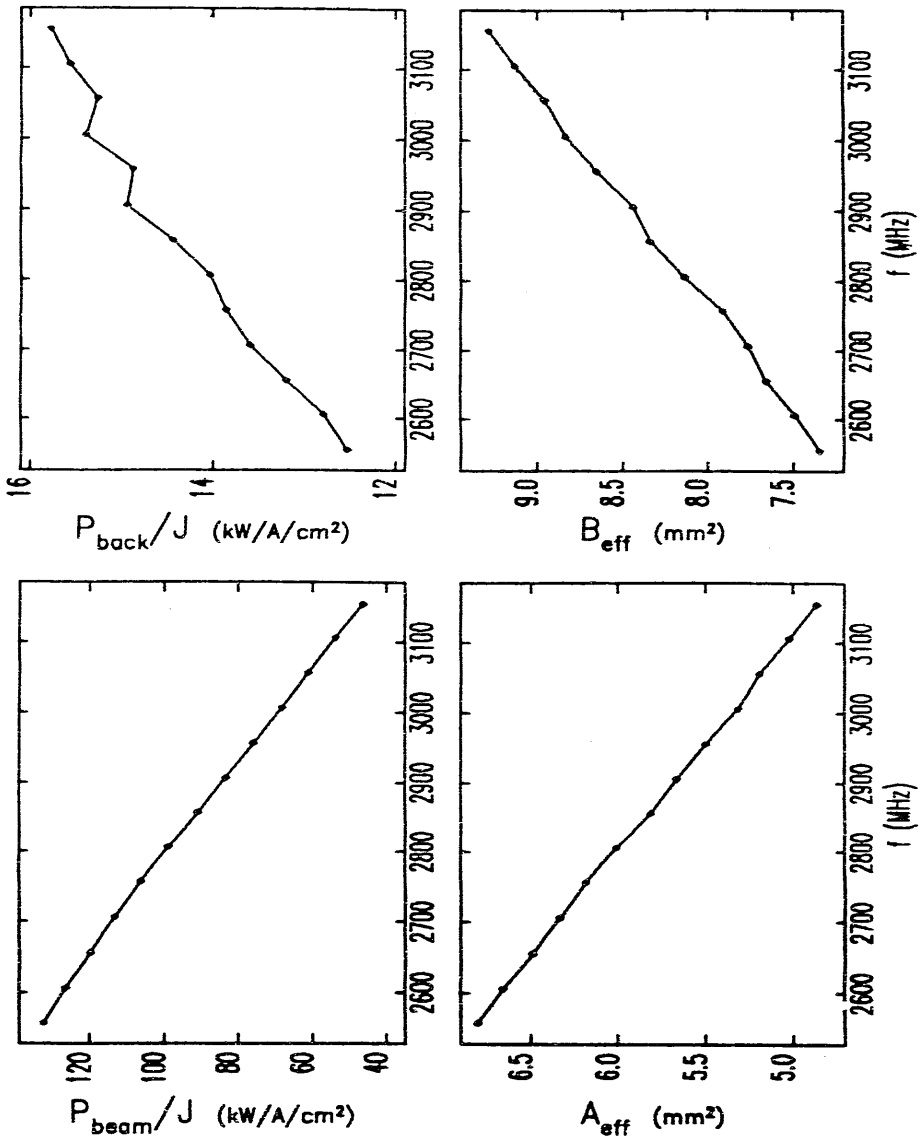

Figure 2.13: Dependence of Beam Properties on RF Frequency

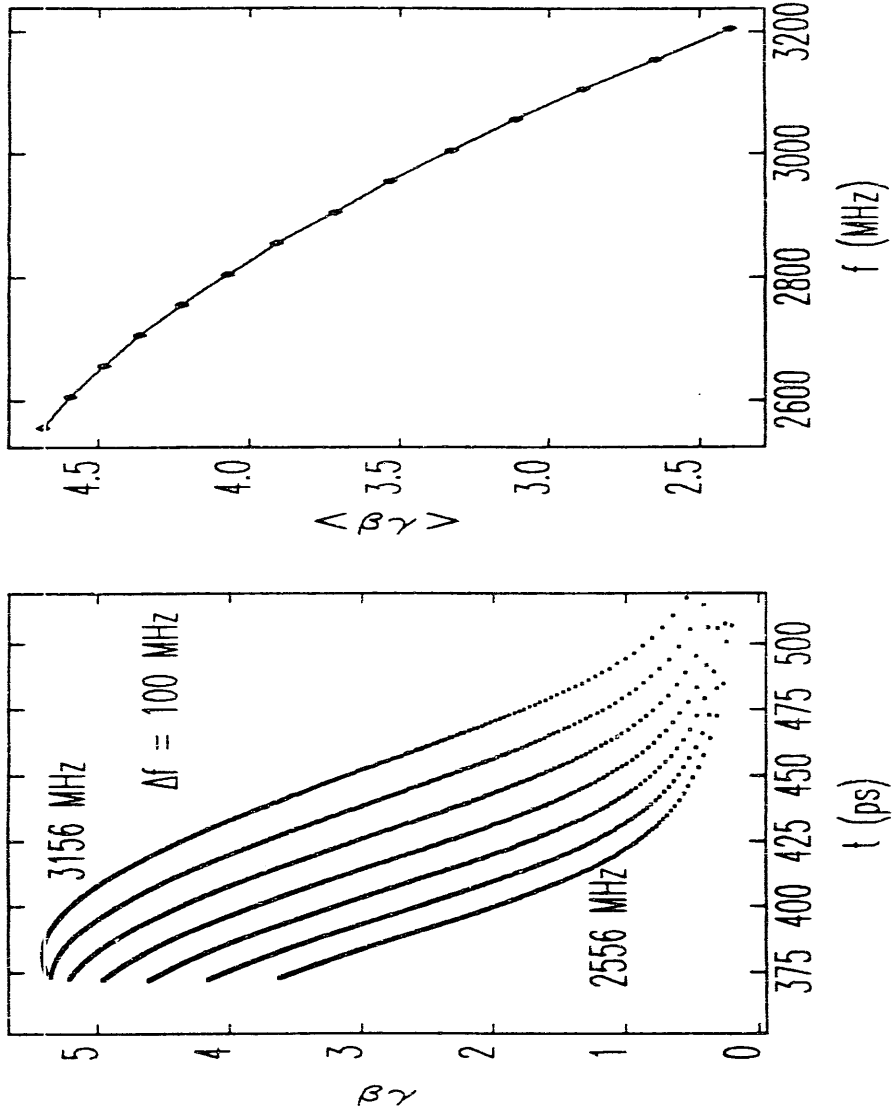

Figure 2.14: Dependence of Longitudinal Phase-Space on RF Frequency

as I did for the MASK simulations of the first cell. Figure 2.14 shows the longitudinal phase-space as a function of the RF frequency, while Figure 2.13 shows other beam properties as a function of frequency. Higher frequencies are equivalent in some ways to having a phase-shift of less than $\pi/2$ between the first and second cells, and hence the macro-electrons are seen to arrive nearer to the RF crest in the second cell. This is why the longitudinal phase-space exhibits an increasingly sinusoidal shape as the frequency is increased. For lower frequencies, the macro-electrons arrive further behind the crest in the second cell; hence, the total momentum gain is less, and the monotonic p(t) curve is retained. Lower frequencies are equivalent to having a lower-than-velocity-of-light structure.

For assessing the effect of frequency errors in MASK, it is only the region around 2856 MHz that is of interest. One sees from the Figure that the phase-space curves for $\pm$100MHz around 2856 MHz are not greatly different from those for 2856 MHz. That is, there is no dramatic effect on the slope or curvature. I find that the change in the average momentum between 2856 and 2836 MHz is 1.7%, while the change in the maximum momentum is 1.0%. The change in the total charge is similarly small, being 2.0%. Hence, the errors made in using 2836 as the frequency in MASK are negligible. Uncertainty in the value of a and in the knowledge of the exact field distribution, plus the effects of higher-order modes induced by the beam, will contribute errors as large as those introduced by the frequency errors.

2.3.3 Effects of Non-Linear Field Terms

Next, I look into the effects of field non-linearities. In particular, Figure 2.15 shows the normalized RMS emittance and normalized transverse brightness for various initial phase intervals (explained presently), where the normalized *transverse* brightness is defined as

$$B_n^* = \frac{Q}{J\epsilon^2}, \tag{2.57}$$

and the normalized RMS emittance is as given in equation (1.9). Note that B_n^* is *not* the same as the normalized brightness, B_n defined in equation (1.21). B_n^* differs in begin normalized to the current density and in having no reference to the bunch length.

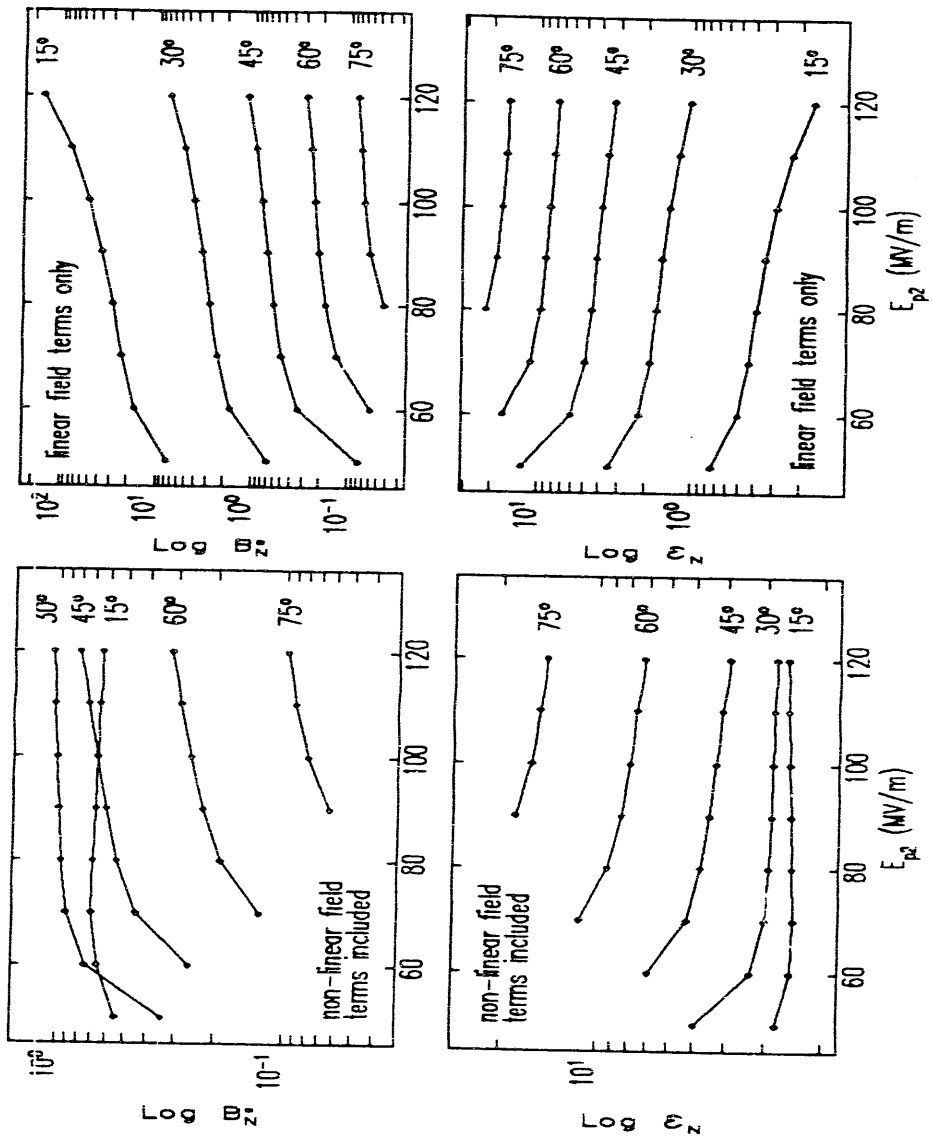

Figure 2.15: rfgun Predictions for $\alpha = 2.9$ and Various Initial Phase Intervals

The units for Q, J, and ϵ used for the graphs are pC, A/cm^2, and $\pi \cdot m_e c \cdot \mu m$, respectively. Results are shown for calculations that include non-linear RF field terms (see equations (2.35) through (2.37)), and for calculations that include only linear terms in E_r, B_θ, and E_z. In order to make a valid comparison between the linear and non-linear cases, it is necessary to ensure that one is looking at the emittance of the same particles in each case. That is, one wants to compare the emittance of particles emitted over the same range of initial phase. This has been done in the Figures, where I plot the results for the linear and non-linear cases for particles emitted during the first $\Delta\phi$ degrees of phase (measured from the beginning of the accelerating phase). As one would expect, the non-linear fields increase the emittance. Just how this occurs, and what the significance of the brightness is, will be discussed shortly.

RF Focusing

To better understand the effects of non-linear fields and time-dependent focussing forces, it is useful to look at a pair of rfgun runs in more detail. In particular, I ran rfgun at $E_{p2} = 75$MV/m with the initial particle distribution given by

$$x_{ij} = i \cdot \delta r, \quad i = 0 \ldots N_r \tag{2.58}$$

$$y_{ij} = 0 \tag{2.59}$$

$$\phi_{ij} = j \cdot \delta\phi, \quad j = 0 \ldots N_\phi, \tag{2.60}$$

where $\delta r = R_c/(N_r - 1)$ and where ϕ_{ij} is the phase of emission from the cathode, $\phi = 0$ being the beginning of the accelerating phase in the first cell. Taking such an initial distribution allows one to see with particular clarity what the effects of the initial coordinates are on the final coordinates. Figure 2.16 shows the resultant transverse phase-space for linear and non-linear fields, with $N_r = 6$, $N_\phi = 1441$, and $\delta\phi = 0.125°$. While at first glance the non-linear fields would appear to decrease the emittance, this is not so. For the linear case, macro-electrons leaving the cathode at the same phase lie along a line in x–x′ space. Hence, the emittance for such a group of macro-electrons is zero, since $\langle x'x \rangle^2 = \langle x'^2 \rangle \langle x^2 \rangle$. For the non-linear case, macro-electrons leaving the cathode at the same phase do not lie along a line in x–x′ space. This increases the emittance by decreasing $\langle x'x \rangle$. This will be seen more clearly in other data, below.

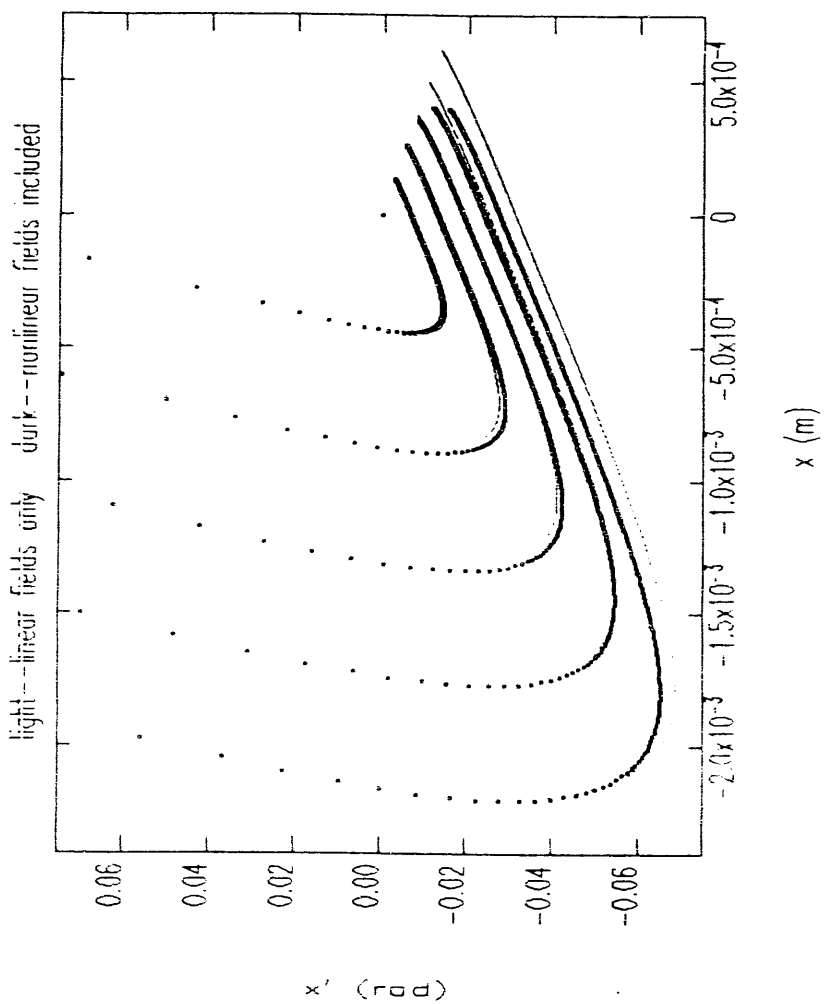

Figure 2.16: rfgun Transverse Phase-Space Results for E_{p2} = 75 MV/m and α = 2.9

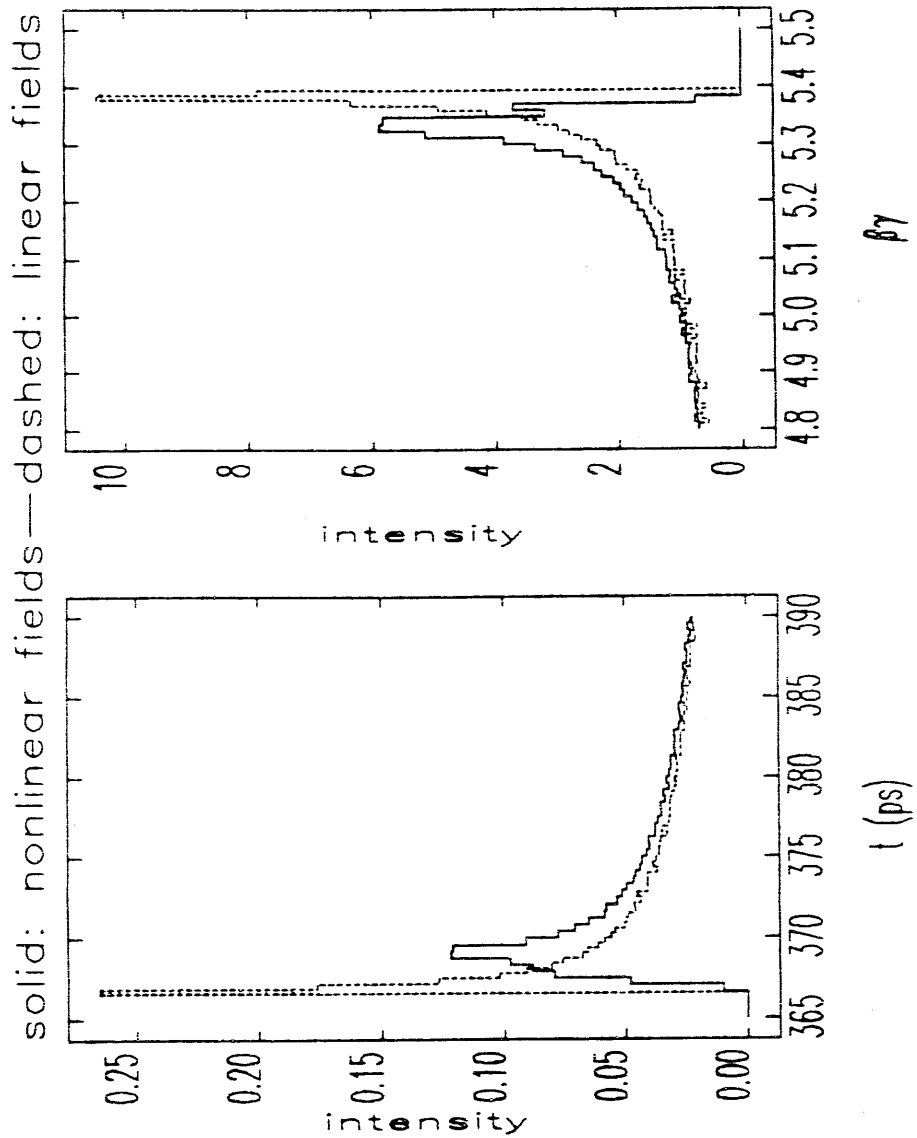

Figure 2.17: rfgun Longitudinal Phase-Space Results for E_{p2} = 75 MV/m and α = 2.9

This Figure also illustrates the effects of "RF focusing" on the emittance. The first macro-electrons emitted from the cathode are those on the right side of the figure ($x = 0$, $x' < 0$). Subsequent particles are focused differently by the time-varying (and hence initial-phase-dependent) E_r and B_θ fields. As a result, the RF sweeps the particles in a clockwise sense in the Figure. This is all consistent with Figure 2.15 where one sees that the emittance increases as a larger initial phase interval is considered, and that, for the linear case, the emittance falls increasingly rapidly as $\Delta\phi$ is decreased, whereas for the non-linear case, the emittance seems to reach a lower limit as $\Delta\phi$ is decreased. The brightness clearly saturates in the non-linear case, whereas it does not do so in the linear case. From Figure 2.16 it is also clear that curvature of the x-x' path traced by particles starting at the same radius is a result of the sinusoidal nature of the RF fields, rather than non-linearities in r.

Longitudinal Phase-Space

Figure 2.17 shows the longitudinal phase-space for the non-linear and linear cases. The non-linear fields produce a broadening of the momentum peak and of the time distribution as well, since particles starting at different radii experience different accelerating fields. In the linear case, all particles starting at the same phase receive, to first order, the same momentum and take the same time to exit the cavity. I infer from the sharpness of the momentum and time peaks for the linear case that any deviations in momentum gain or time-of-flight that result from longer path-lengths due to transverse motion are small. The momentum and exit-time distributions predicted by MASK are considerably broader than those predicted by `rfgun`, even with non-linearities included in `rfgun`. The reasons for this will be seen later in this chapter.

Transverse Phase-Space Evolution

It is instructive to look at the evolution of the transverse phase-space as the beam travels through the gun. To do this, I have run `rfgun` with an initial particle distribution defined by equations (2.59) through (2.60), with $N_r = 41$, $N_\phi = 19$, and $\delta\phi = 10°$. Figure 2.18 shows the resultant transverse phase-space at a series of z positions in first cell. The lines connect particles emitted at the same phase but different cathode

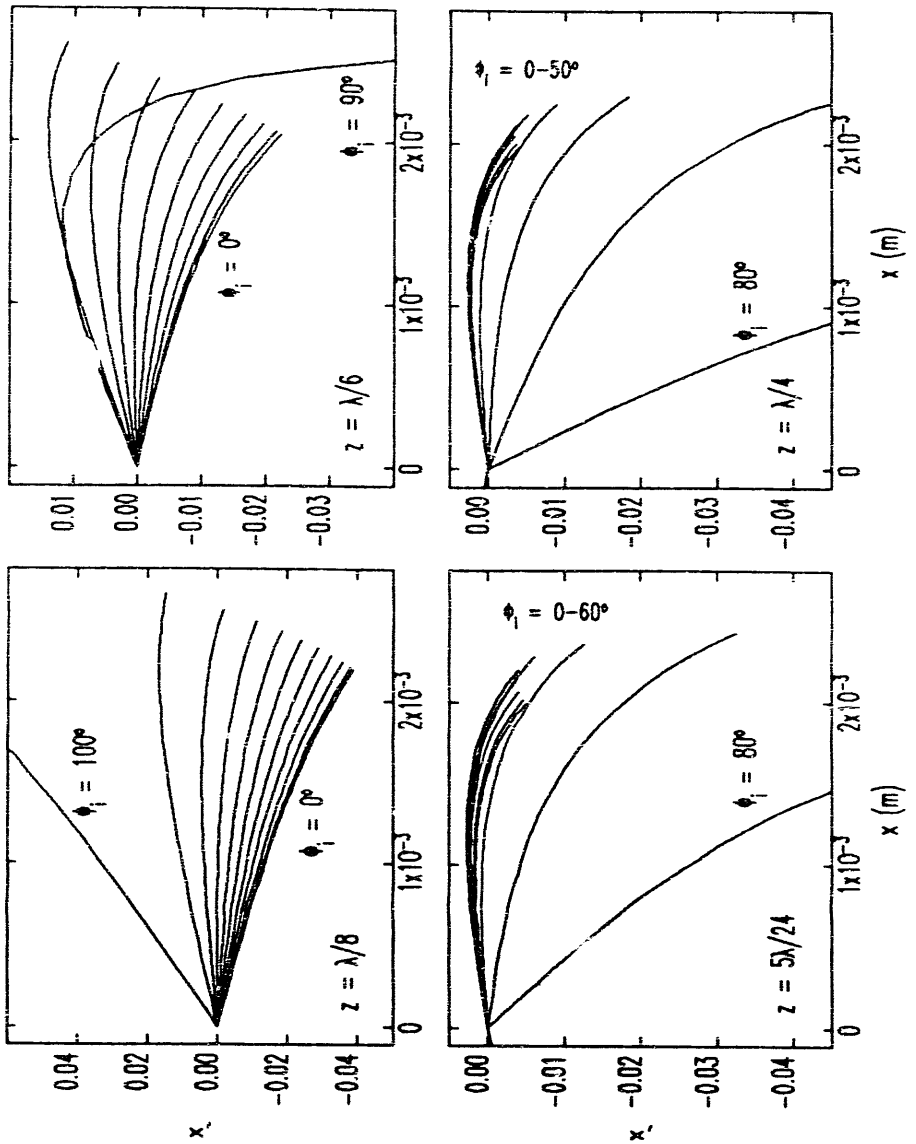

Figure 2.18: `rfgun`: Transverse Phase-Space Evolution in the First Cell, for E_{p2} = 75 MV/m and α = 2.9

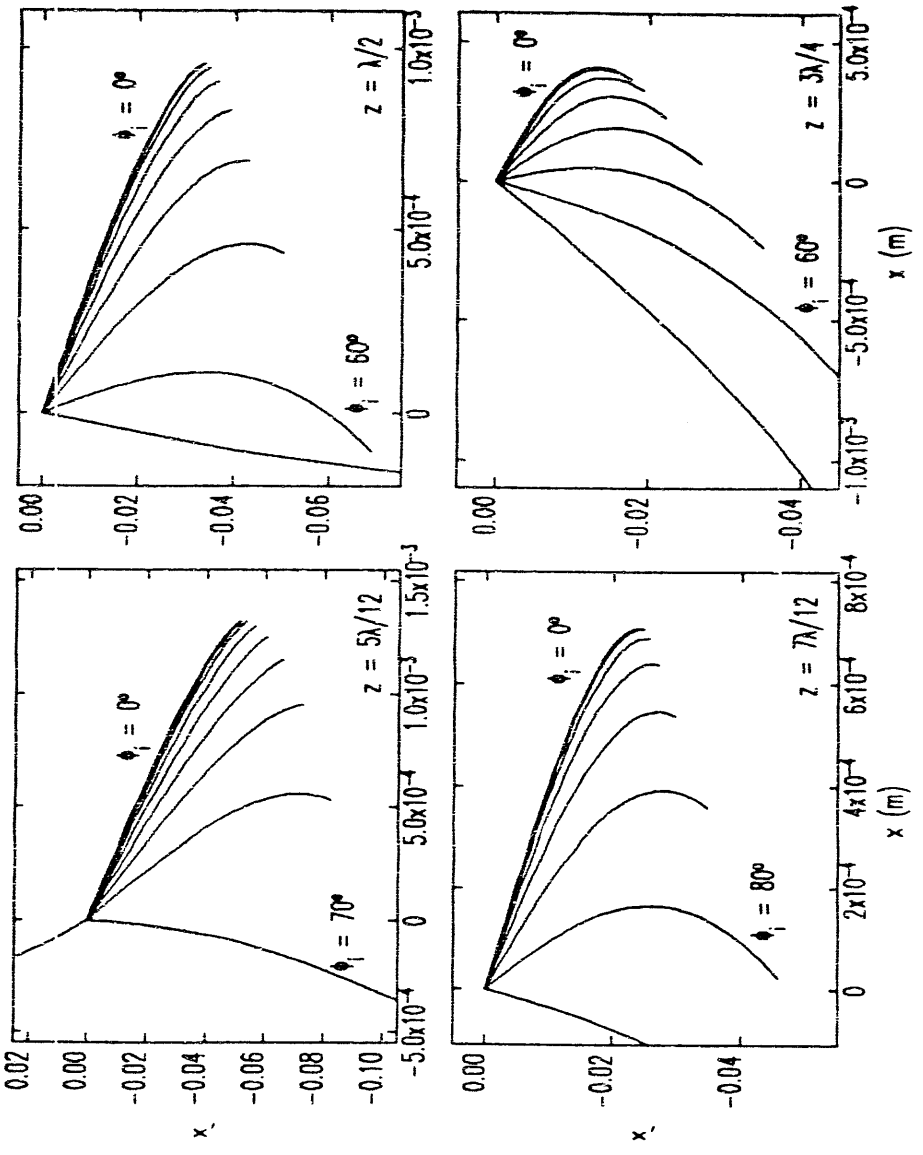

Figure 2.19: rfgun: Transverse Phase-Space Evolution in the Second Cell, for E_{p2} = 75 MV/m and α = 2.9

radii. (All of the graphs have fewer than N_ϕ lines, because particles emitted with ϕ increasingly near to 180° traveling increasing short distances in the first cell, and hence do not reach the z-planes of these phase-space plots.)

Even from the first of these plots one can see both non-linear and RF focusing effects, the later causing the fanning out of the lines, while both contribute to the curvature (I shall show below why this last point is true). As the bunch proceeds, one sees that curvature is more severe for the particles emitted at later phases, a result primarily of the large phase-spread these particles end up with. One also sees that the RF focusing is not simply fanning out the beam according to initial phase, but is also "mixing" the beam in transverse phase-space. This is simply a result of the sinusoidal variation of the focusing forces (i.e., if the variation in time were monotonic, there would be no mixing). The number of lines is seen to decrease with successive graphs, due to the slowing down and back-accelerating of the later parts of the beam.

The evolution of the beam in the second cell is shown in Figure 2.19. One sees that here is a dramatic increase in the curvature of the lines, i.e., the effect of non-linear fields, that occurs in this cell. While would appear that this change is a result of fields in the center of the second cell, this is an incorrect conclusion. What happens is that as the beam continues to converge, the effect of non-linearities from the first cell and the beginning of the second cell becomes much more evident. As one sees from Figure 2.9, the non-linear fields in the center of the second cell are very small.

Relative Importance of Different Non-Linear Field Components

I stated earlier that the curvature of the lines is due only partly to non-linearities, and partly to the time-variation of the fields. The reason is that the non-linear E_z terms cause particles starting at the same phase but at different radii to be accelerated at different rates, and thus to go through the gun at different phases relative to the RF. rfgun allows one to selectively "turn off" the non-linear E_z, E_r, and/or B_ϕ fields. By turning off the non-linear E_z terms in RF gun, I have verified that this is significant effect.

Figure 2.20 shows the effect of turning off each of E_r, E_z, and B_ϕ in turn. Each graph has both the results for all non-linear fields (in the lighter pen) along with the results with one non-linear field eliminated. Non-linear E_z terms have a dominant

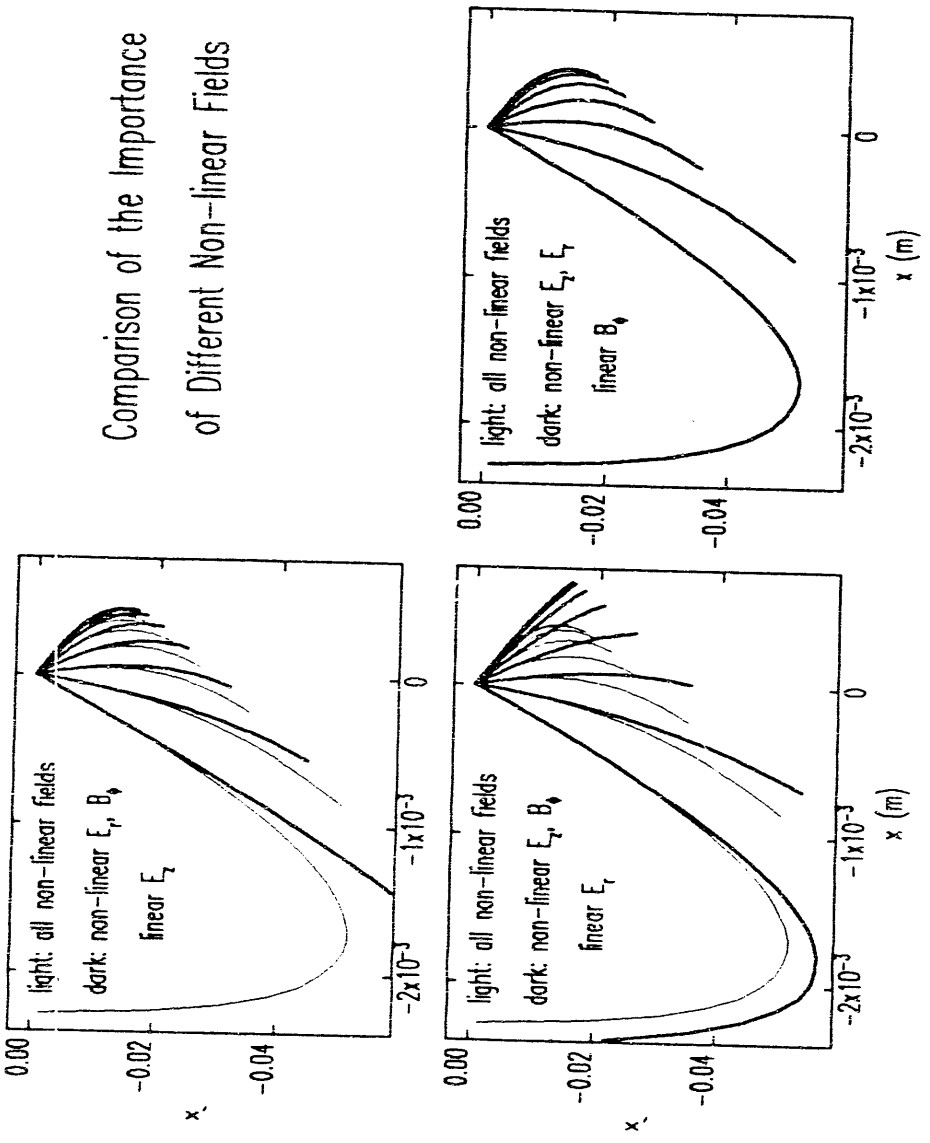

Figure 2.20: rfgun: Effects of Different Non-Linear Fields

effect on later particles, while E_r is the dominant source of curvature for the particles emitted closer to $\phi = 0$. This is plausible, since the velocity spread for the later particles (which have lower momenta) will be larger, since they are less relativistic. The effect of nonlinear B_ϕ terms is seen to be insignificant (the differences caused by removing non-linear B_ϕ terms cannot be seen on the graph, though there are differences). This is also plausible, since B_ϕ is 90° out of phase with the accelerating field (i.e., E_z), while the particles themselves are largely in phase with the accelerating field; in addition, one sees from Figure 2.9 that the non-linear B_ϕ terms are smaller compared to the linear B_ϕ terms than are the non-linear E_r terms compared to the linear E_r terms.

2.3.4 Effects of Parameter Errors in MASK Runs

Next, I look at the effect of using $\alpha = 3$ and f = 2836MHz in MASK. I use rfgun to do this evaluation, since it is faster (and since I don't have simulation cells tuned for 2836MHz for use in MASK). In particular, Figure 2.21 shows phase-space plots for $z = \lambda/4$ and $z = 3\lambda/4$ for $\alpha = 3.0$ and f = 2856MHz, and for $\alpha = 3.0$ and f = 2836MHz, compared to the results for $\alpha = 2.9$ and f = 2856MHz. One sees that while there are effects, they are confined to the particles that come later in the beam—i.e., the highest momentum particles seem least effected. The principle effect is a rotation in phase-space. The curvature of the lines is not noticeably changed.

Figure 2.22 shows the normalized emittance and normalized brightness for $\alpha = 3$ and f = 2836MHz, along with those for $\alpha = 2.9$ and f = 2856MHz. The emittance is somewhat smaller, and the brightness correspondingly larger, for the former than for the later. Thus, one expects that the MASK simulations will under-estimate the emittance by perhaps as much as 15% (though generally less) in the range $E_{p2} \geq 60$MV/m, with the error decreasing as E_{p2} increases, and decreasing as a smaller initial phase interval is taken. This difference is overwhelmingly a result of the difference in α, rather than the difference in frequency. This indicates that the difference is due to the more-rapid acceleration for $\alpha = 3$ than for $\alpha = 2.9$, which reduces the effect of non-linear E_r fields by increasing the momentum of the particles.

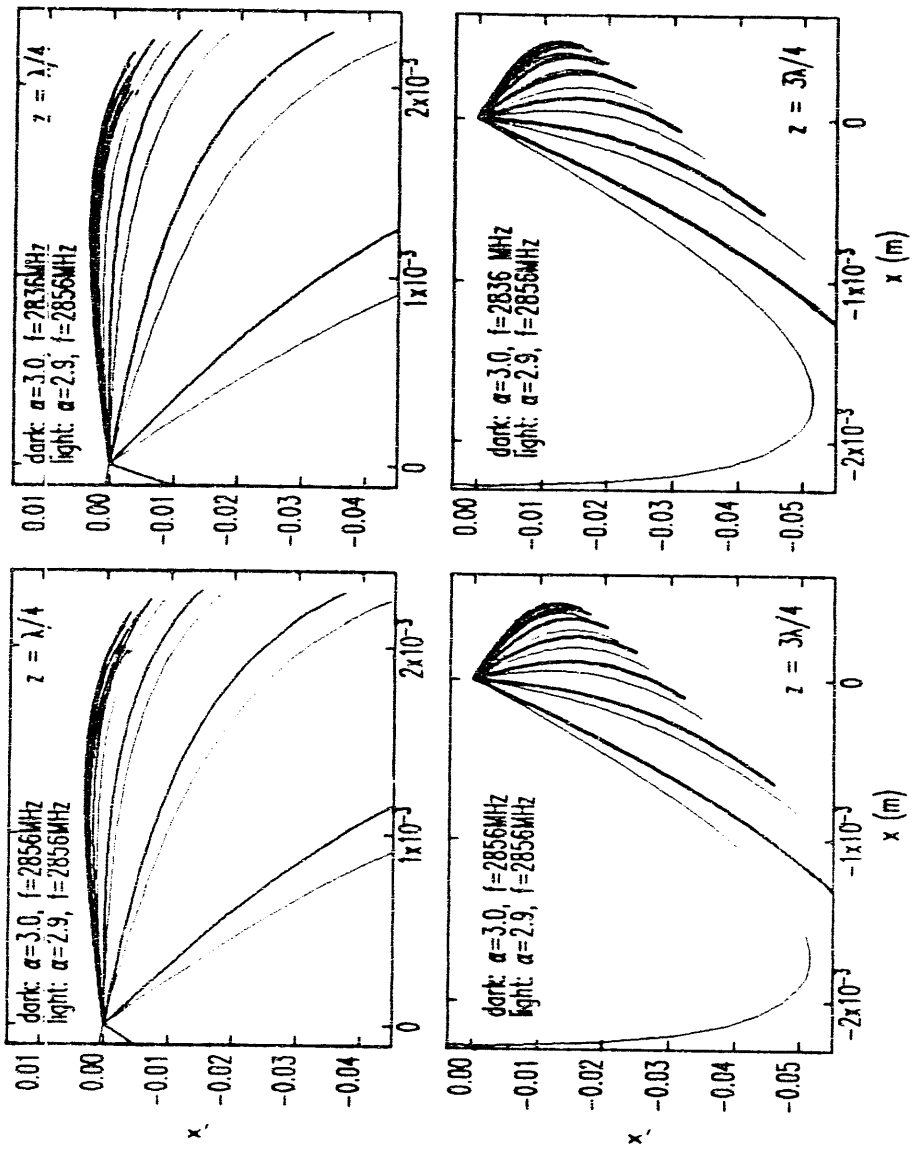

Figure 2.21: rfgun: Effects of a and Frequency Errors

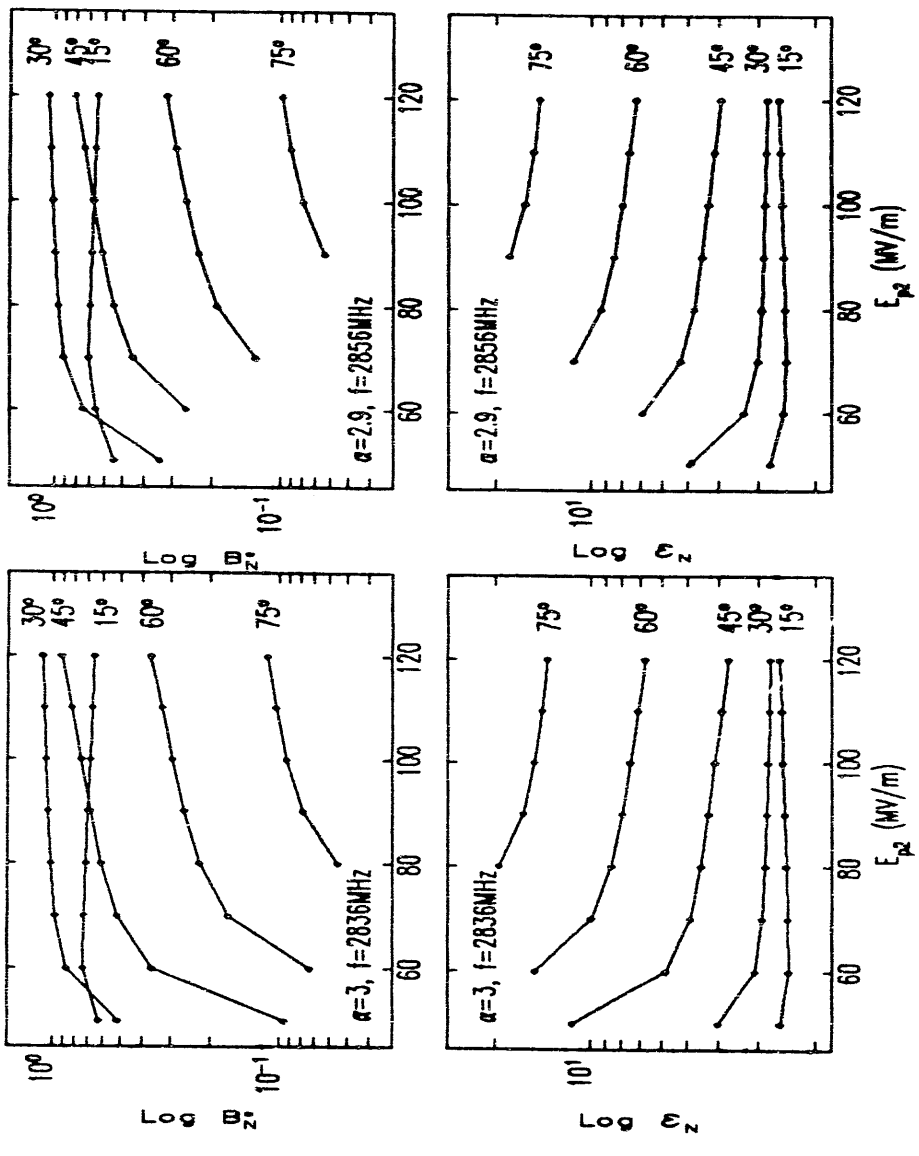

Figure 2.22: Emittance and Brightness for $\alpha = 2.9, f = 2856$ MHz, and $\alpha = 3$, f = 2836 MHz.

2.3.5 rfgun Predictions versus Momentum Spread

In the foregoing, I have looked at properties of subsets of the beam based on initial phase. This is useful, but in reality one cannot choose such subsets directly. However, since there is a high degree of correlation between initial phase, ϕ, and final momentum, one can to a large extent filter for initial phase by momentum filtration. In the above, I have chosen to work in terms of ϕ directly, because it makes the analysis more straight-forward. Because the non-linearities affect the momentum distribution as well as the transverse coordinates, the final momentum interval corresponding to a given $\Delta\phi$ for the linear case is different from the final momentum interval corresponding to the same $\Delta\phi$ for the non-linear case. Hence, if the above analysis were done with momentum filtration, the difference between the linear and non-linear cases would have been blurred.

Because MASK does not provide the initial phase of particles, it is not impossible to do the analysis based on initial phase for MASK results. Hence, for the MASK results given below, I employ momentum filtration. For comparison, I do the same for the non-linear case in **rfgun**. Figure 2.23 shows the normalized RMS emittance, normalized transverse brightness, and normalized charge per bunch, for fractional momentum intervals defined by

$$\frac{1-f}{1+f}p_{max} \leq p \leq p_{max}, \tag{2.61}$$

where p_{max} is the maximum momentum in the beam and $\pm f$ is the fractional momentum range about p_{max} $(1-f)$. The central momentum for the interval is $p_{max}/(1+f)$, and is not the same as the average momentum of the particles in the momentum interval. For typical gun operating parameters of $E_{p2} = 75$MV/m and $J = 10$A/cm^2, **rfgun** predicts a normalized RMS emittance of less than about $4\ \pi \cdot m_e c \cdot \mu$m and useful charge of as much as 100 pC, for momentum spread of less than $\pm 10\%$. Other relevant data for comparison of **rfgun** results with MASK can be gleaned from the data for $\alpha = 3$ presented in Figures 2.11 and 2.12.

The brightness is useful in comparing **rfgun** and MASK results. The merit of this quantity is that it is related to the density of particles in phase-space, rather than simply the area. It should thus be less insensitive than the emittance to momentum filtration "errors" (i.e., the inclusion of different subsets of the beam in the same

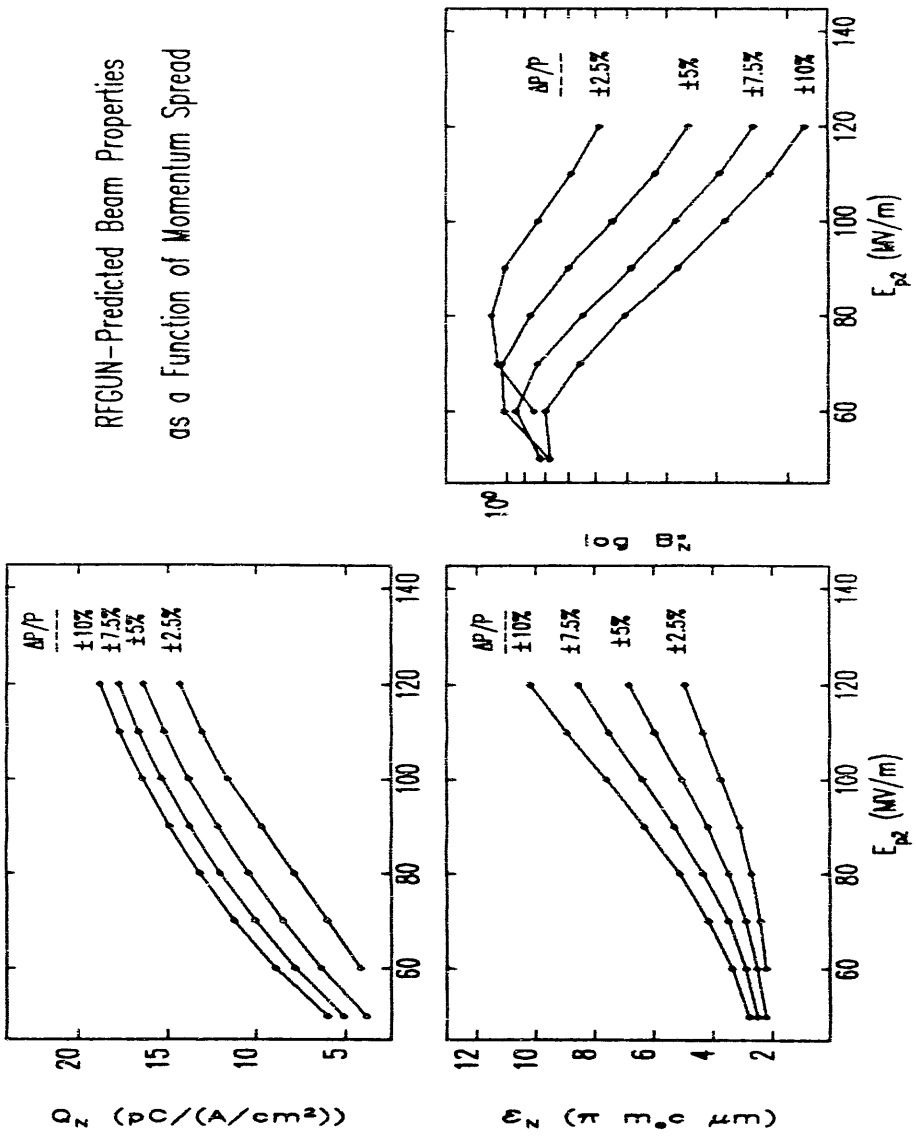

Figure 2.23: **rfgun** Predictions for $\alpha = 2.9$ and Various Final Momentum Intervals

momentum fraction f). Of course, the emittance numbers are important as well, and having them as a function of the momentum interval is also important, since in reality one is constrained by the momentum acceptance the beamline after the gun, and since chromatic effects in this beamline will worsen with larger momentum intervals. (This is discussed in Chapter 5.)

2.3.6 MASK Beam Snapshots

rfgun is able to simulate the gun with non-linear fields and with accurate longitudinal and transverse dynamics. However, it does not include any of the effects of space-charge. To look into these effects, I next discuss the results of MASK simulations. I have discussed my methodology in using MASK in the previous section. Here, I concentrate on what MASK predicts and on tests of the validity of my methodology.

Figures 2.24 and 2.25 show a series of beam snap-shots taken at various RF phases during the RF gun cycle, where a phase of 0° marks the beginning of the accelerating phase in the first cell. These were made for the nominal operating parameters of $E_{p2} = 75MV/m$ and $J = 10A/cm^2$. The graphs in Figure 2.24 have a vertical to horizontal aspect ratio of 1, while those in Figure 2.25 have an aspect ratio of 3. The dots represent individual macro-electrons, while the solid line is the actual cavity shape (which differs slightly from the mesh approximation used in the simulation, as seen in Figures 2.5 and 2.6). (In some cases, dots appear inside the "metal" of the cavity walls or outside the simulation boundary; this is because MASK dumps the particle coordinates before checking for particles that have been lost.)

The beam in these graphs appears to be hollow in part because each macro-electron represents a ring of charge and because each ring represents the same amount of charge, so that more macro-electrons are needed at larger radii in order to achieve the same current density. The desired current density is a constant, J. Hence, the number of particles inside a radius R is $\pi R^2 J$, and the number within an annulus of width ΔR about radius R is $2\pi R \Delta R J$. Hence, the number of macro-electrons within an interval ΔR about R increases linearly with R. Another reason for the hollow appearance is that non-linear focusing terms that cause an increase in the radial field magnitude with increasing r, tend to produce an increasing particle density with

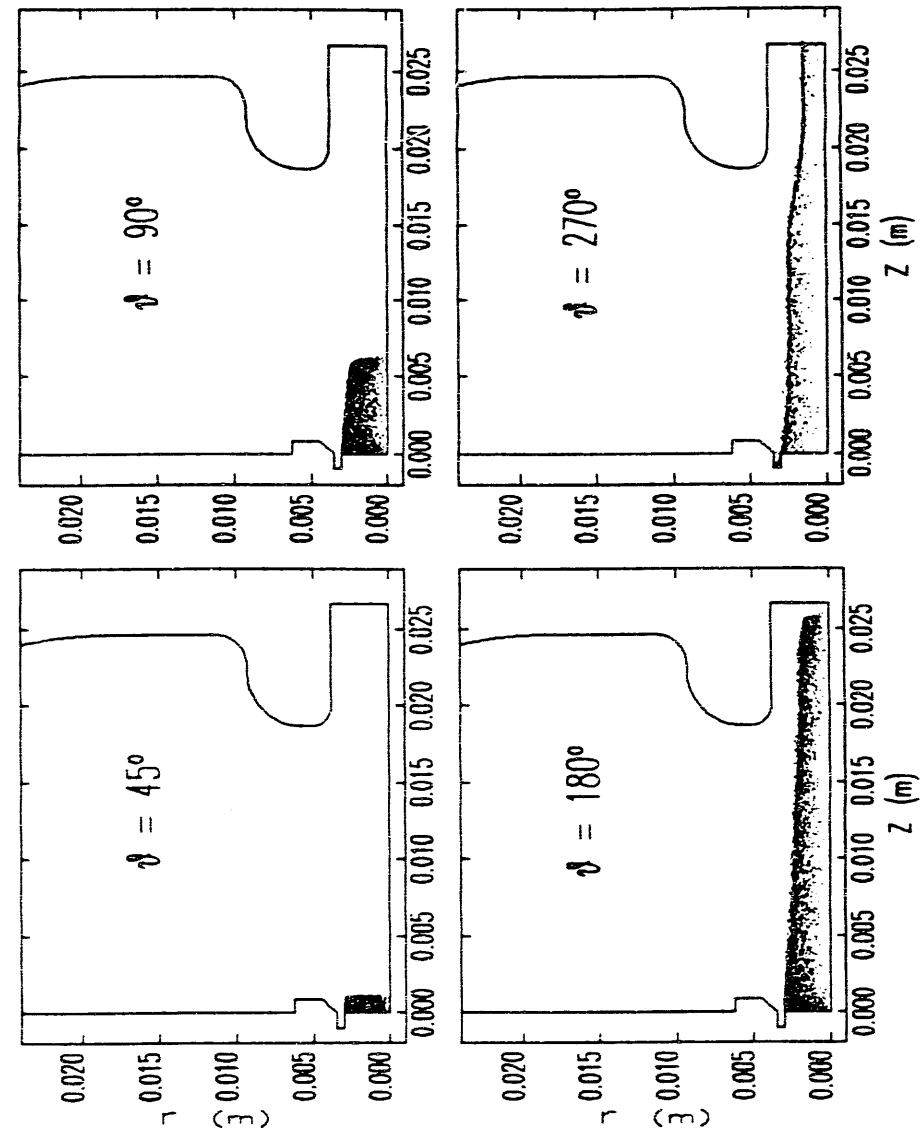

Figure 2.24: MASK Beam Snapshots at Various Phases—First Cell

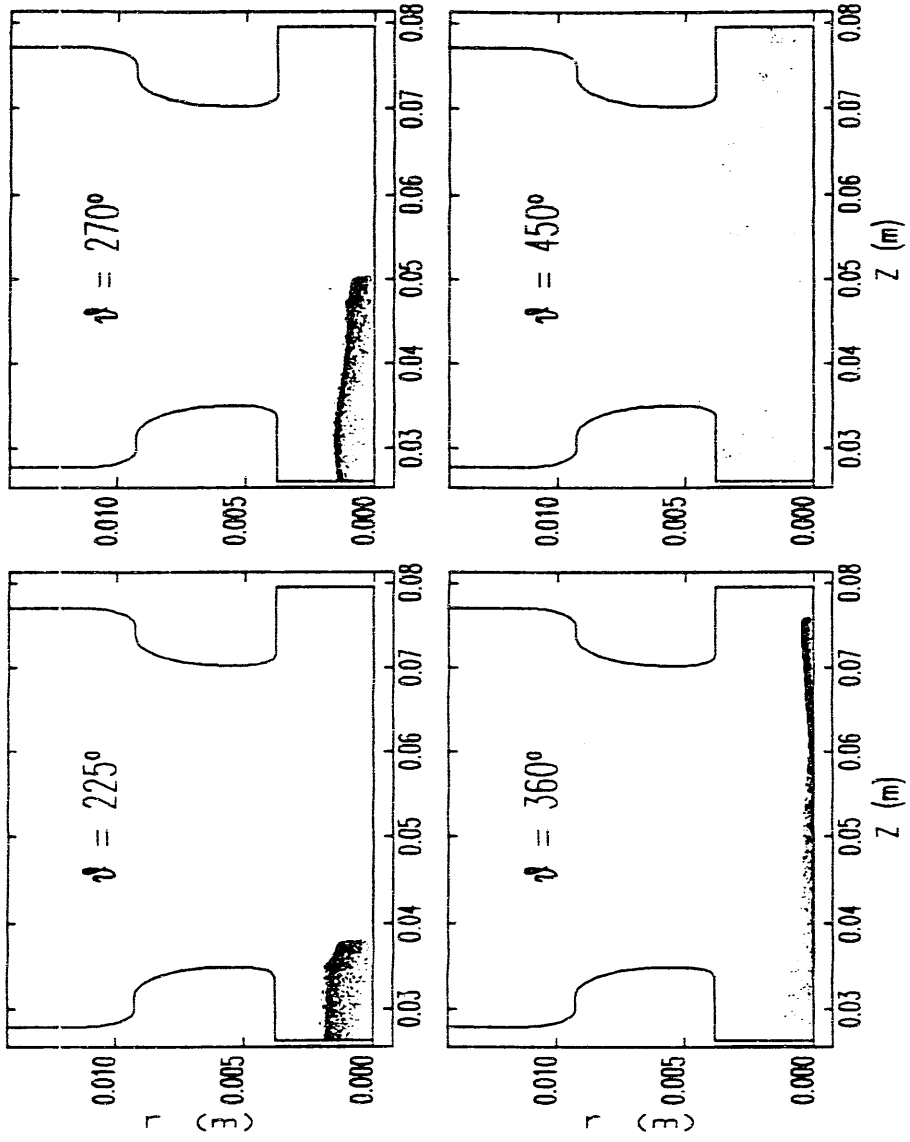

Figure 2.25: MASK Beam Snapshots at Various Phases—Second Cell

increasing r (see Sections 2.3.11 and 2.3.12).

Notice that the lead particles in the beam have not moved more than about 6mm after the first 90° of the cycle, while after another 90° of the cycle, the lead particles are about to exit the first cell, having traveled more than 25mm. This testifies to the rapid increase in velocity that accompanies the particles' trip across the first cell. Indeed, for $E_{p2} = 75\text{MV/m}$, the lead particles have $\beta \approx 0.8$ upon exiting the first cell. Note that the beam travels most of the length of the second cell in 180°, even though the second cell is twice the length of the first. After 270° of the cycle have passed, the RF in the first cell is at the decelerating crest. One sees that there are still many particles in the first cell. These particles are in fact being accelerated back into the cathode.

The relatively slow initial motion of the particles also underlines the importance of the cathode region in determining beam properties, since particles spend a disproportionate amount of time in the region of the cathode. It is this initially slow motion in a region with large non-linear fields that, for example, leads to the large effect of the non-linear terms in E_z on the momentum spread. As the beam travels through the first 5mm of the first cell, the front edge of the beam takes on a cupped shape, due to the non-uniformity of the longitudinal field across the cathode.

The Figures clearly show the effect of the focusing noses in producing a converging beam: the beam radius has decreased by a factor of about 2 by the time the beam exits the first cell. If the focusing noses were not in place around the cathode, the beam would fill the aperture of the beam tube. With the focusing noses, the beam converges and passes easily through the beam tube, even for very high current densities. The transverse beam size continues to decrease as the beam travels through the second cell, partially due to additional focusing forces encountered in passing the first cavity nose in the second cell. These are visible in Figure 2.25 by virtue of the "kink" they produce in the radial beam envelope, as seen in the graph for 270°.

2.3.7 Calculating Emittance from Cylindrical Coordinates

Since MASK (unlike `rfgun`) works in cylindrical coordinates, the dynamics of macro-electrons is calculated in terms of radial, azimuthal, and longitudinal momenta. Since

do not impose any external magnetic fields (e.g., a solenoidal field along the axis), the azimuthal momenta are identically zero. Hence, the relevant phase-space coordinates of any macro-electron are (z, r, p_z, p_r). Macro-electrons that pass through the origin still have positive r coordinates, but have p_r reversed in sign. To see how to obtain the emittance, note that the x and y coordinates are related to r by

$$x = r\cos\theta_s \quad \text{and} \quad y = r\sin\theta_s, \tag{2.62}$$

where the subscript s stands for the *spatial* coordinates. Similarly, the x and y momenta are related to p_r by

$$p_x = p_r\cos\theta_p \quad \text{and} \quad p_y = p_r\sin\theta_p, \tag{2.63}$$

where the subscript p stands for the momentum. Because there is no beam rotation, however, one must take $\theta_s = \theta_p$. The normalized emittance in the x plane is given by equation (1.9), which implies

$$\epsilon_{nx} = \pi m_e c\sqrt{\langle r^2\cos^2\theta\rangle\langle p_r^2\cos^2\theta\rangle - \langle p_r r\cos^2\theta\rangle^2}. \tag{2.64}$$

Averaging over θ, one sees that

$$\epsilon_{nx} = \frac{\pi}{2}\sqrt{\langle r^2\rangle\langle p_r^2\rangle - \langle p_r r\rangle^2}. \tag{2.65}$$

Clearly, $\epsilon_{nx} = \epsilon_{ny}$.

2.3.8 Tests of the Independent Bunch Assumption

I mentioned above that the MASK simulations are done with the implicit assumption that each bunch is independent. To test this assumption, I simulated the first cell for five cycles with beam (in addition to the cycles necessary to excite the cell). It was not possible to simulate both cells, since the frequency mismatch between the two cells in the simulation would have made the phase between the cells drift, thus obscuring the effect. In any case, one expects that the predominant effect will come in the first cell, where the energy of the beam is lowest and where more charge is present during more of the RF cycle. In order to bring out any effect, I deliberately chose a very high current density of 80A/cm^2, though 10–20 A/cm^2 is the range used

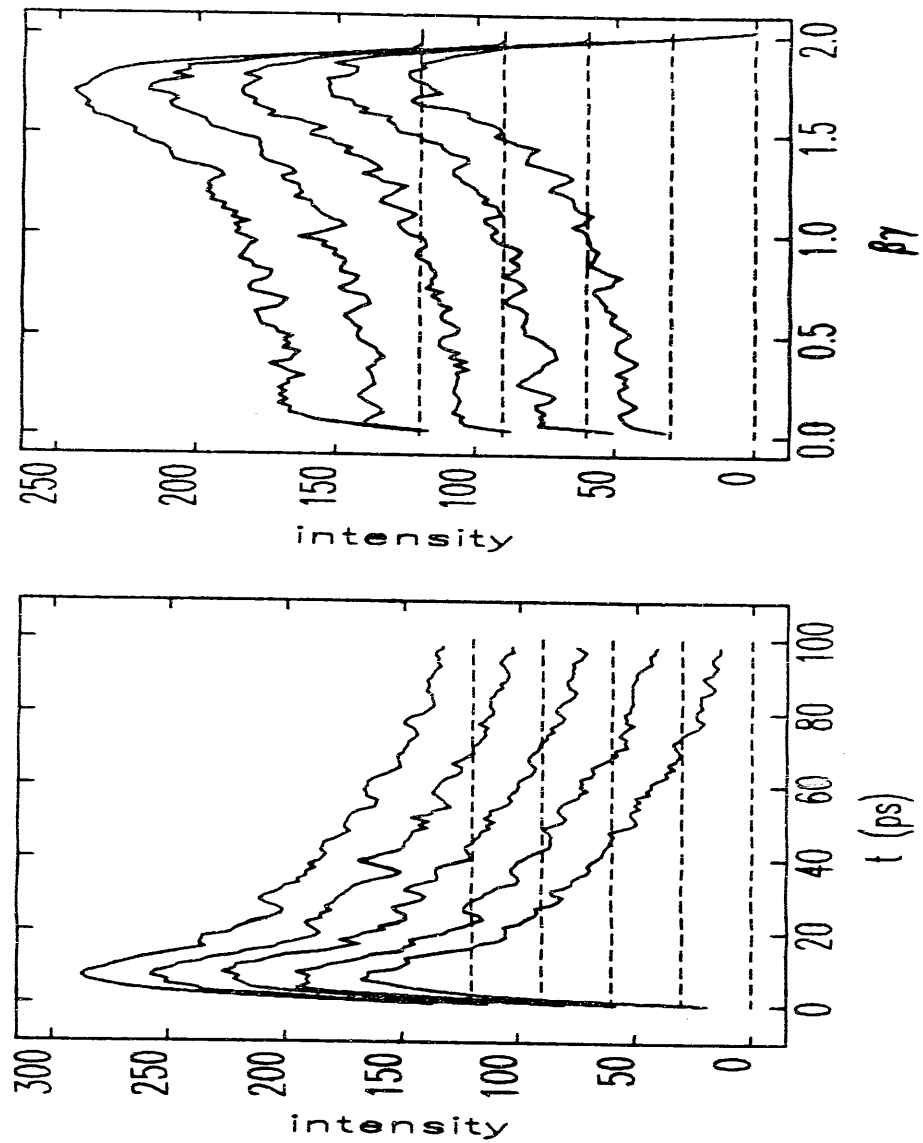

Figure 2.26: Exit-Time and Momentum Histograms as Functions of Bunch Number

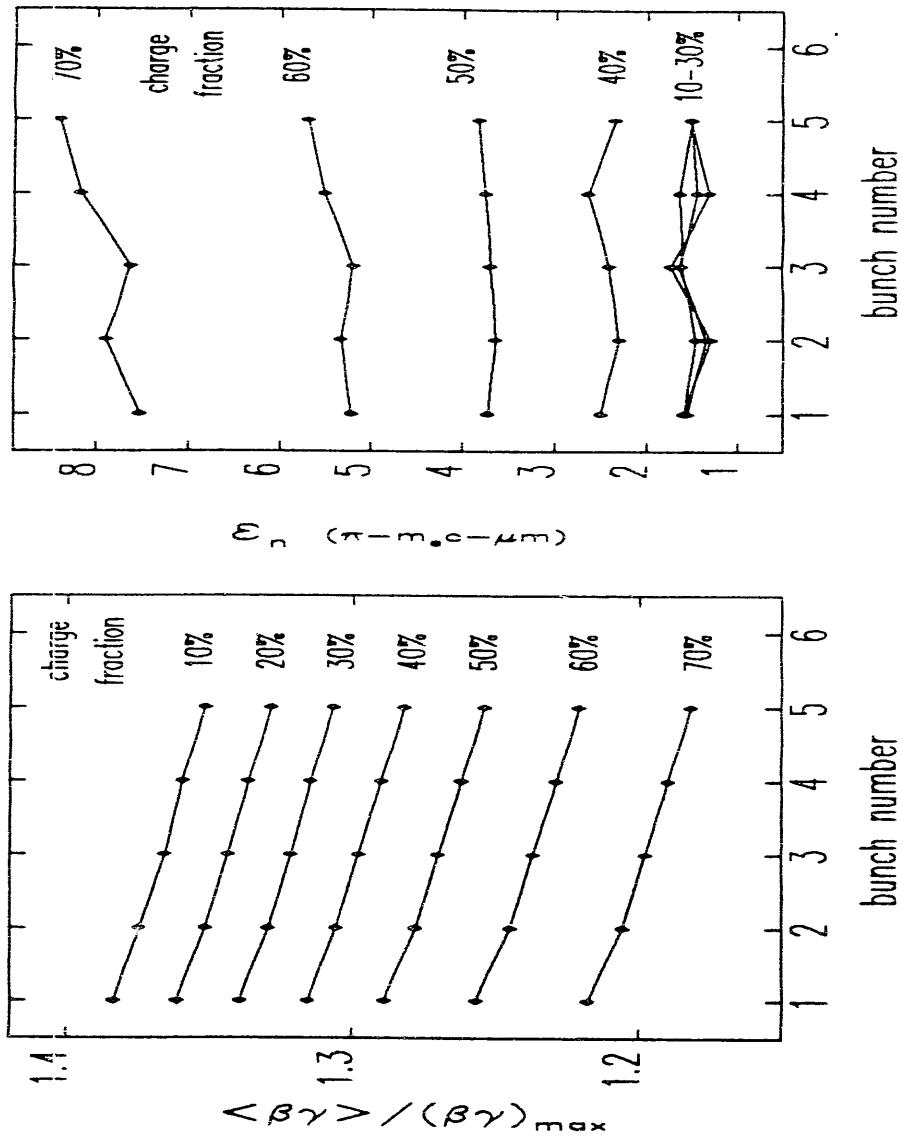

Figure 2.27: Normalized Average Momentum and Normalized Emittance as Functions of Bunch Number

in actual operation of the gun at SSRL. In Figure 2.26, I show histograms of exit-time (relative to the first particle in the bunch) and momentum for the five bunches, offset for clarity, with later bunches being offset by a larger, positive amount. No dramatic changes are evident, though the momentum peak is clearly occurring at smaller momenta for later bunches. This is a result of the extraction of energy from the cavity by previous bunches.

Figure 2.27 shows additional data. Here, I am compare the normalized average momentum, $\langle\beta\gamma\rangle/(\beta\gamma)_{max}$, and the normalized emittance for successive bunches, for various fractions of the total charge in the first bunch, starting with the most energetic particles in each bunch. This is less ambiguous than using momentum-spread intervals, since in the present case the average momentum and the momentum distribution are changing. This is roughly equivalent to choosing the same initial phase interval (ignoring longitudinal mixing caused by non-linear E_z terms.)

One sees that the effect on the normalized emittance is not dramatic, producing a spread of about 0.5 $\pi \cdot m_e c \cdot \mu m$ and no clear trend toward an increase for less than 60% of the charge in the beam. There is an trend in the normalized average momentum, but the regularity of the trend suggests that it is simply a result of the natural change in the momentum distribution as a function of field level in the cell. If the effect were due to high-order modes, one would expect it to display less regularity. Without proving this, I believe the data presented confirm the reasonableness of using MASK in the single-bunch, assumed-steady-state mode, especially for current densities significantly less than 80 A/cm^2.

2.3.9 Transverse Beam Evolution

In order to better understand the bunch evolution within the gun, I have done MASK runs with "emittance windows" at various locations inside the gun. These windows are user-defined planes of constant z, such that whenever a macro-electron passes one of the planes while traveling in a specified direction (i.e., toward positive z or negative z), MASK dumps the macro-electron's phase-space coordinates (i.e., radius, time, and radial and longitudinal momentum). Because MASK checks the z coordinate of each macro-electron against the z coordinate of each window at every time step, using

too many windows is expensive in terms of CPU time. Hence, I placed only four windows in the first cell and four in the second. In the first cell, one window was in front of the cathode, to keep track of back-bombarding macro-electrons. The other three were equispaced by $\Delta z = \lambda/12$, being placed at $z = \lambda/12, \lambda/6$, and $\lambda/4$. The last of these was used to record macro-electron coordinates for re-injection into second-cell simulations. In the second cell, the windows were similarly placed. One window was at $z = \lambda/4$, in order to keep track of macro-electrons that get back-accelerated into the first cell. The others equispaced by $\Delta z = \lambda/6$, being placed at $z = \lambda/4 - \lambda/6, \lambda/4 + \lambda/3$, and $3\lambda/4$. The last window was used to record the macro-electron coordinates at the exit of the gun.

Figure 2.28 shows the MASK-calculated transverse phase-space distributions in the first cell from a simulation with $J = 10^{-5}A/cm^2$ (i.e., essentially turning off the space-charge) and $E_{p2} = 75MV/m$. Comparison with Figure 2.18 shows that the MASK results are qualitatively similar to the `rfgun` results, but quantitatively different.

2.3.10 Accuracy of MASK Field Calculations

These differences are a result of differences in the fields calculated by `rfgun` and MASK. Recall that `rfgun` uses an off-axis expansion, starting from SUPERFISH-calculated values of $E_z(z, r = 0)$.

Figure 2.29 shows derivatives of MASK- and SUPERFISH-calculated fields for the mesh spacings given earlier in this chapter (i.e., $\Delta z = 0.43$mm and $\Delta r = 0.25$mm). SUPERFISH predicts somewhat smaller $\partial_z^2 E_z(z, r = 0)$ and $\partial_z^3 E_z(z, r = 0)$ near the cathode and in the vicinity of the cell noses.

Figure 2.30 compares the longitudinal fields calculated by MASK to those calculated by URMEL [65], showing that $E_z(z, r)$ has a increasingly large apparently linear term in z as r increases toward R_c. Thus, the fields calculated by MASK do not exactly satisfy (2.54). (This comparison could not be done between MASK and SUPERFISH, because SUPERFISH uses an adaptive, and hence irregular, triangular mesh, which makes it difficult to obtain the off-axis fields. The URMEL fields could not be used in `rfgun` because they are too noisy to permit accurate higher-order

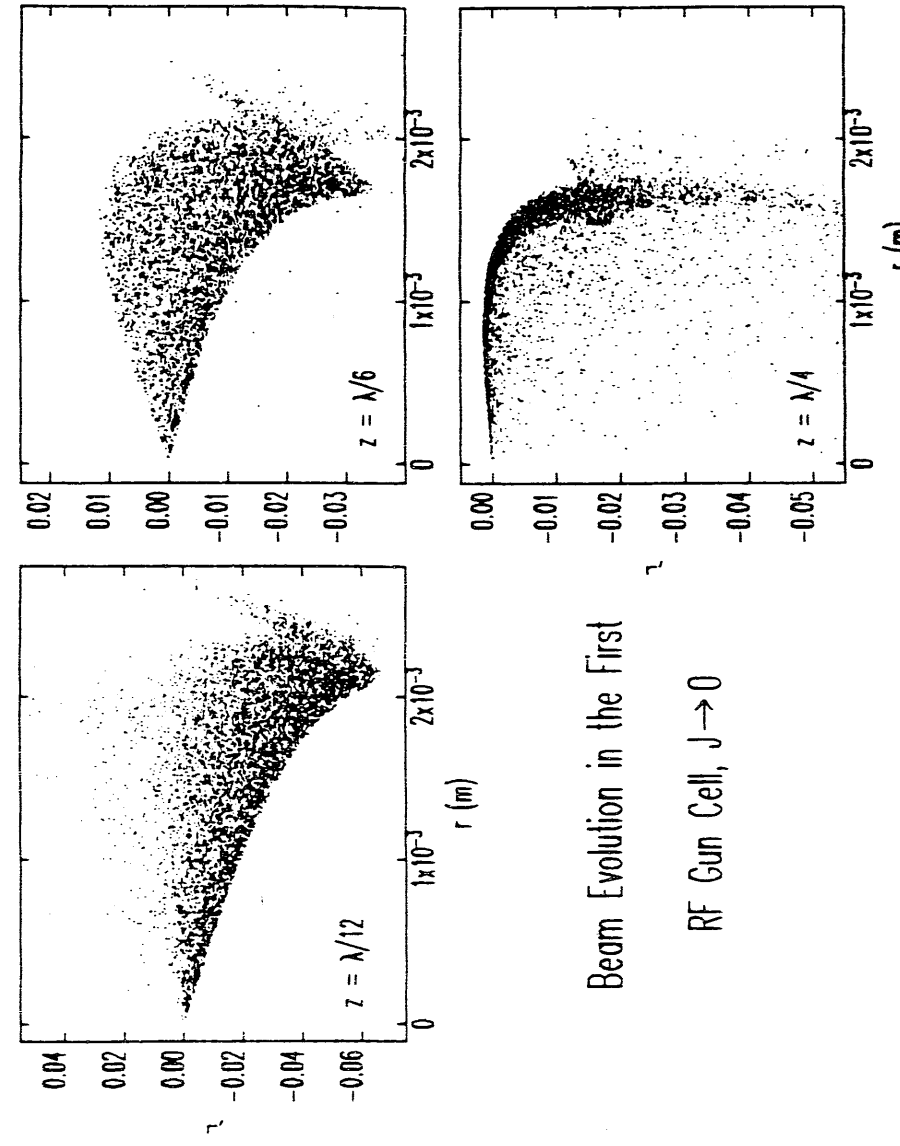

Figure 2.28: MASK Transverse Phase-Space Evolution in the First Cell, for J — 0

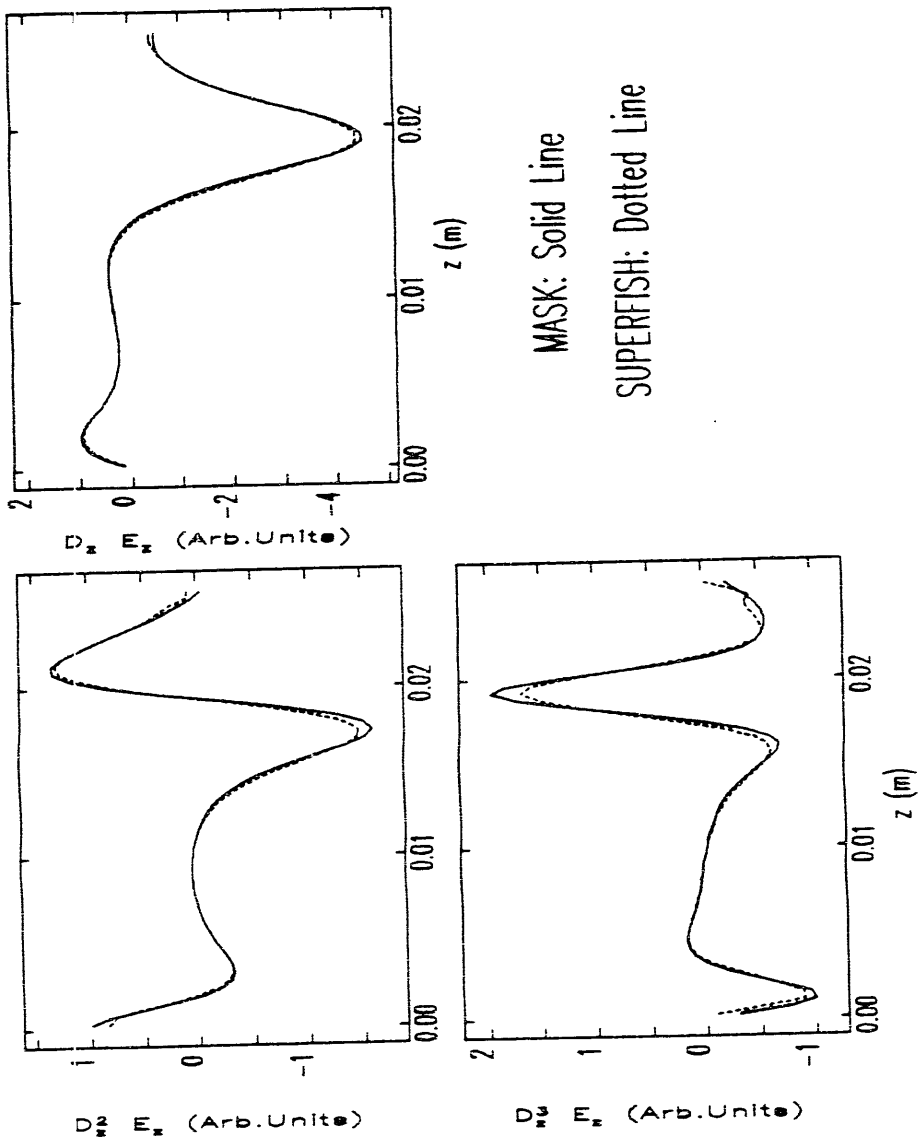

Figure 2.29: Comparision of Derivatives of On-Axis Longitudinal Fields in the First Cell as Calculated by MASK and SUPERFISH

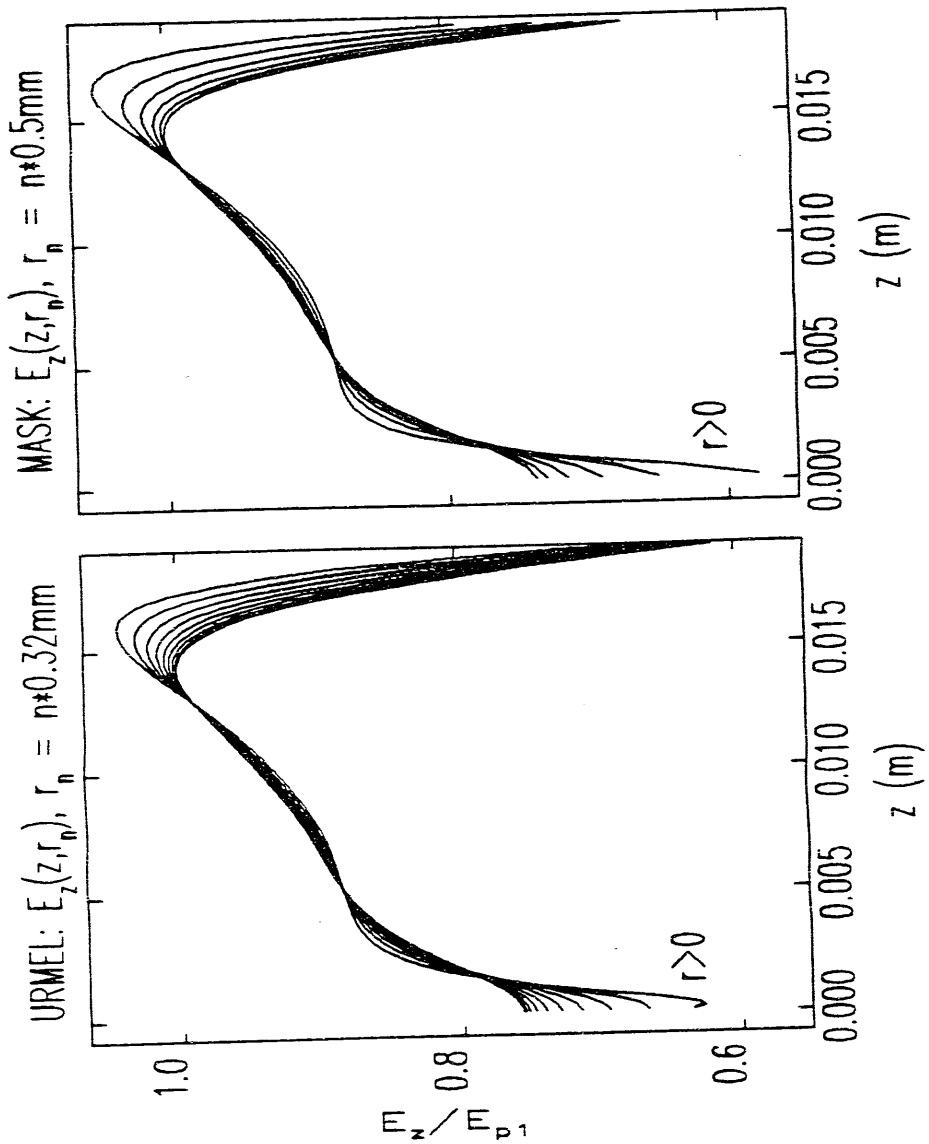

Figure 2.30: Longitudinal Fields in the First Cell as Calculated by MASK and URMEL

numerical differentiation.) If the longitudinal mesh-spacing is decreased in MASK, the MASK and URMEL results come into much better agreement, but the predicted particle motion is not greatly changed. Hence, another explanation must be sought for the differences between MASK and rfgun predictions.

2.3.11 Adequacy of Off-Axis Expansion

The real source of the discrepancies between MASK and rfgun is the inadequacy of the off-axis expansion used in rfgun, which for E_r is only third order in r. Figure 2.31 shows four different calculations of $E_r(z)$ at r = 2.87mm, which is close to the cathode radius (3mm). An explicit MASK result was obtained by running MASK with a finer mesh (Δz = 0.215mm and Δr = 0.164mm) and sampling $E_r(z)$ at r=2.87mm directly; MASK predicts larger a E_r than any of the other calculations. The other three curves are various calculations of $E_r(z)$ at the same radius using off-axis expansions. Two third-order calculations are shown, one starting with SUPERFISH-calculated on-axis longitudinal fields, $E_{z0}(z)$, the other starting with the same result from MASK. The other expansion is a fifth-order expansion starting with the SUPERFISH-calculated $E_{z0}(z)$ (the MASK data was too noisy to allow a fifth-order expansion, as the noisiness of the third-order expansion shows). Clearly, the fifth-order expansion is the closest to the explicit MASK result. Hence, the conclusion is that MASK is making more accurate predictions of the effect of non-linear fields than rfgun is making.

In order to test this diagnosis, I used the $E_z(z, r = 0)$ profile from MASK in rfgun, and repeated some of the analysis done above. Figure 2.32 shows the normalized RMS emittance and the normalized brightness for the two cases. A significant, though hardly dramatic, change in the predicted emittance is obtained when using the MASK-calculated on-axis field profile. For larger initial-phase intervals, the predicted emittance is smaller, while for smaller initial-phase intervals, it is larger. The brightness follows the opposite pattern, as expected.

Figure 2.33 shows a comparison of MASK results for J → 0 with rfgun results obtained using the MASK-calculated $E_z(z, r = 0)$. One sees that MASK predicts larger emittances than rfgun, though not dramatically larger. The reason for this

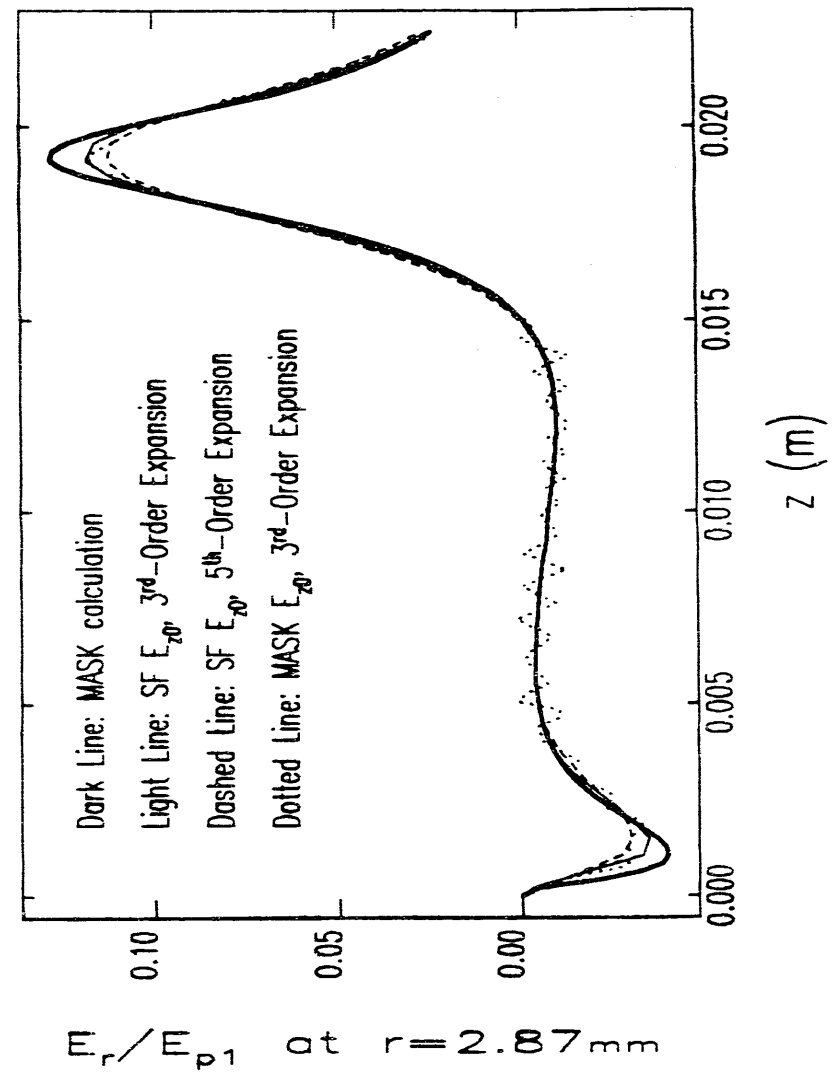

Figure 2.31: Comparison of $E_r(z)$ for r=2.87mm, as Calculated by MASK and Using Off-Axis Expansions of Various Orders

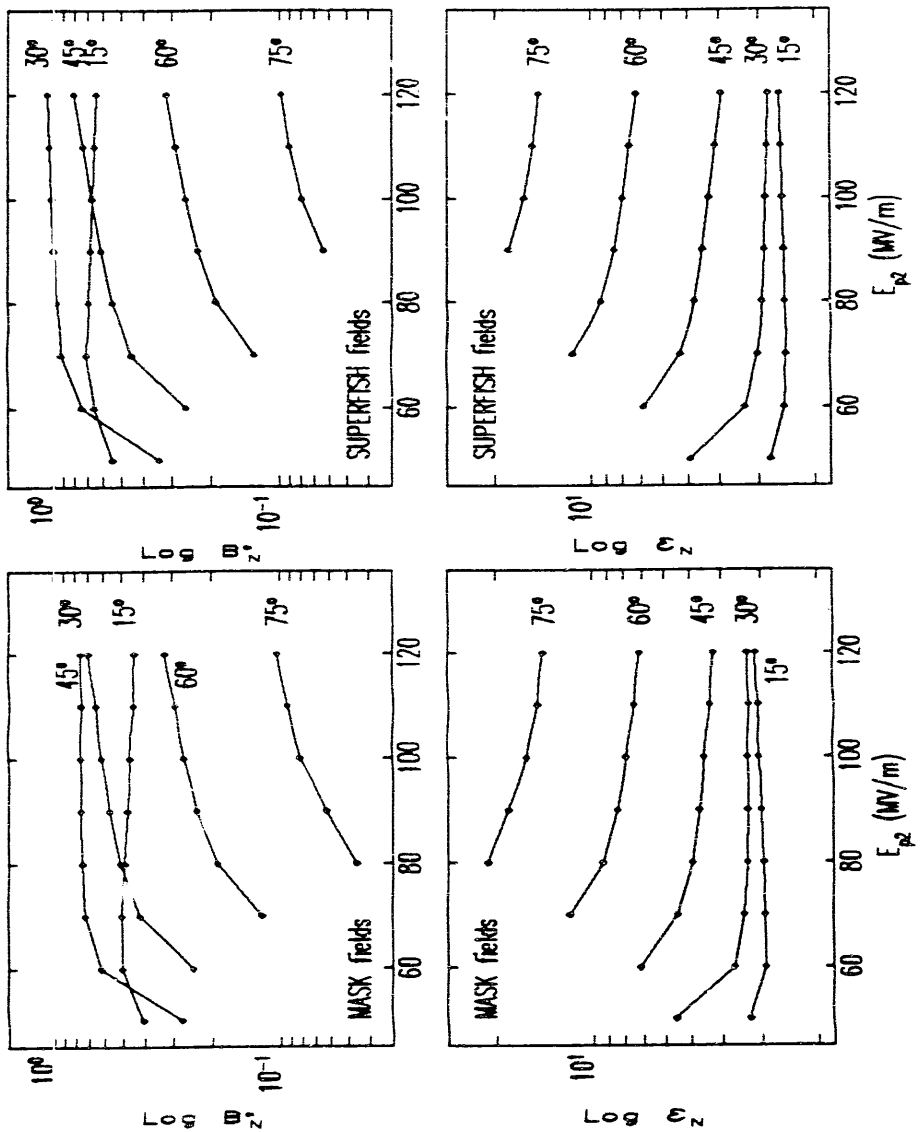

Figure 2.32: Comparison of rfgun Emittance Predictions for Various Initial Phase-Intervals, for MASK- and SUPERFISH-Calculated Fields

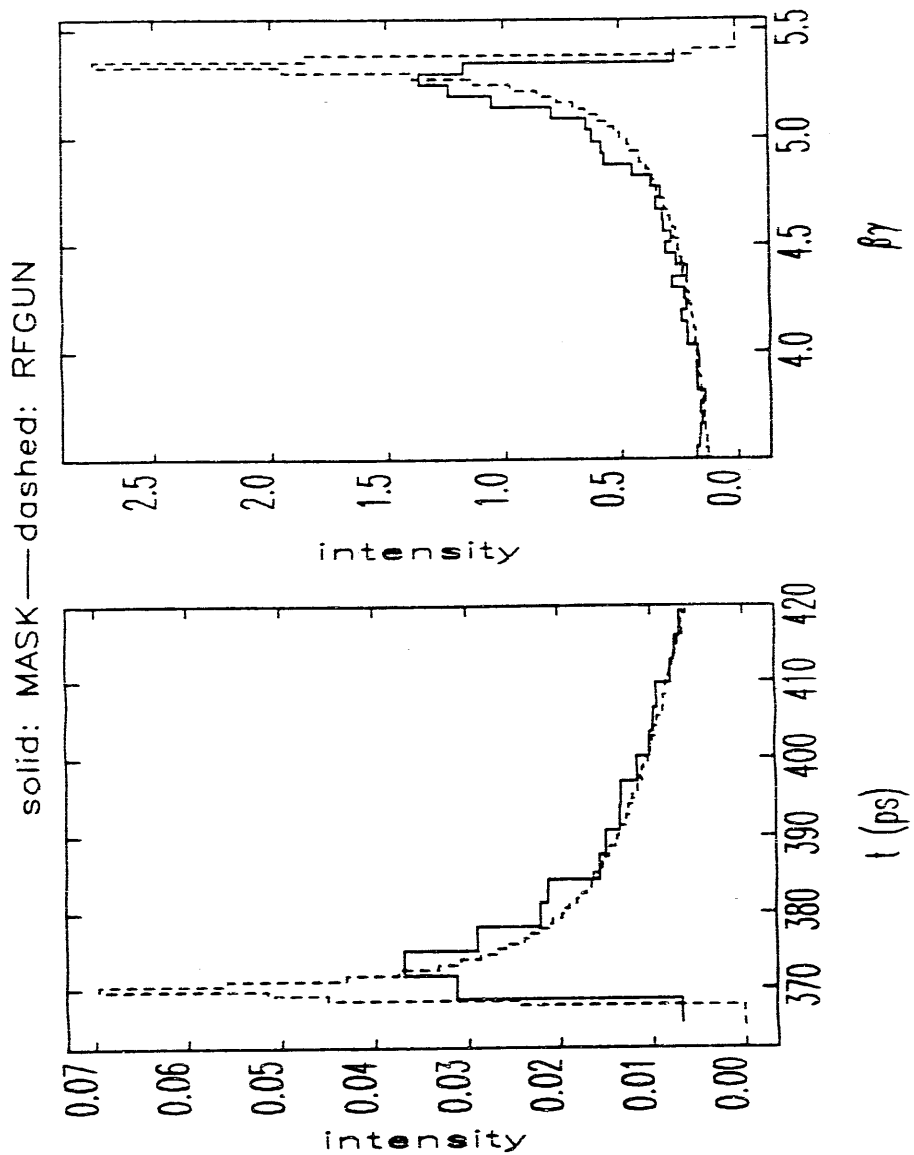

Figure 2.33: Comparison of rfgun results for MASK-Calculated Fields with MASK Calculations for J — 0

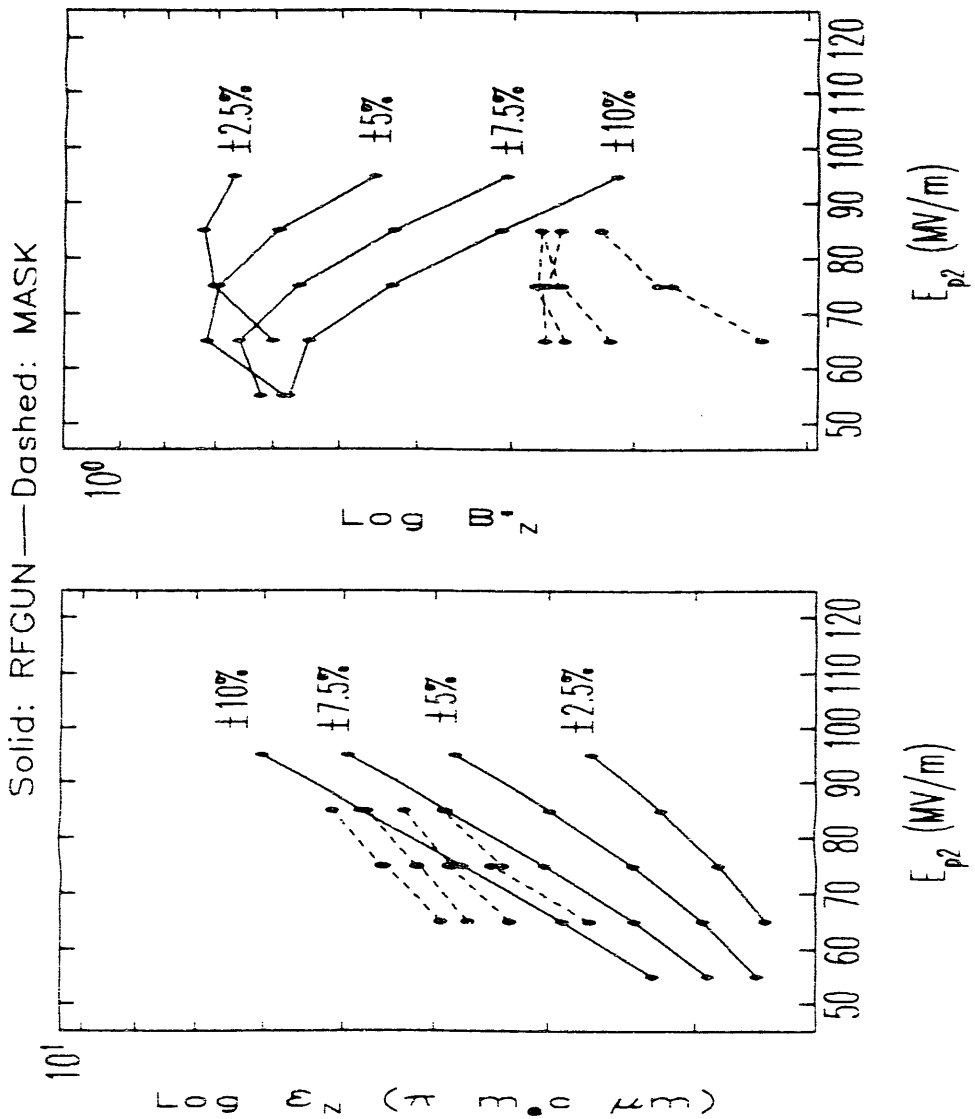

Figure 2.34: Comparison of rfgun results for MASK-Calculated Fields with MASK Calculations for J — 0

discrepancy is that, even using the MASK-calculated on-axis field profile, **rfgun** cannot duplicate the fields used in MASK, since the expansion used in **rfgun** is not of sufficiently high order.

Next, I show in Figure 2.34 a comparison of the momentum and time distributions for MASK and **rfgun** with E_{p2} = 75MV/m and J — 0 in MASK, where the MASK-calculated fields are again used in **rfgun**. As one would expect, **rfgun** predicts narrower spectra because of the larger non-linear fields in MASK. I have also found that MASK consistently predicts about 3% greater maximum momentum than is predicted by **rfgun** for the same value of E_{p2} (with a = 3 and f = 2836MHz in **rfgun** in order to match MASK). (Because of this discrepancy, I ran **rfgun** with E_{p2} = 77.1MV/m in order to match the peak momentum to that of **rfgun** for comparison of the time and momentum spectra in Figure 2.34.) One possible explanation for this is that the phasing of the first and second cell is imperfect. Because of the frequency mismatch between the two cells (discussed in the previous section), the cells drift out of frequency by about 0.5° during one RF period. This would seem to be too small to have the observed effect, however.

Another confirmation of the effect of the larger non-linear fields in MASK is obtained by running MASK with a smaller cathode. While I will not take the space to show these results, I have found that running MASK with R_c — $R_c/2$ produces a noticeably smaller momentum and time spread in the final beam. Quantitative results can be found in Chapter 4.

2.3.12 Space-Charge Effects

Transverse Phase-Space

I turn now to the effects of space-charge as predicted by MASK. Figure 2.35 shows the evolution of the transverse phase-space in the first cell for E_{p2} = 75MV/m and J = 80A cm^{-2}. These are to be compared to those shown in Figure 2.28 for J — 0. The effect of the space-charge forces for this high current density are clearly evident. From these two figures, it is apparent that the space-charge forces tend to counter the cavity fields, since the slopes are significantly more positive for J = 80A/cm^2. This is as expected, since the space-charge forces are radially defocusing. One effect of this

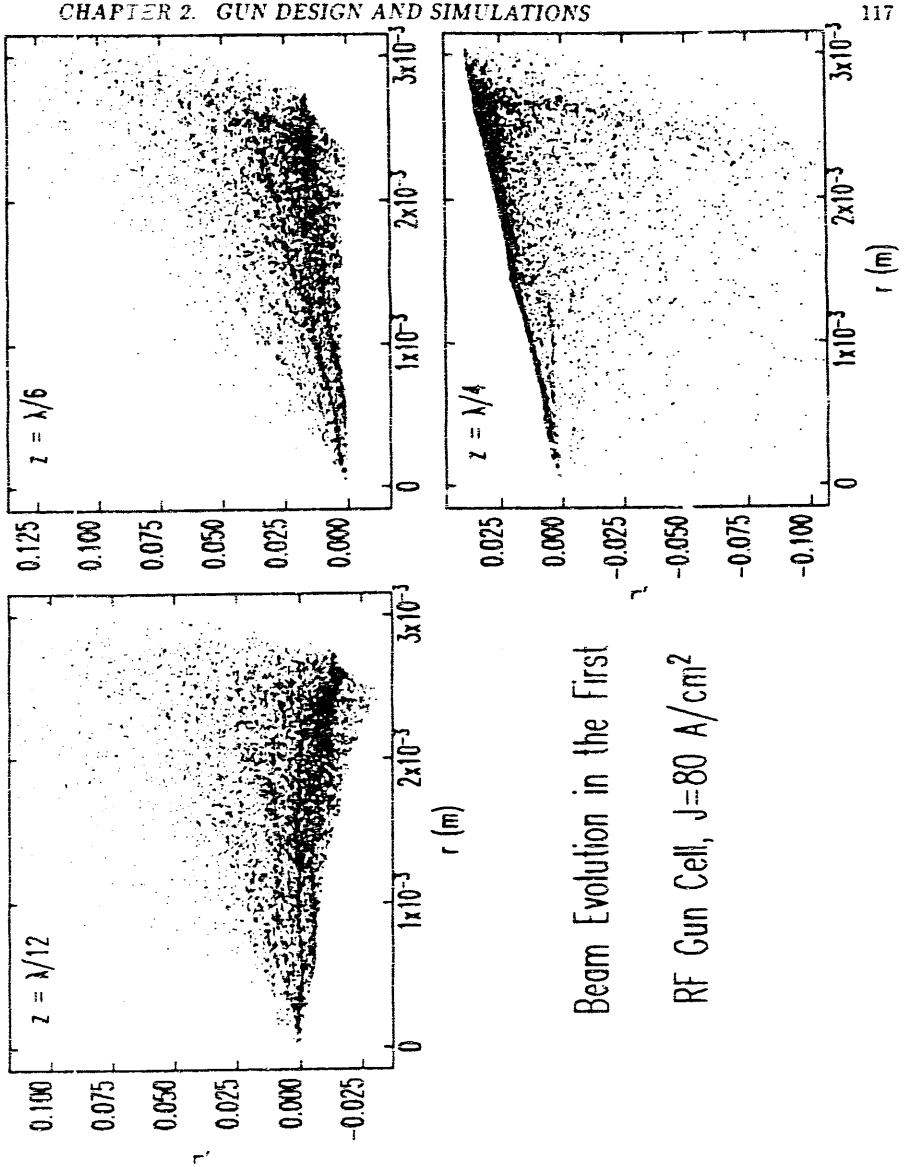

Figure 2.35: MASK-Calculated Transverse Phase-Space Evolution in the First Cell, for J = 80A/cm^2

defocusing is a significantly larger beam size at the exit of the first cell, as well as a larger spread in slopes.

It would appear from the Figure that the space-charge forces not only counter the linear focusing forces, but also compensate for non-linear fields as well. For this to be the case, one would require that at some point the radial distribution of charge be non-uniform and that the charge density increase with radius. To see why, consider that for a longitudinally uniform, cylindrically symmetric beam the radial electric field is given by Gauss's law[31]:

$$E_{r,beam}(r) = \frac{1}{\epsilon_o r}\int_0^r \rho(\bar{r})\bar{r}\,d\bar{r} \tag{2.66}$$

where $\rho(r)$ is the charge density per unit cross-sectional area. For a uniform radial distribution, $\rho(r) = \rho_o$ out to to some radius R_b (the edge of the beam), after which it falls to zero. Hence, for this case,

$$E_{r,beam}(r) = \frac{r\rho_o}{2\epsilon_o}, \qquad r \le R_b \tag{2.67}$$

which is simply an additional linear field term.

Next, consider what happens if $\rho(r) = \rho_o(1 - \eta r^n)$ for $r \le R_b$, where n is an integer:

$$E_{r,beam}(r) = \frac{r\rho_o}{2\epsilon_o}\left(1 + 2\eta\frac{r^n}{n-1}\right) \qquad r \le R_b \tag{2.68}$$

As I discussed in the previous section, $E_{r,beam}$ must be an odd function of r, and hence n must be even. In order for the space-charge forces to increase at a greater-than-linear rate with radius, η must clearly be positive, which implies that the beam must be somewhat more hollow than a radially uniform beam.

Figure 2.36 shows histograms of the intensity vs radius for $z = \lambda/12$ for the cases $J \to 0$ and $J = 80A/cm^2$. Each bin in the histograms represents an annulus, with the height of the histogram being proportional to the charge in that annulus. For a uniform distribution, one would expect a linear function of r, since for this distribution the height of the bin that starts at $r = n\Delta r$ is

$$H_n = 2\pi\rho_o\int_{n\Delta r}^{(n+1)\Delta r}\rho(\bar{r})\bar{r}\,d\bar{r} = \pi\rho_o\Delta r^2(2n-1). \tag{2.69}$$

For $J \to 0$, one sees that H_n increases faster than linearly. This is due to the non-linear increase in focusing fields with radius, and might have been anticipated from

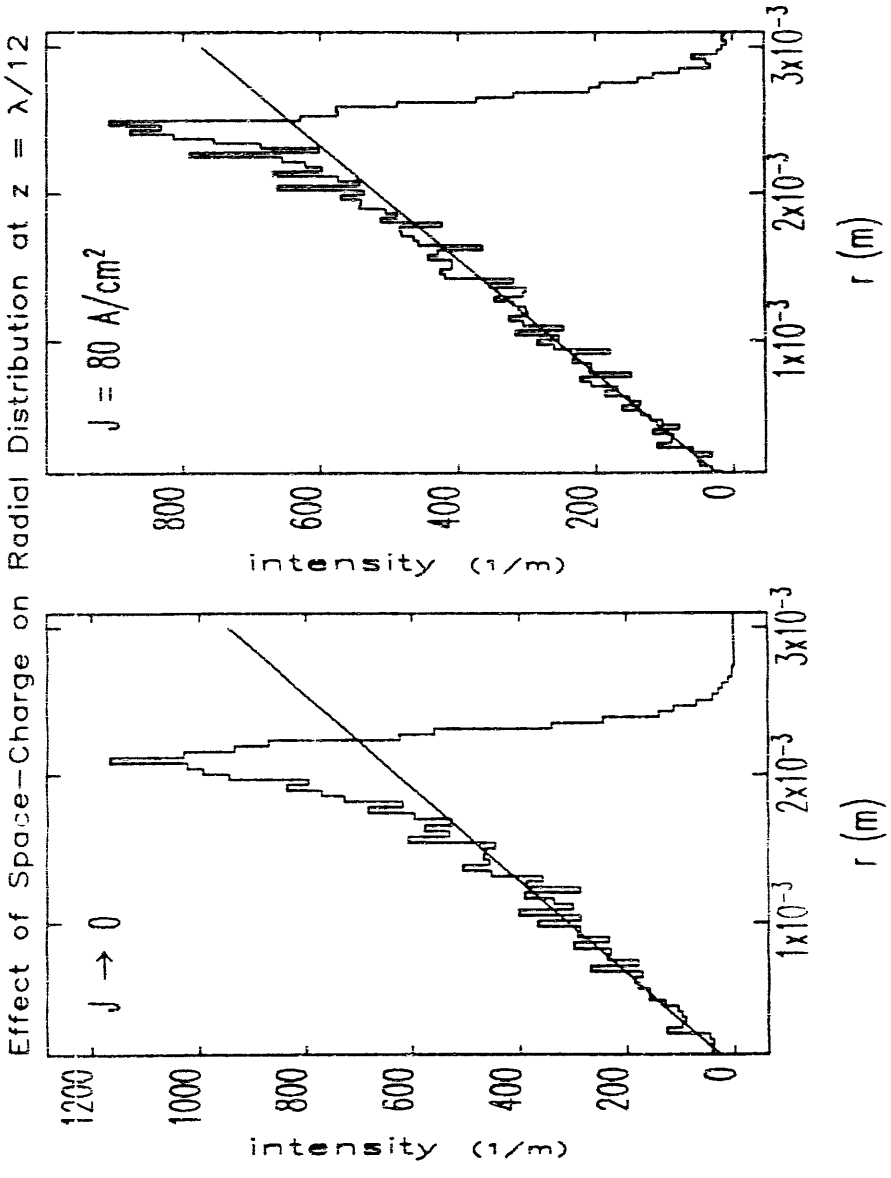

Figure 2.36: MASK-Calculated Transverse Beam Distribution at $z = \lambda/12$, for J — 0 and $J = 80 A/cm^2$

Figure 2.28. For J = 80. the distribution is clearly much more uniform. The non-linear cavity fields and space-charge forces in this case tend to balance each other, since the non-uniform distribution that the non-linear cavity fields try to create is just the kind of distribution that is necessary to counter these self-same non-linear cavity fields. One expects the radial beam distribution to have just enough radial non-uniformity to compensate the non-linear cavity fields.

It is not at all apparent from these Figures what the net effect on the emittance is. The beam is larger over-all for the case with high space-charge, but the correlation would also seem to be higher. I will show below that the emittance is in fact substantially larger for the high space-charge case.

Longitudinal Phase-Space

I next look at the effects of space-charge on the longitudinal phase-space. Figure 2.37 shows the effect of space-charge on the longitudinal phase-space at the gun exit. The longitudinal space-charge forces are seen to broaden the momentum and time distributions, much as the non-linear forces do. This is to be expected, since particles at the head of the beam are accelerated by the particles that follow, while trailing particles are decelerated. This broadens the momentum spectrum because it amplifies the existing distribution, namely that leading particles have more momentum than trailing particles. It broadens the time distribution simply because momentum is monotonically related to time-of-flight in the gun. Further broadening occurs because the time-varying nature of the cavity fields results in additional acceleration of those particles that are pushed ahead, and less acceleration of those that are pushed back.

In order to get a more detailed look at the phase-space distributions for the two cases, I have compressed the longitudinal phase-space using a ideal alpha-magnet and drift space system, as described in Chapter 3. The results are shown in Figure 2.38. Several effects are apparent in this Figure. First, the highest-momentum part of the beam is more energetic for the high space-charge case, due to acceleration by the fields of trailing particles that occurs in the gun; as a result, the "top" of the beam falls further behind the centroid during compression, because the delay in the alpha-magnet increases with increasing momentum. Second, the time-spread for a given small momentum slice is significantly broadened; this is a result of the

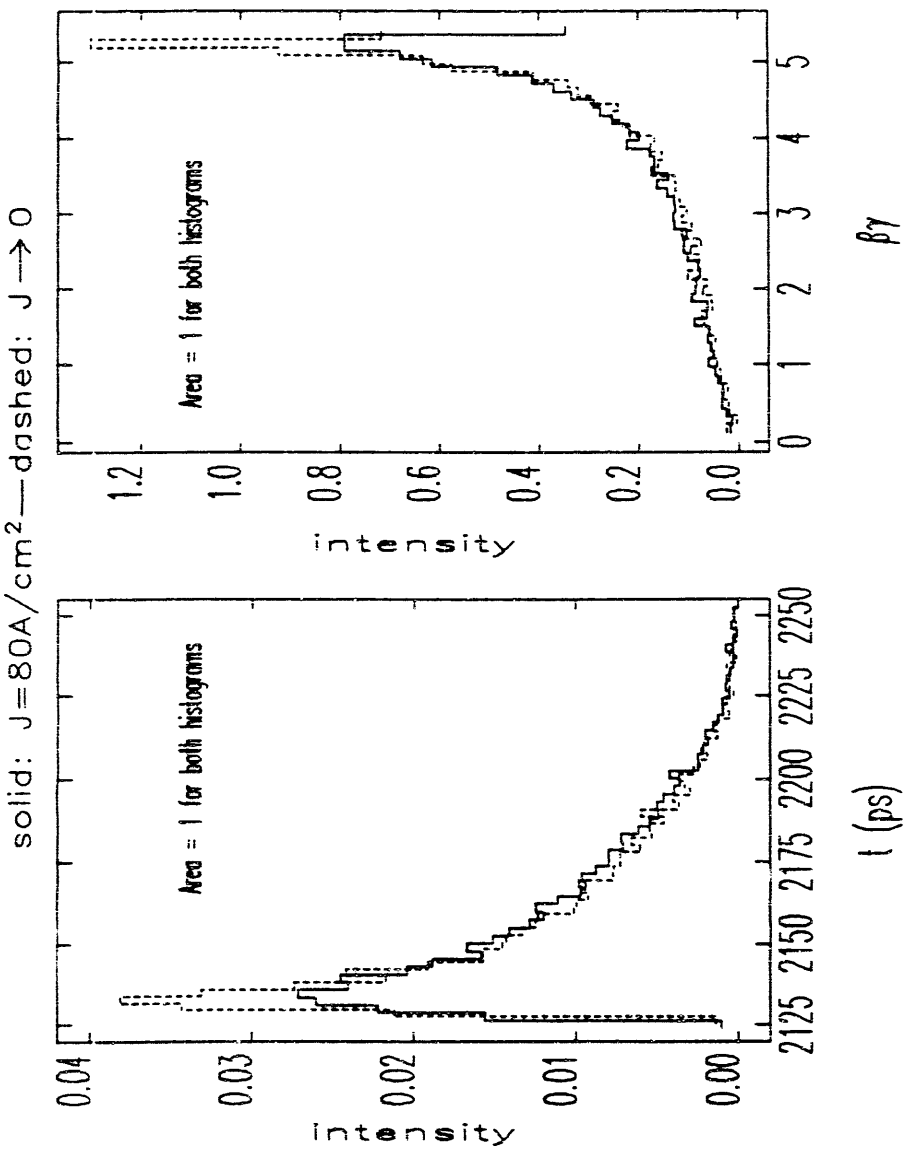

Figure 2.37: MASK Longitudinal Phase-Space Distributions at the End of the Second Cell

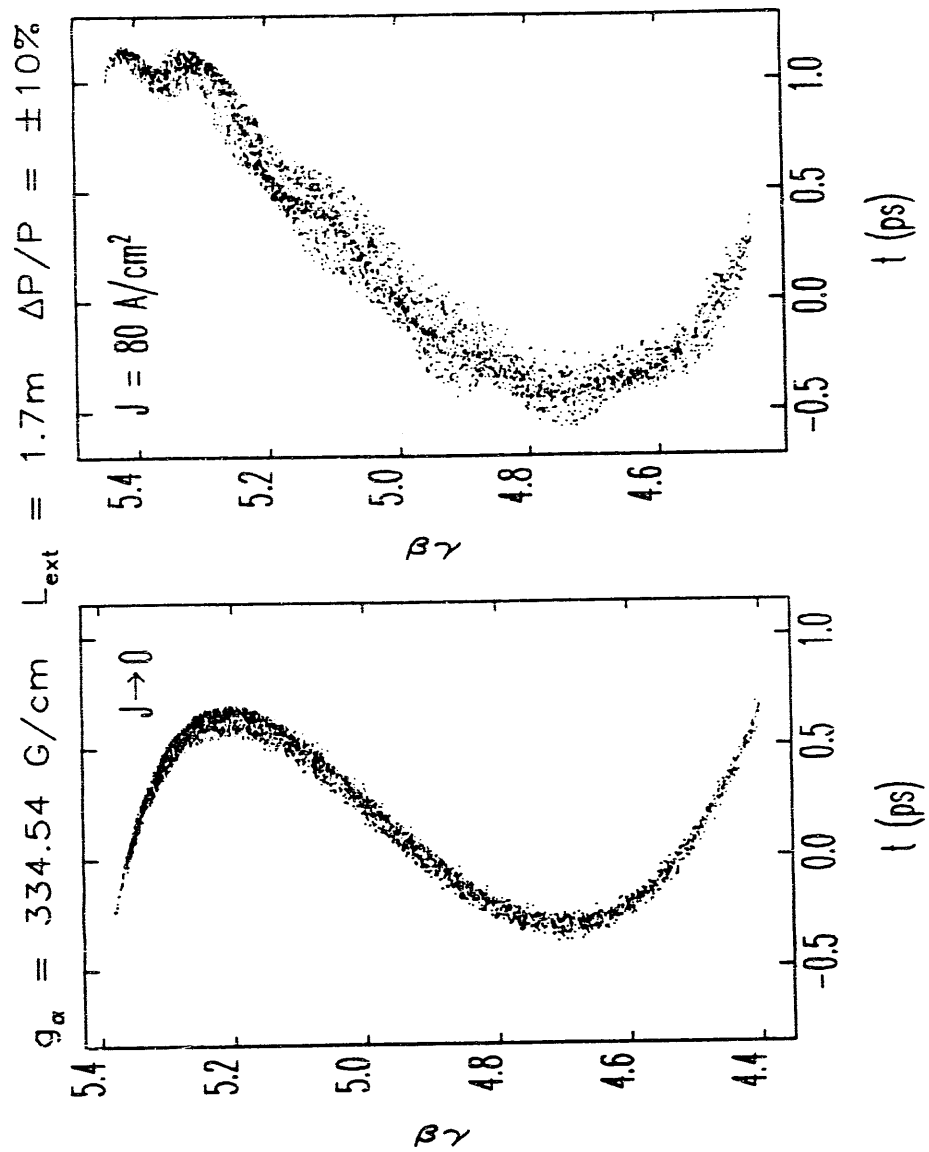

Figure 2.38: MASK Longitudinal Phase-Space Distributions after Alpha-Magnet-Based Compression

longitudinal space-charge forces being a non-linear function of radius, which produces longitudinal mixing. Third, there is a clear oscillation in the p(t) curve, especially evident near the top of the beam; this is apparently a plasma oscillation, resulting from the longitudinal space-charge forces. Finally, one sees that the space-charge forces broaden the beam significantly relative to the broadening by non-linear fields (without which the phase-space distribution for $J \rightarrow 0$ would be a line).

2.3.13 MASK Predictions of Gun Performance

Figures 2.39 through 2.41 show results for the emittance, brightness, and charge as calculated by MASK, as a function of current density and peak on-axis electric field, for various final momentum fractions. The smaller range of E_{p2} in these results (as compared to those for `rfgun`) is a result of my concentrating computer resources on the range that is of most interest for actual running at SSRL. One sees that the normalized charge per bunch decreases as current density increases, a result of the longitudinal forces in the beam, which may be thought of as decreasing the current density by forcing the electrons apart. As one might expect, this effect lessens as the cavity fields are increased, since this decreases the strength of the particle-induced fields relative to the cavity fields and results in faster acceleration, thus decreasing the effect of the particle-induced fields further.

The trends in emittance and brightness hold some surprises. In particular, the emittance does not always increase when the current density is increased: for small momentum intervals, the opposite occurs. There are two effects that may explain this. First, as was seen above, the particle-induced fields tend to counter the non-linear cavity fields, which would in turn tend to limit emittance growth due to those non-linear fields. Second, space-charge related changes in the longitudinal phase-space result in there being a larger phase-interval represented in a given final momentum fraction for small current density than for a large current density. As was seen above, emittance depends strongly on the initial phase-interval one considers. Hence, it should not be surprising that when one takes a very small final momentum fraction, this effect becomes apparent, since for small final momentum fractions the initial phase-interval is smaller, whereas the effects of longitudinal space charge are great

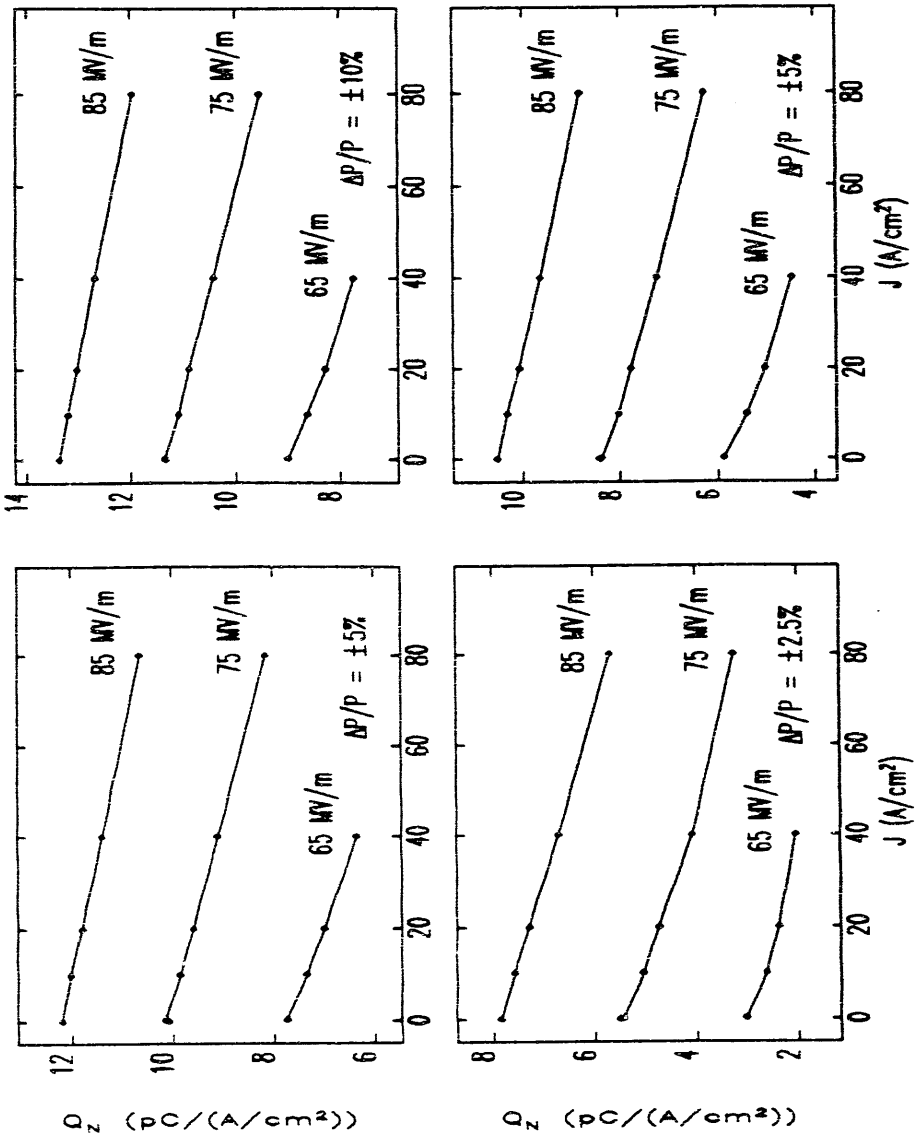

Figure 2.39: MASK Results for Normalized Charge Per Bunch

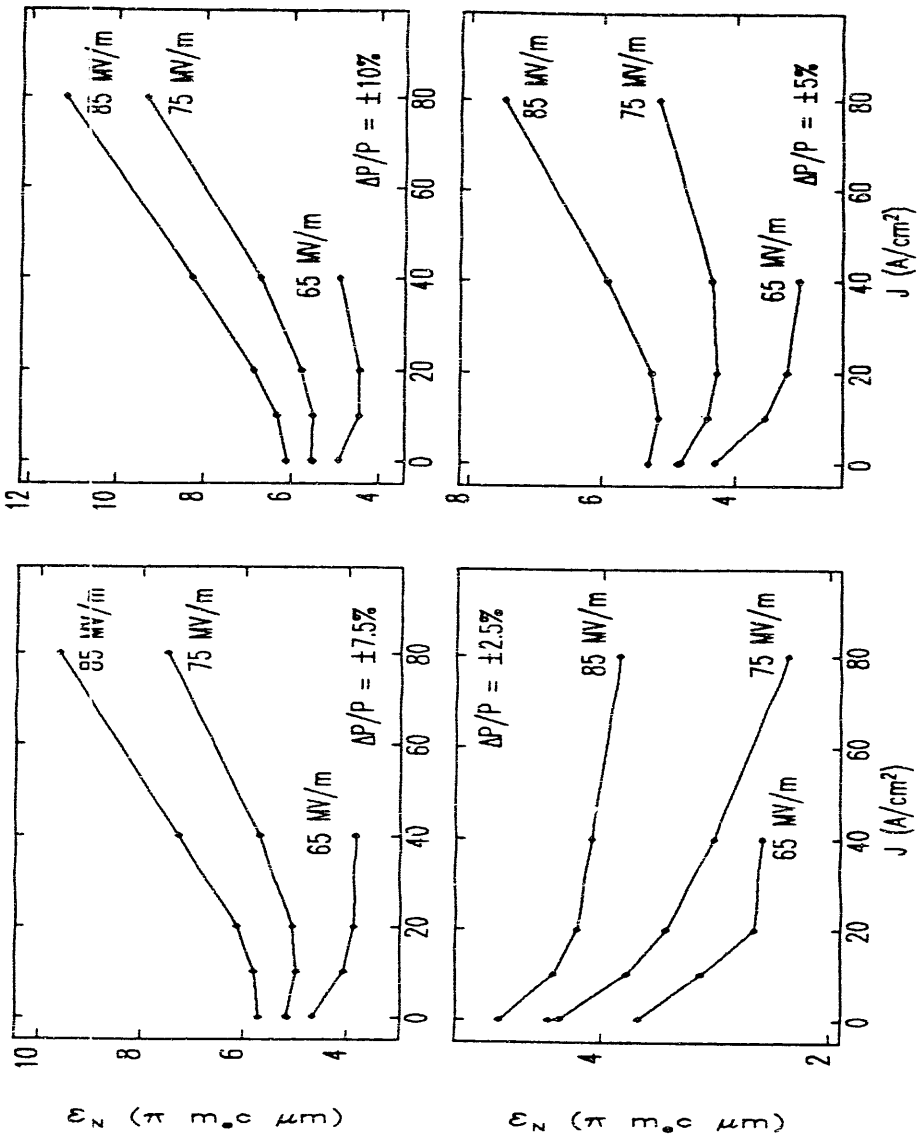

Figure 2.40: MASK Results for Normalized RMS Emittance

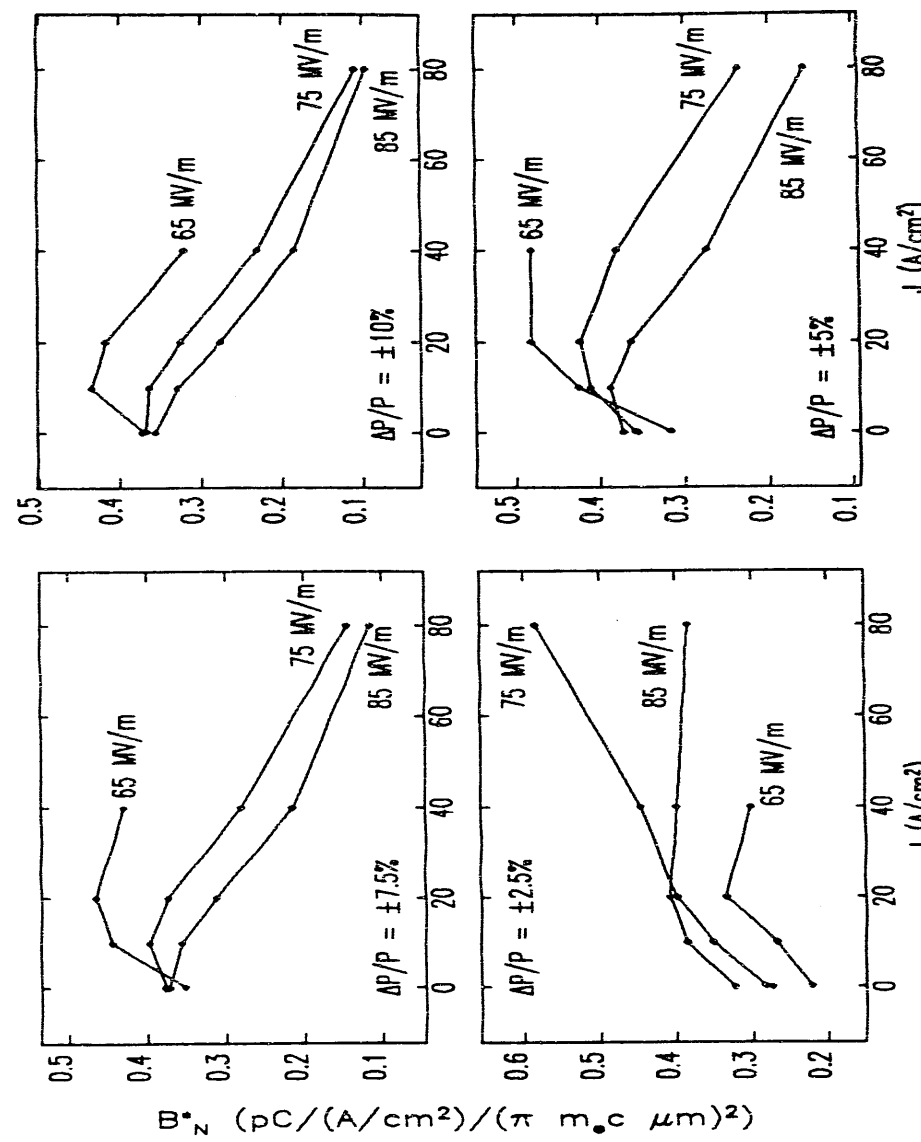

Figure 2.41: MASK Results for Normalized Brightness

(since the charge distribution is so strongly peaked).

Note that the brightness shown here, as throughout this section, is normalized to the current density. Hence, Figure 2.41 does not show the actual brightness decreasing with increasing current density. Figure 2.42 shows the transverse brightness, $B_t = B_n^* J$, without normalization to the current density. One sees that the transverse brightness increases rapidly with current density at first, then saturates as space-charge effects on the emittance overcome the increasing charge per bunch.

Finally, to provide a more complete picture, Figures 2.43 through 2.45 show transverse phase-space distributions for $E_{p2} = 75$MV/m and $\Delta P/P = 10\%$, for a range of current densities. As noted previously, the distributions for higher current density show less curvature due to the balancing of non-linear cavity fields by space charge. Printed on the graphs are the RMS beam-sizes and beam-divergences. One sees that the beam is predicted to be quite small at the gun exit, but that the RMS divergence is rather large.

Additional performance data will be presented in Chapter 4, where I include the effects of the gun-to-linac transport line, and in particular the effects of the alpha-magnet. In addition, Chapter 4 gives comparative data for other RF gun and DC gun systems.

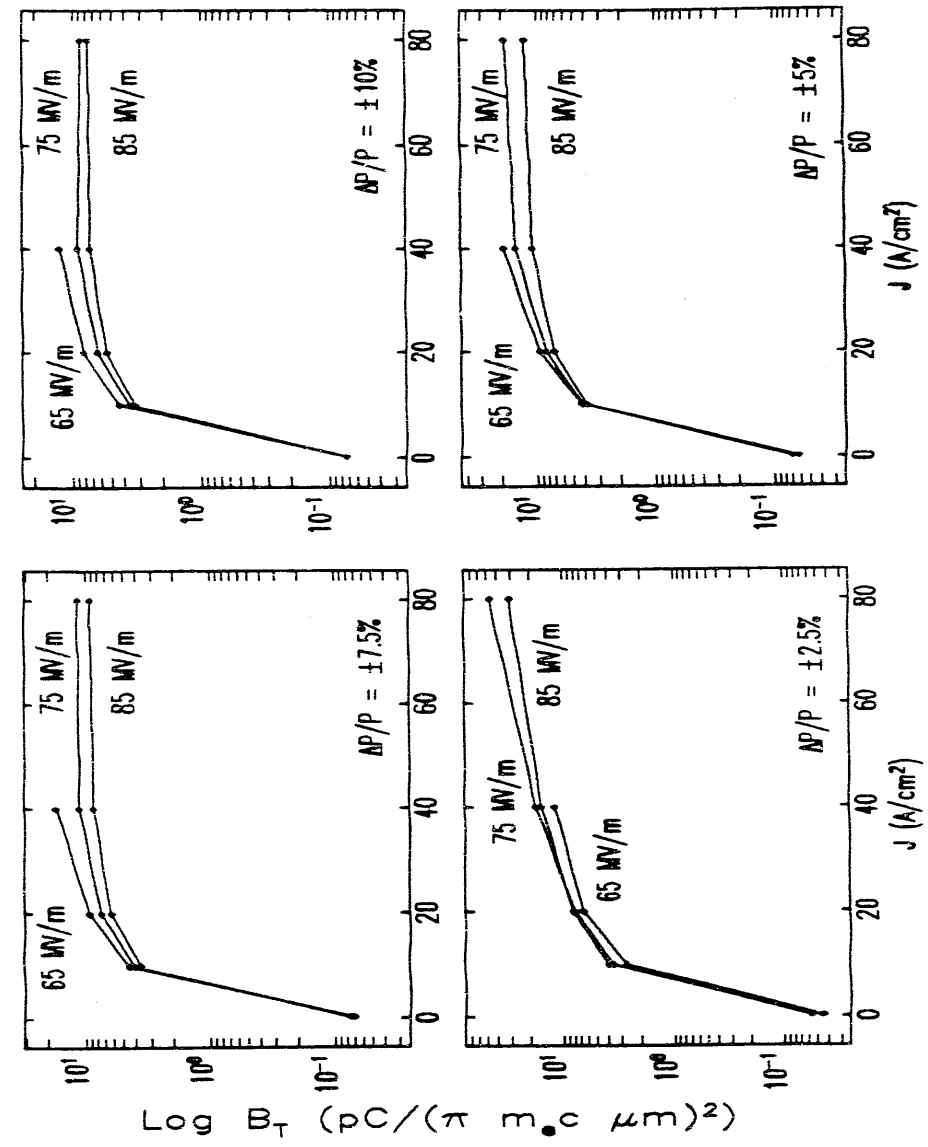

Figure 2.42: MASK Results for Transverse Brightness

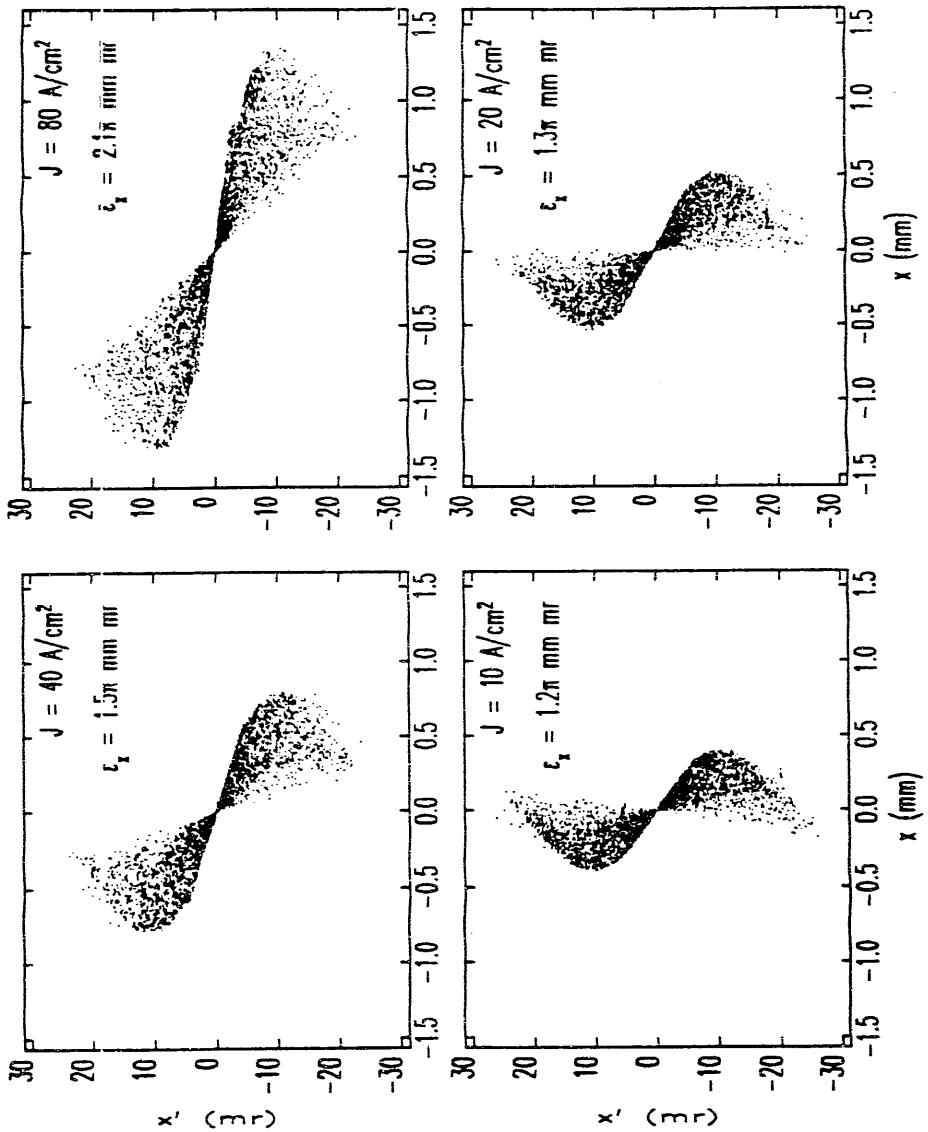

Figure 2.43: MASK Results for Transverse Phase-Space Distribution, for E_{p2} = 75 MV/m and $\Delta P/P$ = 10%—Part 1

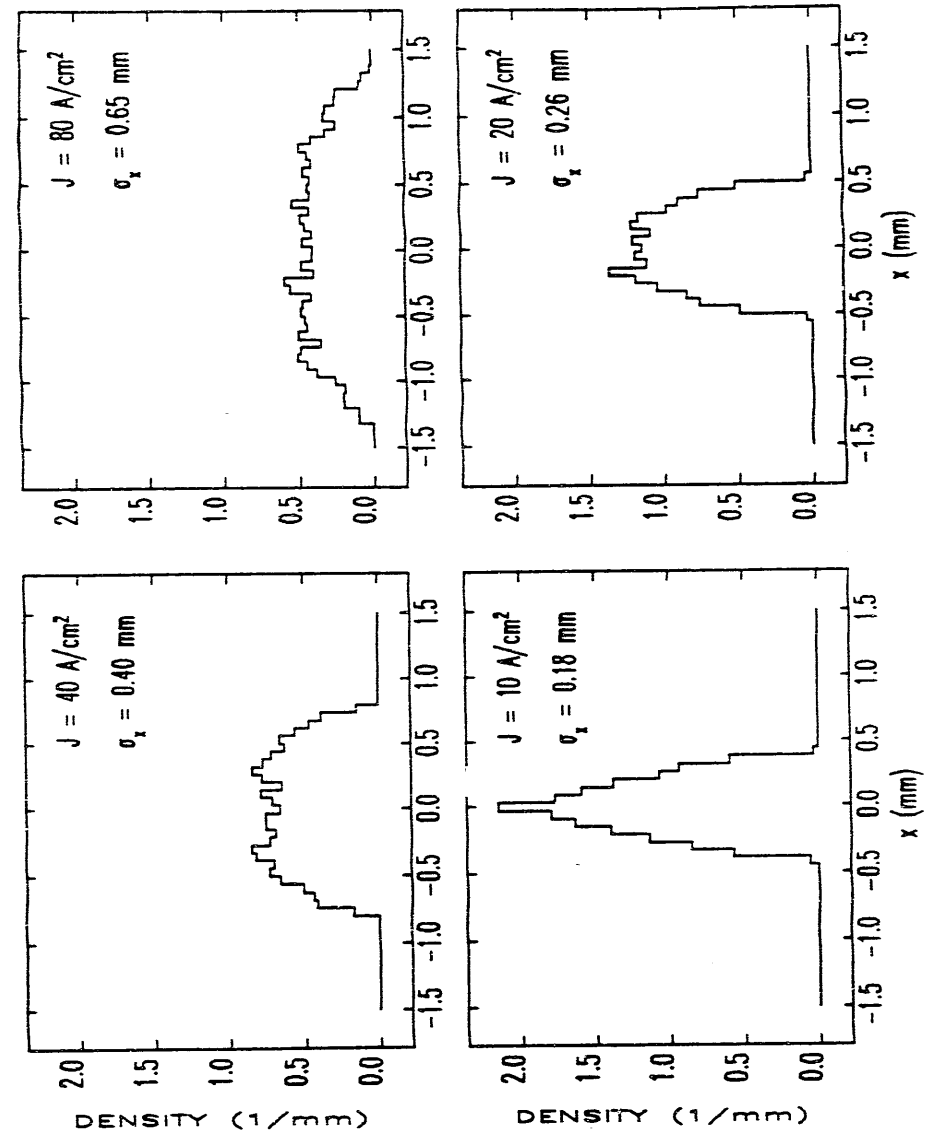

Figure 2.44: MASK Results for Transverse Phase-Space Distribution, for E_{p2} = 75 MV/m and $\Delta P/P$ = 10%—Part 2

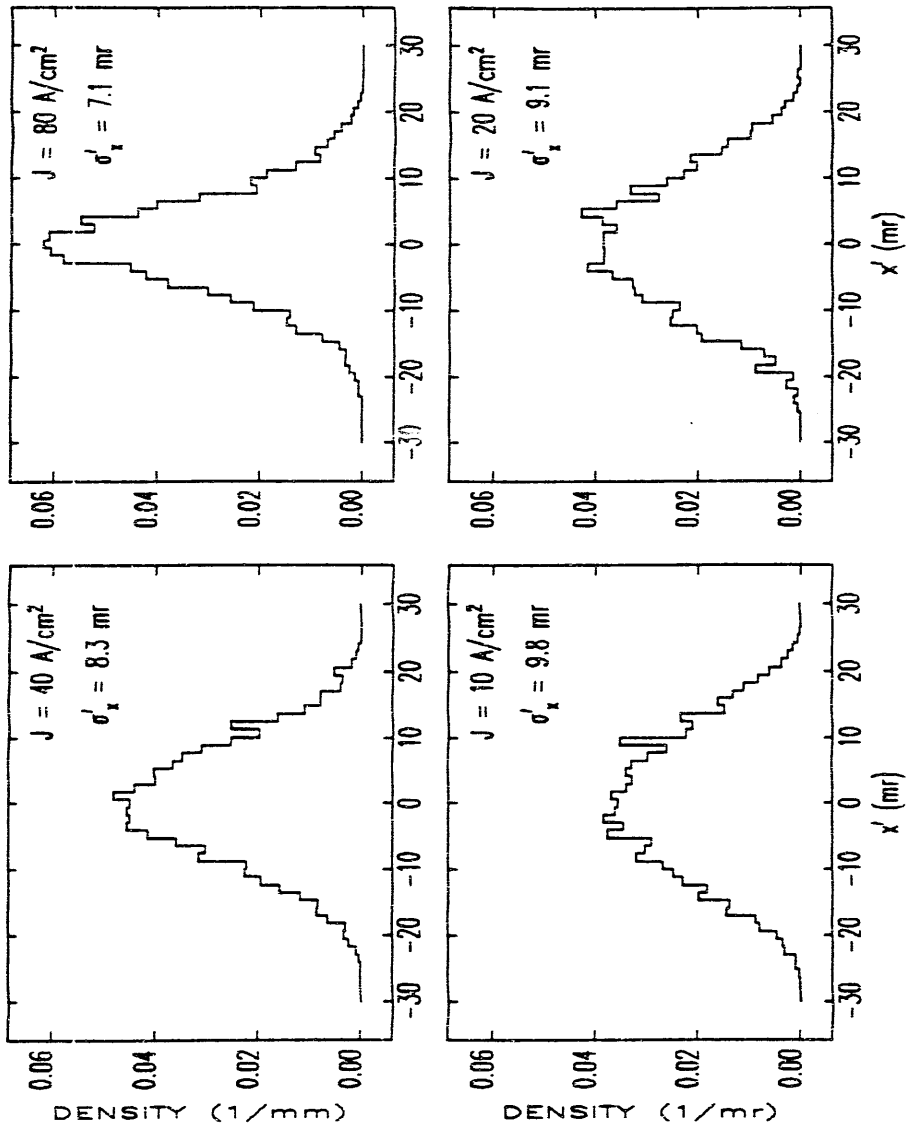

Figure 2.45: MASK Results for Transverse Phase-Space Distribution, for E_{p2} = 75 MV/m and $\Delta P/P$ = 10%—Part 3

Chapter 3

The Alpha-Magnet

As will be discussed fully in Chapter 4, the beam directly out of the gun is not suitable for injection into a S-band linear accelerator section. Doing so would produce an accelerated beam with a large energy spread because of the large phase-spread the particles coming into the accelerator section would have in the absence of compression. Magnetic bunch compression is one solution to this problem, and the one which is most suitable for use with the RF gun. Indeed, the possibility of using magnetic compression, as opposed to RF bunching, is one of the attractive features of the RF gun.

The theory of magnetic compression will be discussed fully in the next chapter, along with the motivation for using an alpha-magnet. In this chapter, I will describe the alpha-magnet and derive its main properties. First, I will discuss the magnetic design of the SSRL alpha-magnet, which is an asymmetric quadrupole, and contrast this design with an alternative design, namely a Panofsky quadrupole. Second, I will present the equation of motion in an alpha magnet, and show how a scaled form of the differential equation can be used to deduce some of the magnet's properties, without integration. I will prove that the transport matrices for any alpha magnet can be expressed in terms of transport matrices for this scaled equation of motion. I will show how these latter transport matrices can be derived from fits to the results of numerical integration of the scaled equation of motion for an appropriately selected ensemble of particles. I will present the results of a calculation of alpha-magnet transport matrices to third order, along with discussion of the accuracy of the results. Having

calculated matrices for a perfect alpha-magnet. I then discuss how to extend the treatment to imperfect alpha-magnets, specifically those with multipole and beam-hole-induced field errors. Finally, I present the results of experimental measurements of the SSRL alpha-magnet, including magnetic measurements and measurements of some first-order matrix elements.

3.1 Magnetic Characteristics and Design of the Alpha-Magnet

The alpha-magnet and its properties were first described by Enge[45]. It is essentially half of a quadrupole magnet, with a symmetry plane at $q_1 = 0$, i.e., with a vertical mirror plane along the longitudinal axis. This mirror plane provides the symmetry necessary to obtain quadrupole-like fields in the interior of the magnet. Figure 3.1, a simplified cross-sectional view of the alpha-magnet designed for the SSRL project, illustrates these points and anticipates the discussion to follow. Rather than inject the beam along the quadrupole axis (as might be done if the magnet where to be used as a combined-function dipole and quadrupole), the beam is injected through the "front-plate", i.e., through the iron piece that functions as an approximation to an ideal magnetic mirror-plane.

3.1.1 Asymmetric Quadrupole Design

To understand this in more detail, it is convenient to use the approximation that the permeability of iron is infinite. In this case, Maxwell's equations at a material boundary mandate that the magnetic field $\mathbf{H}$ just outside the iron be perpendicular to the iron surface. (For a full discussion of several of the points that follow, see J.D.Jackson, [31].) It follows that the iron surfaces are equipotentials of the magnetic scalar potential Φ_M, which is related to the magnetic field by

$$\mathbf{B} = \mathbf{H} = -\nabla\Phi_M, \tag{3.1}$$

where I employ Gaussian units, and use the fact that $\mathbf{B} = \mathbf{H}$ in air.

An infinitely-long quadrupole magnet is defined as one that has a magnetic field given by

$$\mathbf{B} = g(q_1\hat{q}_3 - q_3\hat{q}_1), \tag{3.2}$$

where g is the quadrupole gradient, and where $\hat{q}_1$, $\hat{q}_2$, and $\hat{q}_3$ form a right-handed coordinate system (The reason for the unusual choice of coordinates—(q_1, q_2, q_3) instead of the usual (x, y, z)—is for consistency with subsequent sections of this chapter.) The

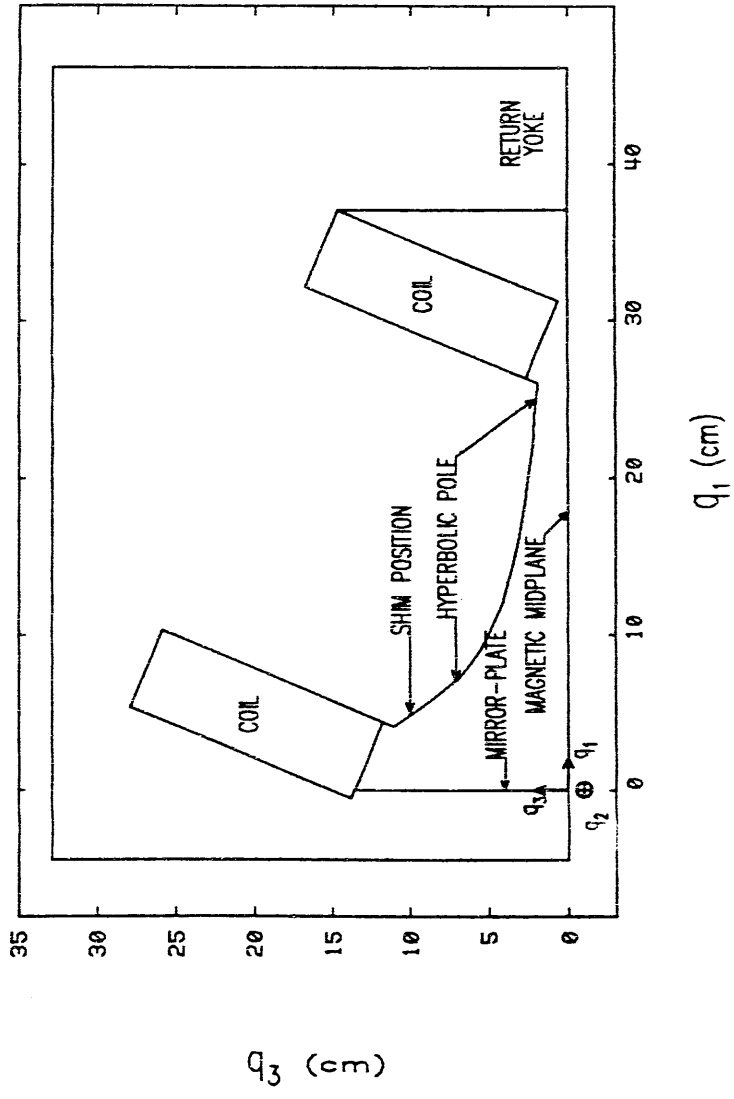

Figure 3.1: Simplified Cross-sectional view of the SSRL alpha-magnet.

reader can verify that this field satisfies Maxwell's equations. and also that it can be derived from the magnetic potential

$$\Phi_Q = -gq_1q_3. \tag{3.3}$$

Knowing the magnetic potential necessary to produce quadrupolar magnetic fields allows one to specify the location of equipotential surfaces that will produce such a field. That is. if one arranges magnetic surfaces and suitable driving currents so as to obtain equipotentials of a quadrupolar field on the magnetic surfaces, then the region inside the boundary formed by the magnetic surfaces will contain a quadrupolar field distribution. While it is by no means essential to do so. this is typically accomplished by a four-fold symmetric arrangement of iron. where alternate poles of the magnet have the same potential except for a change in sign. Since the magnet poles are equipotentials, they must be hyperbolic in shape. (This brief exposition does not show the full power of the equipotential method in treating multi-pole fields, for which the reader should consult other sources.[6])

From the definition of the quadrupole field, it follows that the lines $q_1 = 0$ and $q_3 = 0$ are equipotentials with $\Phi = 0$. Hence, if a magnetic surface is placed along the line $q_1 = 0$ extending into $q_1 < 0$, then the field in the region $q_1 > 0$ is unchanged, since the locations and shapes of the equipotentials are unchanged. This is what is done for the asymmetric quadrupole alpha-magnet design used for the SSRL project. The reader is referred again to Figure 3.1, which exhibits the truncated hyperbolic poles and the mirror-plate along $q_1 = 0$. This design is called "asymmetric" because the hyperbola extends further horizontally than vertically, in order to obtain a large horizontal good field region. The deviation from the hyperbolic equipotential surface that is implied by truncation of the hyperbola is made up for by "shiming" the pole with additional magnetic material near the upper end of the hyperbola. This is a trail-and-error process that was carried out using the magnet code POISSON[66].

The resultant calculated gradient in the $q_3 = 0$ plane is shown in Figure 3.2, along with measurements performed on the magnet before the beam entrance/exit hole was cut in the mirror plate. Note that the way the data is normalized means that one should compare the shapes of the curves rather than the absolute agreement. I used a linearized Hall probe for these measurements (as well as those presented below),

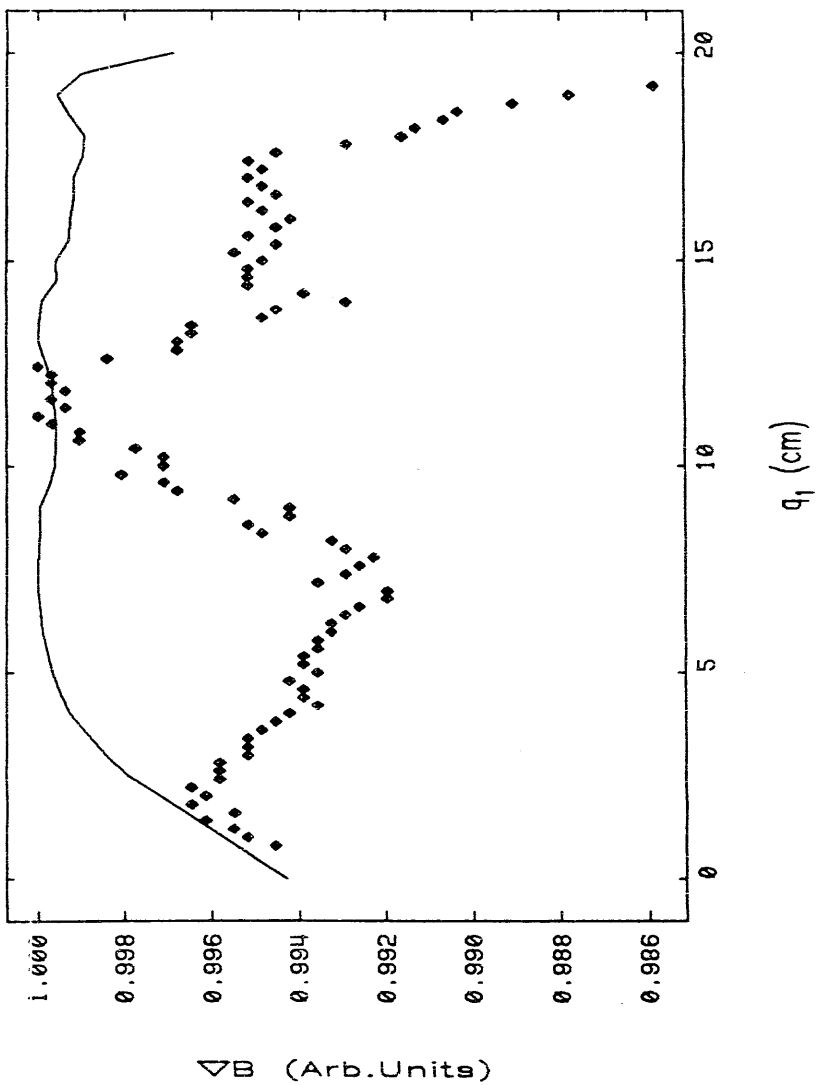

Figure 3.2: Computed and Measured Gradient of the SSRL Alpha-Magnet

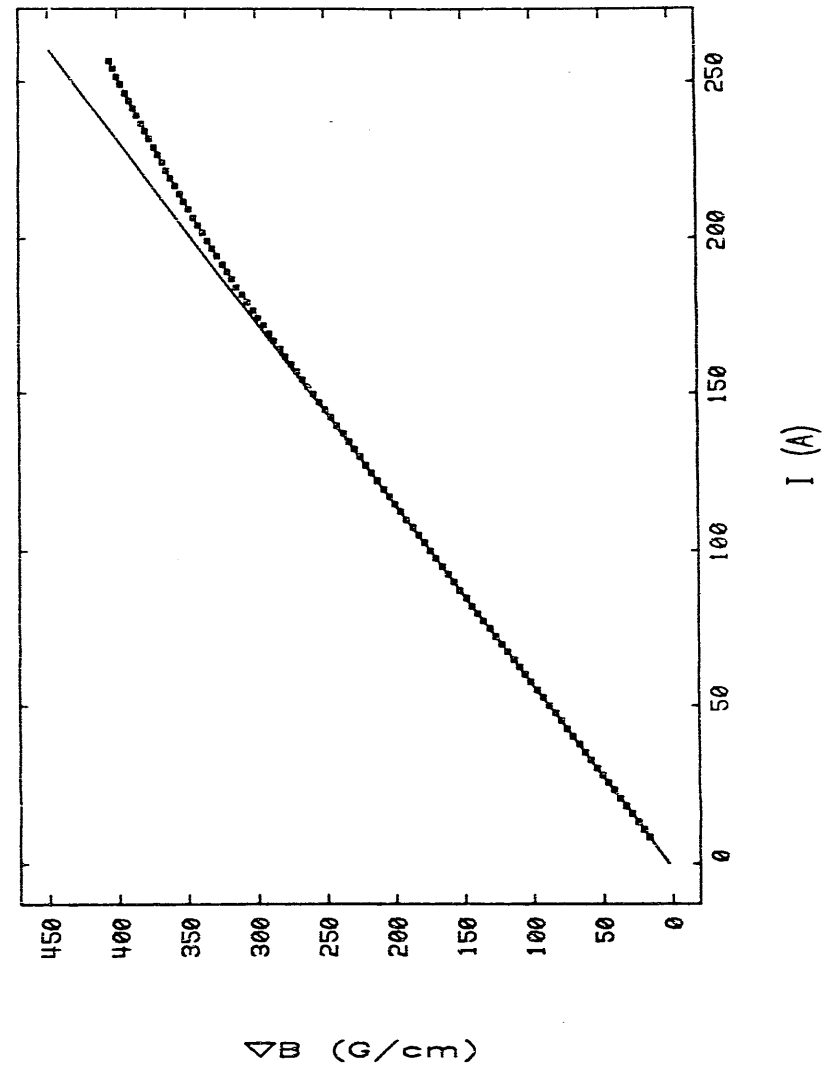

Figure 3.3: Measured Excitation Curve of the SSRL Alpha-Magnet

to ensure that spurious non-linearities did not appear in the data. The discrepancies are believed to be due in part to construction errors in the magnet, which resulted in deviations of the pole profile from the design. Some of the discrepancies are also due to round-off errors and convergence problems in POISSON, which cause the gradient near $q_1 = 0$ to become non-uniform. In any case, the non-uniformities of the gradient for the magnet without a beam port are dwarfed by those introduced when the beam port is cut into the front plate. I will return to this topic later in this chapter. Figure 3.3 shows the measured excitation curve, along with a line showing extrapolating the low-current region of the curve to high currents, which illustrates the effect of saturation. Selected magnet parameters are listed in Table 3.1.

Table 3.1: **SSRL Alpha-Magnet Design Parameters.**

number of turns	80
maximum current	260 A
maximum gradient	405 G/cm
inscribed pole radius	10 cm
good-field region (extent in q_1)	20 cm
gradient uniformity without beam port	.5%
depth (extent in q_2)	40 cm
resistance per coil @ 45°C	40 mΩ

3.1.2 Panofsky Quadrupole Design

Another magnet design that might be employed instead of the asymmetric quadrupole used here is a half Panofsky quadrupole [67] depicted in Figure 3.4. Unlike standard quadrupole designs where the quadrupole field is obtained through the approximately hyperbolic shape of the poles, the Panofsky quadrupole relies on uniform sheets of current to produce a quadrupole field. From 3.4 it can be seen that $J \neq 0$ at the pole surfaces, from which it follows that the fields in the magnet gap are *not* determined solely by the shape of the poles, in contrast to the situation for a standard quadrupole design. The most straight-forward way to calculate the fields in a Panofsky

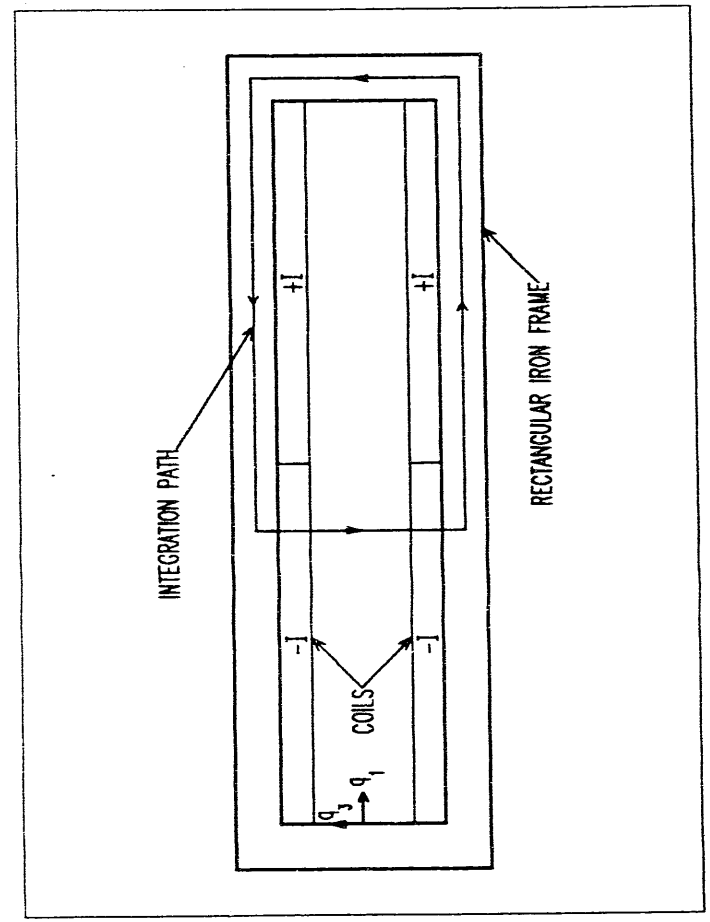

Figure 3.4: Panofsky quadrupole

quadrupole is to use the integral form of Ampere's law:

$$\int \mathbf{H} \cdot \mathbf{dl} = \frac{4\pi}{c} I. \tag{3.4}$$

In more practical units, this can be written as[6]:

$$\int \mathbf{H} \cdot \mathbf{dl} = 0.4\pi I, \tag{3.5}$$

where $\mathbf{H}$ is in Gauss, l is in cm, and I is in Ampere-turns. Taking the integration loop as shown in 3.4 and assuming infinite permeability and that H_y is a function of x only (which must be approximately true for a magnet that is wide compared to its gap-height), one obtains

$$H_3 = 0.8\pi \frac{Jtq_1}{h}, \tag{3.6}$$

where h is the full gap of the magnet, J is the current density in the current sheets, and t is the thickness of the current sheets. The linear dependence of H_3 on q_1 demonstrates that this is indeed a quadrupole. In order to obtain H_1, one employs $\nabla \times \mathbf{H} = 0$, from which it follows that

$$\mathbf{H} = \frac{8\pi}{ch} Jt(q_1 \hat{q}_3 - q_3 \hat{q}_1). \tag{3.7}$$

By comparison with equation (3.2), it is seen that the magnet in Figure 3.4 is, in fact, a quadrupole, with gradient

$$g = \frac{0.4\pi}{h} Jt, \tag{3.8}$$

where J is in A/cm^2, g is in G/cm and t and h are in cm.

3.1.3 Comparison of the Two Designs

A major difference between the Panofsky and asymmetric quadrupole designs for the alpha magnet is the amount of power consumed to produce a given gradient in a specified region. It is this difference that lead to the adoption of the asymmetric design for the SSRL project.

To investigate this, I will assume that what is desired is an alpha magnet with depth D (as perceived in Figures 3.1 and 3.4), useful gap h_u, and good field region G.

using coils made from a metal with resistivity ρ and metal packing-fraction f. Then for the Panofsky quadrupole design, the power consumed is

$$P_{PQ} = 10fJ^2 G g h_u \rho \frac{D+G}{2fJ\pi - 5g}. \tag{3.9}$$

where J is the current density in the conductors, and where I have made the optimistic assumption that the good-field region is the same as the half-width of the coil window. The thickness of the current sheets is

$$t = \frac{5}{2} \frac{g h_u}{2fJ\pi - 5g}. \tag{3.10}$$

where

$$J > \frac{5}{2} \frac{g}{f\pi} \tag{3.11}$$

must hold in order to obtain a meaningful solution. Taking J as a free parameter of the design, the minimum power consumption is obtained when J takes the value

$$J_{PQ,opt} = 5\frac{g}{f\pi}, \tag{3.12}$$

for which the power is

$$P_{PQ,min} = \frac{50 g^2 G h_u \rho (D-G)}{f\pi^2}. \tag{3.13}$$

For an asymmetric quadrupole design, the power consumed is

$$P_{AQ} = \frac{5}{\pi} K_1 G g h_u \rho J (D - K_2 G), \tag{3.14}$$

where K_1 and K_2 are constants that give, respectively, the ratios of the maximum x extent of the pole and the pole-root-width to the good-field region. For the SSRL alpha-magnet, we have $K_1 \approx 1.3$ and $K_2 \approx 1$. Note that the power consumption of the asymmetric quadrupole can be decreased indefinitely by decreasing J (which is not the same as the current density in the Panofsky quadrupole), at the expense of larger coils; obviously, this is limited by practical considerations such as the cost of materials, water pressure drop, etc.

If one takes the ratio of P_{AQ} to $P_{PQ,min}$, one obtains

$$\frac{P_{AQ}}{P_{PQ,min}} = \frac{\pi K_1 f (D - K_2 G)}{10(D-G)} \frac{J}{g} \approx \frac{J}{3g}. \tag{3.15}$$

where I ignore factors of order unity in making the approximation. The maximum gradient desired in the SSRL application was 350 G/cm^2. Hence, the Panofsky quadrupole would have used more power unless the current density for the asymmetric quadrupole were above about 1000 A/cm^2. In fact, the coils in the magnet could be made large enough to achieve J $\leq$ 175A/cm^2, from which one can conclude that a comparable Panofsky quadrupole would consume about six times as much power as the design used.

3.2 Particle Motion in the Alpha-Magnet

3.2.1 Scaled Equation of Motion

Particle motion in the alpha-magnet is best described with the aid of a diagram such as Figure 3.5, which shows the central particle trajectory and the coordinate system. In terms of these coordinates, the magnetic field for $q_1 > 0$ is

$$\mathbf{B} = g(q_3\mathbf{q}_1 + q_1\mathbf{q}_3). \tag{3.16}$$

where the constant g is the alpha-magnet gradient. The equation of motion is obtained from the Lorentz force

$$\mathbf{F} = -e\mathbf{E} - \frac{q}{c}\mathbf{v} \times \mathbf{B}, \tag{3.17}$$

with $\mathbf{E} = 0$, and is

$$\frac{d\gamma\mathbf{v}}{dt} = -\frac{e}{m_e c}\mathbf{v} \times \mathbf{B}. \tag{3.18}$$

Since the magnetic field does no work, γ is constant and can be taken outside the derivative. Since the magnitude of the velocity is also constant, one can rewrite the derivatives as derivatives with respect to path-length, $s = \beta ct$, instead of time. Combining these, one obtains

$$\frac{d^2\mathbf{q}}{ds^2} = -\frac{e}{m_e c^2 \beta\gamma}\frac{d\mathbf{q}}{ds} \times \mathbf{B}. \tag{3.19}$$

I now define a constant α by

$$\alpha^2 = \frac{eg}{m_e c^2 \beta\gamma}, \tag{3.20}$$

or, in more practical units

$$\alpha^2 = 5.86674 \times 10^{-4}\text{cm}^{-2}\frac{g(\text{G/cm})}{\beta\gamma}. \tag{3.21}$$

The equation of motion becomes

$$\frac{d^2\mathbf{q}}{ds^2} = -\alpha^2\frac{d\mathbf{q}}{ds} \times \frac{\mathbf{B}(\mathbf{q})}{g} \tag{3.22}$$

$$= -\alpha^2\frac{d\mathbf{q}}{ds} \times (q_3, 0, q_1) \tag{3.23}$$

$$= -\alpha^2\left(\frac{dq_2}{ds}q_1, \frac{dq_3}{ds}q_3 - \frac{dq_1}{ds}q_1, -\frac{dq_2}{ds}q_3\right) \tag{3.24}$$

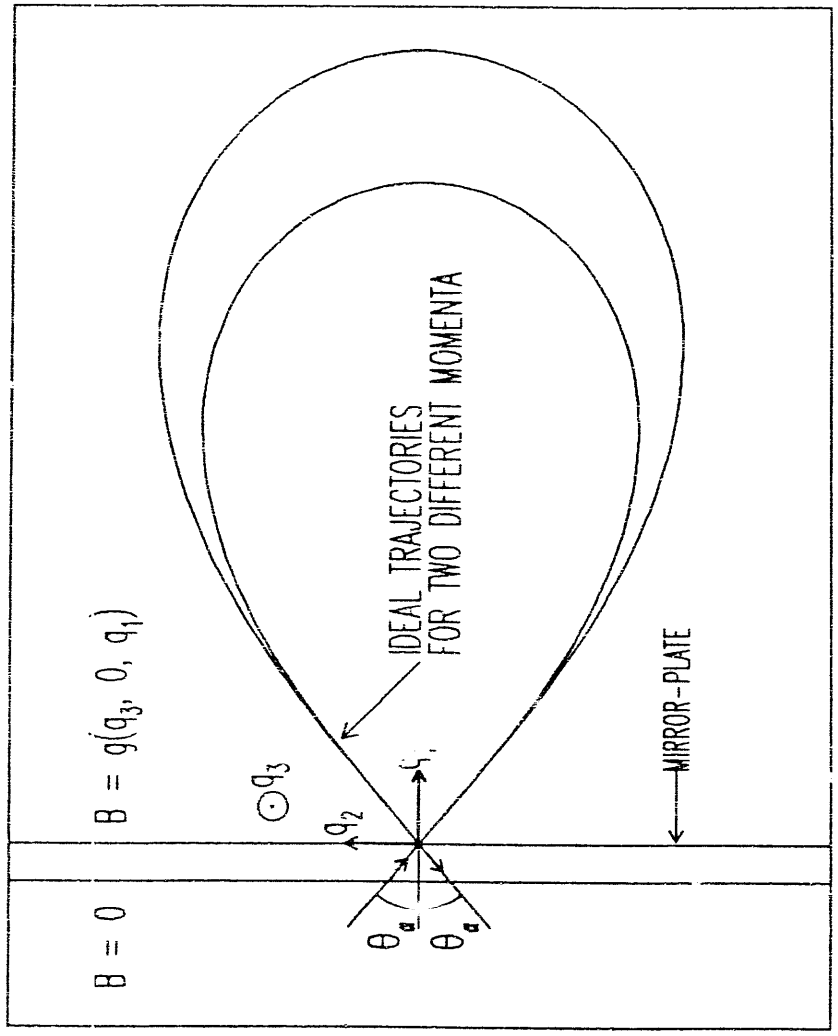

Figure 3.5: Alpha-magnet coordinate system

I chose by convention to make $g > 0$, i.e., I define the e_3 axis to obtain $B_3 > 0$ inside the magnet. This also ensures that $\alpha^2 > 0$, so that α is real and positive. To obtain an α-like trajectory like that exhibited in Figure 3.5, it is then necessary to have initial velocities such that

$$\frac{dq_1}{ds} > 0 \quad \text{and} \quad \frac{dq_2}{ds} < 0. \tag{3.25}$$

I wish to rewrite this equation of motion once more, in such a way as to scale out all explicit dependence on g and $\beta\gamma$. To do this, I define scaled coordinates $\mathbf{Q} = \mathbf{q}\alpha$ and scaled path-length $S = s\alpha$. Using this, I obtain

$$\frac{d^2\mathbf{Q}}{dS^2} = -\frac{d\mathbf{Q}}{dS} \times \mathbf{B}(\frac{\mathbf{Q}}{\alpha})\frac{\alpha}{g} \tag{3.26}$$

$$= -\frac{d\mathbf{Q}}{dS} \times (Q_3, 0, Q_1) \tag{3.27}$$

$$= -\left(\frac{dQ_2}{dS}Q_1, \frac{dQ_3}{dS}Q_3 - \frac{dQ_1}{dS}Q_1, -\frac{dQ_2}{dS}Q_3\right) \tag{3.28}$$

Note that

$$\left(\frac{dQ_1}{dS}\right)^2 - \left(\frac{dQ_2}{dS}\right)^2 - \left(\frac{dQ_3}{dS}\right)^2 = 1, \tag{3.29}$$

a result which will be useful latter, and which in fact does not depend on the scaling (it is true of $\frac{dq}{ds}$ as well).

3.2.2 Ideal Trajectory

From this result, one can deduce that an alpha-magnet can act like an achromatic magnetic mirror, that is, that a zero-emittance beam injected at a specific angle θ_α, to the normal into a perfect alpha-magnet will emerge at the point of injection, at the same angle to the normal and undispersed in momentum.

To see this, first note that the scaled form of the equation of motion does not display any dependence on momentum. Hence, the trajectories of particles with various momenta injected into the magnet at the same angle are simply magnifications or demagnifications of one another. Since the scaled equation likewise does not exhibit any dependence on gradient, the same can be said of particles injected into alpha magnets with differing gradients. Because the scaling involves all coordinates, it

leaves angles unchanged. Hence, if a closed, α-like trajectory does exist, it has the same shape and the same incident and final angles for all values of α (i.e., for particles of all momenta in alpha magnets of all gradients).

Note that the scaling alone is not sufficient to ensure that the magnet can be operated as an achromat. It is also necessary that a trajectory exists which exits at the injection point, since otherwise the scaling would change the exit location relative to the injection point. This would, of course, imply non-zero dispersion upon exiting the magnet.

Next, set $Q_3 = 0$ and note that for trajectories started at $Q_2 = 0$ with $\frac{dQ_1}{dS} = 0$ (implying $\frac{dQ_2}{dS} = 1$) there is some initial value, $\hat{Q}_1$, of Q_1 that results in a trajectory that crosses $Q_1 = Q_2 = 0$. To see that this must be so, imagine starting trajectories from $Q_2 = 0$ at various initial values of Q_1. A trajectory started at infinitesimally small $Q_1 > 0$ will cross $Q_1 = 0$ at infinity, since it "sees" very little magnetic field, and hence is bent toward $Q_1 = 0$ only very gradually. As the starting Q_1 is increased, the trajectory crosses $Q_1 = 0$ at less and less positive values of Q_2, until eventually, for initial $Q_1 = \hat{Q}_1$, the trajectory crosses $Q_1 = 0$ at $Q_2 = 0$.

I will denote this trajectory by $\overline{\mathbf{Q}}(S) = (\overline{Q}_1(S), \overline{Q}_2(S), 0)$, and let $S = 0$ at the start of the trajectory, which is formally defined only for $S > 0$. By construction, $\overline{\mathbf{Q}}(S)$ is a solution to the equations of motion. Consider a new trajectory $\tilde{\mathbf{Q}}(S)$ defined for $S < 0$ as $(\overline{Q}_1(-S), -\overline{Q}_2(-S), 0)$. Upon inserting this trajectory into the equation of motion (with $Q_3 = \frac{dQ_3}{dS} = 0$), one obtains for the left-hand side of equation (3.26), for component 1:

$$\frac{d^2\tilde{Q}_1(S)}{dS^2} = \frac{d^2\overline{Q}_1(-S)}{dS^2} \tag{3.30}$$

$$= \frac{d^2\overline{Q}_1(-S)}{d(-S)^2}\left(\frac{d(-S)}{dS}\right)^2 \tag{3.31}$$

$$= \left(\frac{d^2\overline{Q}_1(S)}{dS^2}\right)_{(S\to -S)} \tag{3.32}$$

$$\tag{3.33}$$

Similarly, for component 2, one obtains

$$\frac{d^2\tilde{Q}_2(S)}{dS^2} = -\left(\frac{d^2\overline{Q}_1(S)}{dS^2}\right)_{(S\to -S)}. \tag{3.34}$$

For the right-hand-side of equation (3.28), one obtains for components 1 and 2, respectively:

$$-\left(\frac{d}{dS}(-\overline{Q}_2(-S))\right)\overline{Q}_1(-S) = \left(-\frac{d\overline{Q}_2(S)}{dS}\overline{Q}_1(S)\right)_{(S\to -S)} \tag{3.35}$$

$$\left(\frac{d}{dS}\overline{Q}_1(-S)\right)\overline{Q}_1(-S) = -\left(\frac{d\overline{Q}_1(S)}{dS}\overline{Q}_1(S)\right)_{(S\to -S)} \tag{3.36}$$

$$\tag{3.37}$$

Combining these last results, one sees that except for the change of variable S to -S, the resultant equations are just those that would be obtained by inserting $\overline{\mathbf{Q}}$ into the equation of motion. Hence, $\tilde{\mathbf{Q}}$ is a solution to the equation of motion, since $\overline{\mathbf{Q}}$ is. Further, the trajectory $\mathbf{Q}_\alpha(S)$, defined by joining $\overline{\mathbf{Q}}$ to $\tilde{\mathbf{Q}}$ at S=0, is also a solution. The subscript α is used from here on to represent properties of the solution $\mathbf{Q}_\alpha(S)$, which is the "α-shaped" trajectory. There should be no confusion with the scaling parameter α, defined by equation (3.20), since the later is not used as a subscript.

A trajectory has thus been demonstrated to exist which starts at $Q_1 = Q_2 = 0$ with such values of $\frac{dQ_1}{dS}$ and $\frac{dQ_2}{dS}$ so as reach $Q_1 = \hat{Q}_1$ and $Q_2 = 0$ with $\frac{dQ_1}{dS} = 0$, and which continues in a mirror symmetric fashion, crossing $Q_1 = Q_2 = 0$ with the negative of the slope with which it started. The absolute value of this slope is denoted by $\tan(\theta_\alpha)$.

Corresponding to $\mathbf{Q}_\alpha(S)$ is an alpha-shaped trajectory for any gradient and particle momentum. These trajectories enter and exit at the angle θ_α, since slopes are not changed by the coordinate scaling.

3.2.3 Numerical Solution of the Equations

It is possible to solve for $\mathbf{Q}_\alpha(S)$ in terms of elliptic integrals[32]. However, this is unproductive, since in the end one obtains a result that can only be used by consulting numerical tables or doing numerical integration. It is better to go directly to numerical integration, especially since the scaled form of the equation allows one to apply the results of a single numerical integration to an infinite number of combinations of $\beta\gamma$ and g.

In order to find the angle θ_α and the maximum value of Q_1 for the trajectory $\mathbf{Q}_\alpha(S)$, I used numerical integration starting at $Q_1 = Q_2 = 0$ and searched for the

value of θ_α that resulted in $\frac{dQ_1(S)}{dS} = 0$ when the trajectory crosses the $Q_2 = 0$ axis again. To gauge the accuracy of the numerical integration, note that at that

$$\left|\frac{dQ}{dS} - 1\right| < 5 \times 10^{-16} \tag{3.38}$$

$$\left(\frac{dQ_1}{dQ_2}\right)_{\text{midplane}} < 5 \times 10^{-15}, \tag{3.39}$$

where the average are taken over the entire integration, which shows that the integration is accurate to 14 decimal places. The Bulirsch-Stoer integration method was employed [61]. Briefly, Bulirsch-Stoer uses the modified midpoint method with polynomial extrapolation of the solution to zero step-size, along with adaptive step-size control.

In this fashion, I obtained

$$\theta_\alpha = 0.71052198004575 \tag{3.40}$$

$$= 40.709910707900^\circ \tag{3.41}$$

$$S_\alpha = 4.64209946506084 \tag{3.42}$$

$$\hat{Q}_1 = 1.81781711509708 \tag{3.43}$$

θ_α is the injection angle for achromatic mirror operation, i.e., the injection angle that results in the trajectory $\mathbf{Q}_\alpha(S)$. S_α is the path length of $\mathbf{Q}_\alpha(S)$ through the entire magnet. $\hat{Q}_1$ is the maximum value of Q_1 reached by $\mathbf{Q}_\alpha(S)$. These quantities are illustrated in Figure 3.6.

3.2.4 Dispersion and Achromatic Path-Length

While these results are not sufficient to fully characterize the optics of the alpha-magnet (see the next section for this), they do allow one to deduce some of the magnet's most important optical properties, namely the dispersion at the vertical midplane and the dependence of path-length on momentum. For this, I revert to unscaled coordinates, and write

$$s_\alpha(\alpha) = \frac{S_\alpha}{\alpha} \tag{3.44}$$

$$\hat{q}_1(\alpha) = \frac{\hat{Q}_1}{\alpha}. \tag{3.45}$$

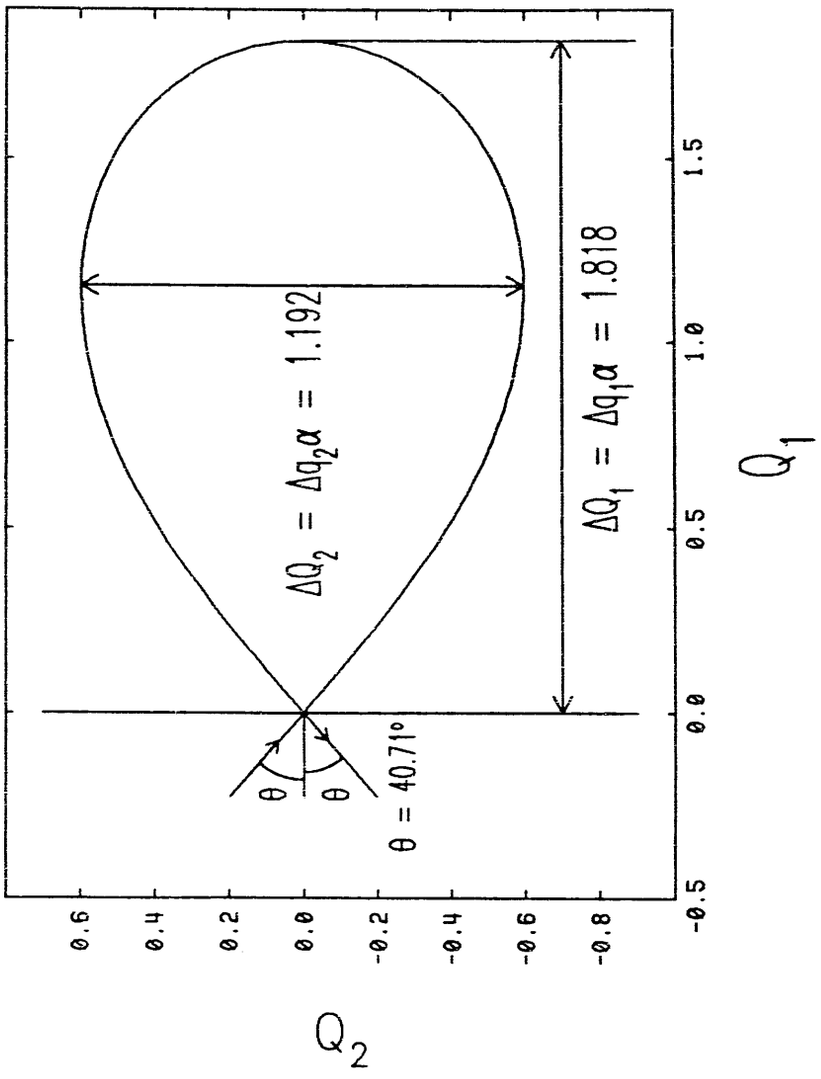

Figure 3.6: Ideal Trajectory in the Alpha-Magnet

In more practical units, and using the numerical values of S_o and $\dot{Q}_1$ given above

$$s_o(\text{cm}) = 191.655\sqrt{\frac{\mathcal{B}\gamma}{g(\text{G/cm})}} \tag{3.46}$$

$$\dot{q}_1(\text{cm}) = 75.0513\sqrt{\frac{\mathcal{B}\gamma}{g(\text{G/cm})}} \tag{3.47}$$

Assuming that the gradient g is fixed, and letting a_o be the value of a for the central particle, of momentum $p_o = (\mathcal{B}\gamma)_o$, the previous equations imply that

$$s(a) = s(a_o)\frac{a_o}{a} \tag{3.48}$$

$$\dot{q}_1(a) = \dot{q}_1(a_o)\frac{a_o}{a}. \tag{3.49}$$

Expanding in $\delta = (p - p_o)/p$, one obtains

$$\frac{a_o}{a} = \sqrt{\frac{p}{p_o}} \tag{3.50}$$

$$\approx 1 + \frac{1}{2}\delta - \frac{1}{8}\delta^2 + \frac{1}{16}\delta^3. \tag{3.51}$$

Using this expansion the dispersive terms of the transport matrix (see the next section) from the entrance of the magnet to the "vertical midplane" (where the ideal trajectory crosses $q_2 = 0$ with $q_1 = \dot{q}_1$) are seen to be

$$r_{16} \equiv -\left(\frac{\partial \dot{q}_1}{\partial \delta}\right)_{\delta=0} \tag{3.52}$$

$$= -\frac{1}{2}\dot{q}_1(a_o) \tag{3.53}$$

$$t_{166} \equiv -\frac{1}{2!}\left(\frac{\partial^2 \dot{q}_1}{\partial \delta^2}\right)_{\delta=0} \tag{3.54}$$

$$= \frac{1}{8}\dot{q}_1(a_o) \tag{3.55}$$

$$u_{1666} \equiv -\frac{1}{3!}\left(\frac{\partial^3 \dot{q}_1}{\partial \delta^3}\right)_{\delta=0} \tag{3.56}$$

$$= -\frac{1}{16}\dot{q}_1(a_o) \tag{3.57}$$

Similarly, the path-length terms for transport through the entire magnet are

$$r_{56} = \frac{1}{2}s(a_o) \tag{3.58}$$

$$t_{566} = -\frac{1}{8}s(a_o) \tag{3.59}$$

$$u_{5666} = \frac{1}{16}s(a_o) \tag{3.60}$$

These will prove useful in checking the results of detailed transport matrix calculations. They are also of interest because the dispersion at the vertical midplane and the momentum-dependence of the path-length are two of the alpha-magnet's most useful features. The dispersion at the vertical midplane allows for momentum selection via a slit or scraper placed at the vertical midplane. The momentum-dependence of the path-length is, of course, necessary for bunch compression, as indicated in the introduction to this chapter.

3.3 Alpha-Magnet Transport Matrix Scaling

In this section I derive results that provide the basis for a calculation of alpha-magnet transport matrices to third order. Transport matrices express particle motion between two points in a beamline as a series expansion about the trajectory of a hypothetical particle that travels along what is considered to be the ideal trajectory for the beamline. Typically this ideal trajectory passes through the center of focusing elements, down the center of the beam-pipe, and so forth. In the case of the alpha-magnet, the ideal trajectory enters and exits at the angle θ_α, with $q_3 = \frac{dq_3}{ds} = 0$.

3.3.1 Curvilinear Coordinates and Matrix Notation

The coordinates used for the transport matrix expansion[10] specify offsets in six-dimensional phase-space of a particle from the ideal trajectory. The coordinate system is curvilinear, i.e., it follows the ideal trajectory. This subject is treated completely in publications on particle beam dynamics, listed in the references. Here, I will simply state that the position of any particle relative to the fiducial particle can be specified in terms of two transverse coordinates, x and y, their derivatives with respect to path length (s_c) for the central trajectory,

$$x' = \frac{dx}{ds_c} \qquad y' = \frac{dy}{ds_c}, \tag{3.61}$$

the longitudinal distance s traveled, and the momentum deviation δ, introduced in the last section. As is usually done, I form a six-dimensional vector from these coordinates:

$$\mathbf{x} = \begin{pmatrix} x \\ x' \\ y \\ y' \\ s \\ \delta \end{pmatrix}. \tag{3.62}$$

This vector gives information about a particle as it crosses a reference plane somewhere in the beamline. The reference plane is constructed so that the fiducial particle

passes through it perpendicularly. I emphasize that the path-length s is *not* the distance of a particle behind the fiducial particle; I depart from convention here in keeping track of the total path-length, for reasons that will be apparent later. This carries no penalty for a beamline composed of static elements, since the expansions in $s - s_o$ are then of no importance.

Transformation of this vector by beamline elements is expressed as a series expansion:

$$x_i \longrightarrow c_i + \sum_j r_{ij}x_j + \sum_{j\geq k} t_{ijk}x_jx_k + \sum_{j\geq k\geq l} u_{ijkl}x_jx_kx_l, \tag{3.63}$$

where c, r, t, and u are the transport matrices for some element, and summation indices run from 1 to 6 unless otherwise indicated. (The reason for the lower-case letters will be seen presently.) The restricted sums are used to obtain expressions that contain only one instance of any term x_jx_k or $x_jx_kx_l$. This is consistent with K.Brown[10], but differs from the definition used by TRANSPORT[68] and some other computer programs, where the matrices are defined in terms of symmetric sums over all indices. The unsymmetric form also has advantages in a computer program, namely reduction of storage used and reduction of the number of arithmetic operations needed to transform particle coordinates. I employ the unsymmetric form exclusively in this work.

The element c is unconventional, and is used to keep track of centroid offsets. It finds application in three ways. First, when used in a tracking program, associating a centroid offset matrix with an element allows one to implement beam misalignments and steering in a straight-forward fashion. In addition, time-of-flight calculations are facilitated by the path-length centroid element, which is useful in a simulation that has time-dependent elements [49]. Second, it is a necessary corrolary of my use of total path-length instead of differential path-length in the vector **x**. Third, in the particular case of the alpha-magnet, the centroid matrix can be used to calculate higher-order dispersive path-length terms, as will be seen below. For the alpha-magnet and all other elements that do not produce orbit distortions, only the c_5 element is non-zero.

3.3.2 Relationships Between Curvilinear and Fixed Coordinates

At this point, the reader might expect the equation of motion to be rewritten in terms of the curvilinear coordinates. This is unnecessary for my purposes. All that I will need in order calculate the matrices (c, r, t, u) is to express the relationship between the curvilinear coordinates $\mathbf{x}$ and the coordinates of the equation of motion, $\mathbf{q}$, at the entrance, vertical midplane, and exit of the alpha-magnet, since it is between these reference planes that I wish to know the transport matrices.

At the entrance of the alpha-magnet (i.e., when the particle crosses the reference plane shown in Figure 3.7), the correspondence between $\mathbf{x}$ and $\mathbf{q}$ is given by

$$\begin{pmatrix} x \\ x' \\ y \\ y' \\ \tau \\ \delta \end{pmatrix} = \begin{pmatrix} \mathrm{sign}(q_1)\sqrt{q_1^2+q_2^2} \\ \tan(\mathrm{atan}(-q_1'/q_2') - \bar{\theta}_\alpha) \\ q_3 \\ q_3' \\ s \\ (p-p_o)/p_o \end{pmatrix}, \tag{3.64}$$

where I have used

$$\bar{\theta}_\alpha \equiv \frac{\pi}{2} - \theta_\alpha. \tag{3.65}$$

The slopes q_1' and q_3' are given by

$$q_1' = \sqrt{1-(q_3')^2}\,\sin\left(\bar{\theta}_\alpha - \mathrm{atan}(x')\right) \tag{3.66}$$

and

$$q_2' = \sqrt{1-(q_3')^2}\,\cos\left(\bar{\theta}_\alpha - \mathrm{atan}(x')\right). \tag{3.67}$$

while the coordinates q_1 and q_2 are given by

$$q_1 = x\sin\theta_\alpha \tag{3.68}$$

and

$$q_2 = x\cos\theta_\alpha. \tag{3.69}$$

The reader may have noticed that the reference plane in Figure 3.7 is partially inside and partially outside the alpha-magnet. Hence, it would seem that in reaching

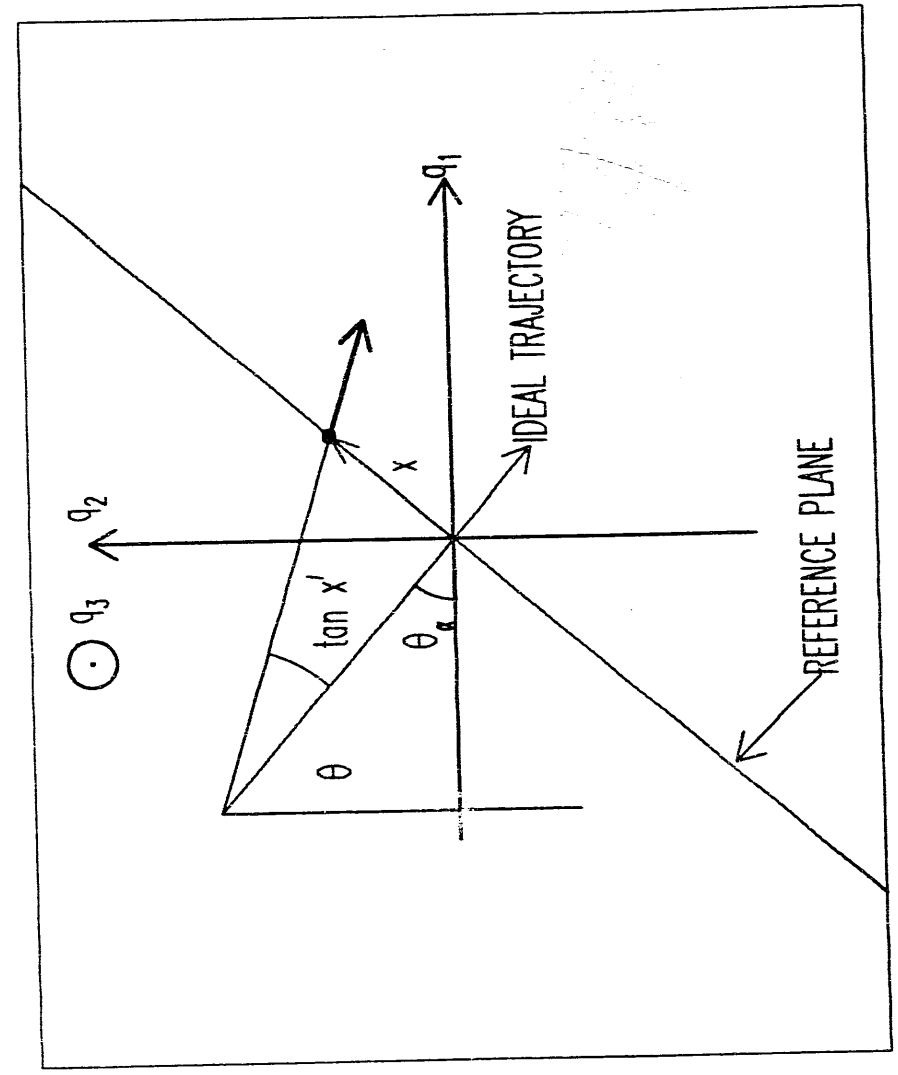

Figure 3.7: Reference plane and coordinates at the entrance

the reference plane, from which transport through the alpha-magnet nominally starts in the transport matrix formalism, some particles have already traversed part of the alpha-magnet's magnetic field. Others (those for which $x < 0$ in figure 3.7), will not yet be inside the alpha-magnet. It would seem that the length of a drift space, for example, prior to the alpha magnet would need to be modified according to the coordinates of the particle, and this is effectively what is done. The prior element in the transport line (presumably a drift space) is considered to deliver all of the particles to the reference plane, with no account taken of the alpha-magnet fields. The computation of the alpha-magnet matrices (see the subsequent sections of this chapter) takes this into account, so that particles that are delivered inside (outside) the alpha-magnet are drifted backward (forward) to the field boundary of the alpha-magnet before numerical integration. As will be seen presently, similar considerations apply at the exit of the alpha-magnet, and an identical procedure is followed for this case.

One could also consider constructing an edge-matrix for the alpha-magnet, similar to what is done for bending magnets, but since the entrance and exit angles for the alpha-magnet do not vary between applications (as they do for bending magnets), this is neither necessary nor useful.

At the vertical midplane of the magnet (i.e., when the particle crosses $q_2 = 0$ inside the magnet, see Figure 3.8), a different relationship holds:

$$\begin{pmatrix} x \\ x' \\ y \\ y' \\ \tau \\ \delta \end{pmatrix} = \begin{pmatrix} \hat{q}_1 - q_1 \\ -q_1'/q_2' \\ q_3 \\ q_3' \\ s \\ (p - p_o)/p \end{pmatrix}. \tag{3.70}$$

The slopes q_1' and q_3' are given by

$$q_1' = -\sqrt{1 - (q_3')^2}\sin(\text{atan}(x')) \tag{3.71}$$

and

$$q_2' = \sqrt{1 - (q_3')^2}\cos(\text{atan}(x')), \tag{3.72}$$

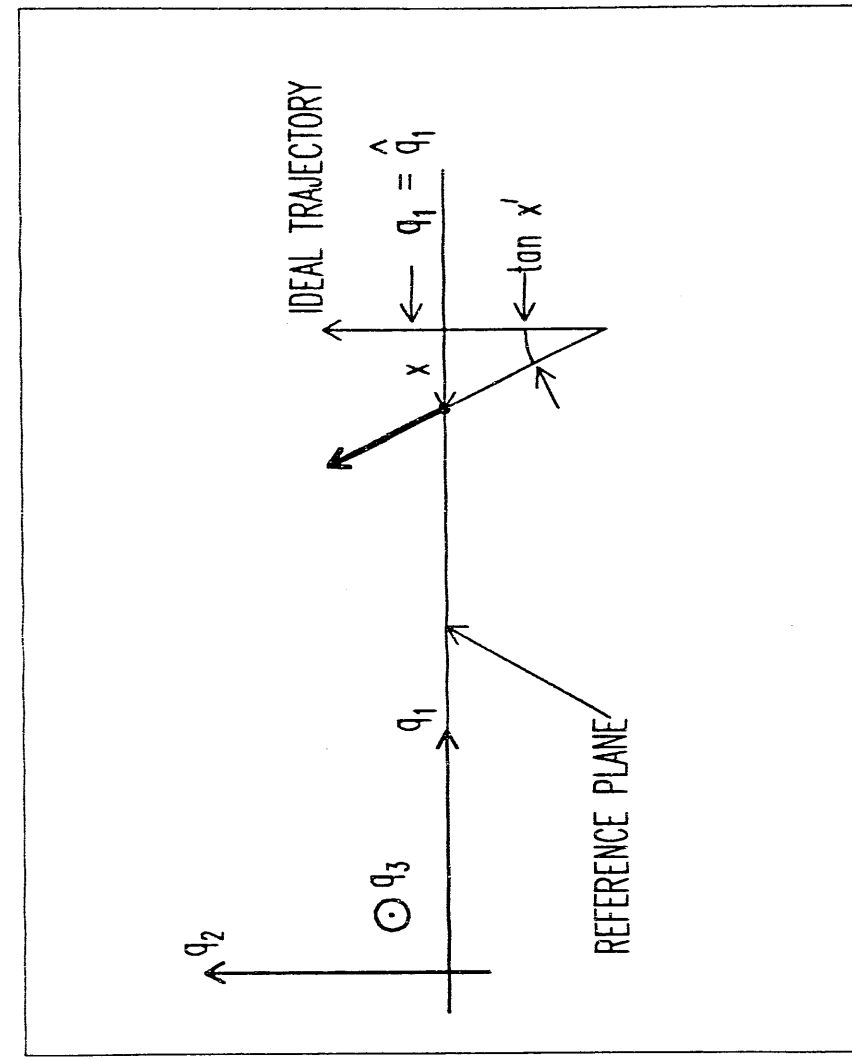

Figure 3.8: Reference plane and coordinates at the vertical midplane

while the coordinates q_1 and q_2 are given by

$$q_1 = \hat{q}_1 - x \tag{3.73}$$

and

$$q_2 = 0. \tag{3.74}$$

Finally, at the exit of the magnet (i.e., when the particle crosses the reference plane shown in Figure 3.9), one obtains:

$$\begin{pmatrix} x \\ x' \\ y \\ y' \\ \tau \\ \delta \end{pmatrix} = \begin{pmatrix} \mathrm{sign}(q_1)\sqrt{q_1^2 + q_2^2} \\ \tan(\bar{\theta}_\alpha - \mathrm{atan}(q_1'/q_2')) \\ q_3 \\ q_3' \\ s \\ (p - p_o)/p \end{pmatrix}. \tag{3.75}$$

The slopes q_1' and q_3' are given by

$$q_1' = -\sqrt{1 - (q_3')^2}\sin(\bar{\theta}_\alpha - \mathrm{atan}(x')) \tag{3.76}$$

and

$$q_2' = -\sqrt{1 - (q_3')^2}\cos(\bar{\theta}_\alpha - \mathrm{atan}(x')), \tag{3.77}$$

while the coordinates q_1 and q_2 are given by

$$q_1 = x\sin\theta_\alpha \tag{3.78}$$

and

$$q_2 = -x\cos\theta_\alpha. \tag{3.79}$$

3.3.3 Coordinate Scaling

Let the gradient in the alpha-magnet and the momentum of the fiducial particle be specified, so that the scaling parameter a takes a definite value, a_o. Then it is possible to define a new vector $\mathbf{X}$ that has the same relationship to $\mathbf{Q}$ that $\mathbf{x}$ has to $\mathbf{q}$. $\mathbf{X}$ is obtained from $\mathbf{x}$ by the transformation

$$\mathbf{X} = \mathbf{A}(a_o)\cdot\mathbf{x}, \tag{3.80}$$

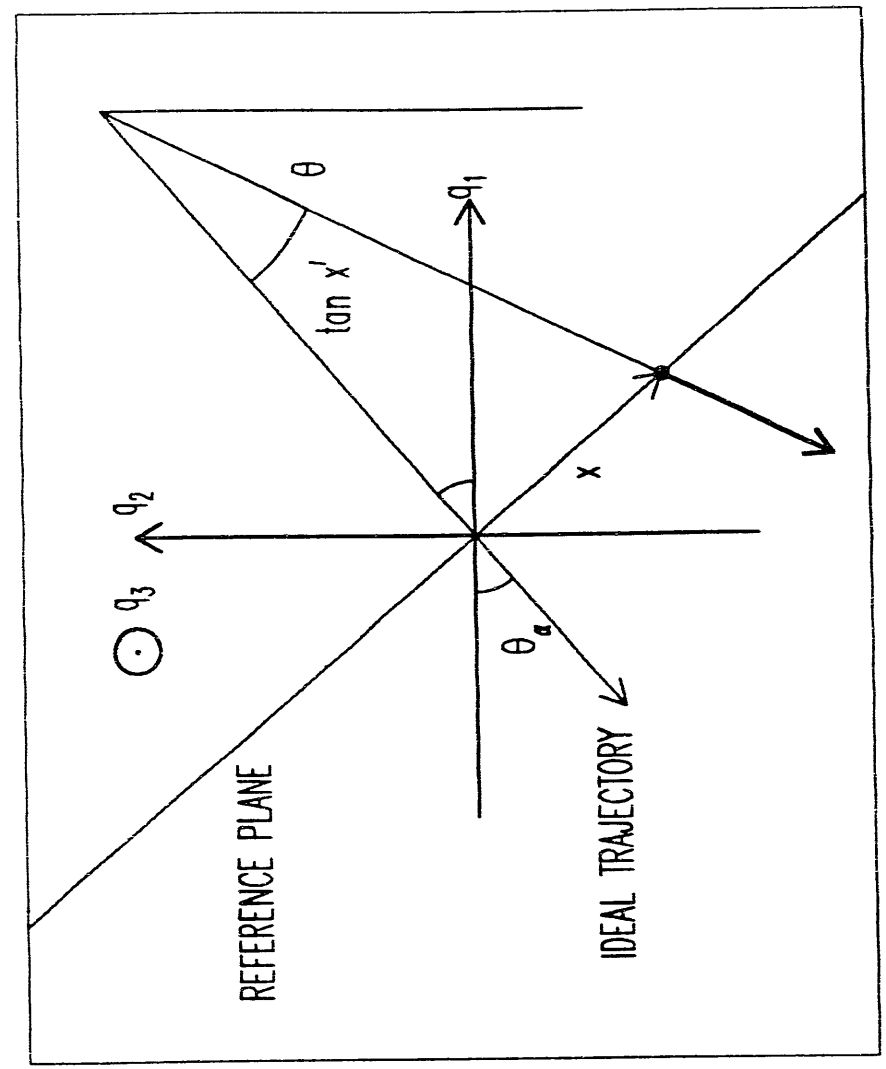

Figure 3.9: Reference plane and coordinates at the exit

where $\mathbf{A}(\alpha_o)$ is a diagonal matrix, given by

$$\mathbf{A}(\alpha_o) = \begin{pmatrix} \alpha_o & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_o & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_o & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.81}$$

The transformation from $\mathbf{x}$ to $\mathbf{X}$ transforms the fiducial particle, which traveled a particular α-like trajectory $\mathbf{Q}_\alpha(s\alpha_o)/\alpha_o$, into the particle that follows the universal trajectory $\mathbf{Q}_\alpha(S)$. To see this more clearly, note that the expression for $\mathbf{X}$ at the vertical midplane is

$$\begin{pmatrix} X \\ X' \\ Y \\ Y' \\ S \\ \delta \end{pmatrix} = \begin{pmatrix} \hat{Q}_1 - Q_1 \\ -\mathrm{atan}(Q_1'/Q_2') \\ Q_3 \\ Q_3' \\ S \\ (p - p_o)/p \end{pmatrix}, \tag{3.82}$$

where $\mathbf{Q}' \equiv \frac{d\mathbf{Q}}{dS}$. (Since angles are unchanged by the scaling, I am free to express the slopes in terms of either the Q_i''s or the q_i''s, even though this "transformation" is not in the matrix $\mathbf{A}$.)

3.3.4 Scaled Equation of Motion with Dispersive Terms

The reader may have noted an apparent inconsistency here: this vector, which is in scaled coordinates, refers to the momentum error, but the scaling was explicitly constructed so as to remove all reference to momentum. The apparent inconsistency stems from the fact that, as developed in the last section, the scaled equation of motion treats every particle (each characterized by some particular scaling constant α) as the fiducial particle (at least as far as momentum is concerned). What is needed to incorporate momentum errors into the scaled equation of motion is to realize that one scales the equation with α_o, the α value for the central momentum, and not with

the particular α of every particle under consideration. The resultant scaled equation, with dispersive effects included exactly, is

$$\mathbf{Q}'' = -\frac{1}{1+\delta}\mathbf{Q}' \times \mathbf{B}(\frac{\mathbf{Q}}{\alpha})\frac{\alpha}{g} \tag{3.83}$$

$$= -\frac{1}{1+\delta}\frac{d\mathbf{Q}}{dS} \times (Q_3, 0, Q_1) \tag{3.84}$$

As foreshadowed at the end of the last section, it is not entirely necessary to include dispersive effects in this fashion. One can obtain all dispersive terms in the matrices by taking derivatives with respect to δ after reverting to unscaled coordinates, though this requires some care if it is to be done correctly. This will be discussed in more detail below. One reason for inserting dispersive effects at this point is to retain the six-dimensional transport matrix formalism. Another reason, as indicated at the end of the last section and as will become more apparent below, is that putting dispersive effects into the formalism provides a check on the calculation of the matrix.

3.3.5 Scaling of the Transport Matrices

One can define transformation matrices for the vector $\mathbf{X}$, with the realization that these transformation matrices apply to the scaled form of the equation of motion:

$$X_I = C_I + \sum_I R_{IJ}X_J + \sum_{J\geq K} T_{IJK}X_JX_K + \sum_{J\geq K\geq L} U_{IJKL}X_JX_KX_L. \tag{3.85}$$

If I now substitute into this relation the definition of $\mathbf{X}$, equation (3.80), I obtain

$$\sum_m A_{Im}x_m = C_I + \sum_{Jj} R_{IJ}A_{Jj}x_j + \sum_{J\geq K}\sum_{jk} T_{IJK}A_{Jj}x_jA_{Kk}x_k + \tag{3.86}$$

$$\sum_{J\geq K\geq L}\sum_{jkl} U_{IJKL}A_{Jj}x_jA_{Kk}x_kA_{Ll}x_l. \tag{3.87}$$

Multiplying from the left by A_{iI}^{-1} and summing over I yields

$$x_i = \sum_I A_{iI}^{-1}C_I + \sum_{Ij} A_{iI}^{-1}R_{IJ}A_{Jj}x_j + \sum_{J\geq K}\sum_{Ijk} A_{iI}^{-1}T_{IJK}A_{Jj}x_jA_{Kk}x_k +$$

$$\sum_{J\geq K\geq L}\sum_{Ijkl} A_{iI}^{-1}U_{IJKL}A_{Jj}x_jA_{Kk}x_kA_{Ll}x_l. \tag{3.88}$$

Using the fact that **A** is a diagonal matrix, this becomes

$$x_i \longrightarrow A_{ii}^{-1}C_i + \sum_j A_{ii}^{-1}R_{ij}A_{jj}x_j + \sum_{j\geq k} A_{ii}^{-1}T_{ijk}A_{jj}x_jA_{kk}x_k + \sum_{j\geq k\geq l} A_{ii}^{-1}U_{ijkl}A_{jj}x_jA_{kk}x_kA_{ll}x_l. \tag{3.89}$$

Comparison with the definition of the matrices (c, r, t, u), equation (3.63), for the normal coordinates gives

$$c_i = A_{ii}^{-1}(\alpha_o)C_i \tag{3.90}$$

$$r_{ij} = A_{ii}^{-1}(\alpha_o)R_{ij}A_{jj}(\alpha_o) \tag{3.91}$$

$$t_{ijk} = A_{ii}^{-1}(\alpha_o)T_{ijk}A_{jj}(\alpha_o)A_{kk}(\alpha_o) \tag{3.92}$$

$$u_{ijkl} = A_{ii}^{-1}(\alpha_o)U_{ijkl}A_{jj}(\alpha_o)A_{kk}(\alpha_o)A_{ll}(\alpha_o), \tag{3.93}$$

where there are no sums in these relations, in spite of the many repeated indices. I have reasserted the dependence of **A** on α_o to emphasize it, since the importance of this result stems from this dependence. Specifically, if the matrices (C, R, T, U) for the scaled equation of motion can be found, then this last result allows one to find the matrices (c, r, t, u) for an alpha magnet run at some gradient g and for some central momentum $p_o = (\beta\gamma)_o$ such that

$$\alpha_o = \sqrt{\frac{eg}{m_e c^2 p_o}}. \tag{3.94}$$

A more easily used form of this result can be obtained by noting that

$$A_{ii} = \frac{1}{A_{ii}^{-1}} = \alpha O_i - \bar{O}_i, \tag{3.95}$$

where O_i ($\bar{O}_i$) is 1 (0) if the index i is an odd (even) integer.

The expression for the matrices becomes

$$c_i = \frac{C_i}{O_i\alpha_o + \bar{O}_i} \tag{3.96}$$

$$r_{ij} = \frac{R_{ij}}{O_i\alpha_o + \bar{O}_i}(O_j\alpha_o + \bar{O}_j) \tag{3.97}$$

$$t_{ijk} = \frac{T_{ijk}}{O_i\alpha_o + \bar{O}_i}(O_j\alpha_o + \bar{O}_j)(O_k\alpha_o + \bar{O}_k) \tag{3.98}$$

$$u_{ijkl} = \frac{U_{ijkl}}{O_i\alpha_o + \bar{O}_i}(O_j\alpha_o + \bar{O}_j)(O_k\alpha_o + \bar{O}_k)(O_l\alpha_o + \bar{O}_l). \tag{3.99}$$

From this one can see that c_i elements may be independent of α_o or else inversely proportional to it. r_{ij} may be proportional to $\alpha_o^{-1}, \alpha_o^0$, or α_o^1. t_{ijk} may be proportional to $\alpha_o^{-1}, \alpha_o^0, \alpha_o^1$, or α_o^2. u_{ijkl} may be proportional to $\alpha_o^{-1}, \alpha_o^0, \alpha_o^1, \alpha_o^2$, or α_o^3.

3.3.6 Alternative Treatment of Dispersive Terms

I indicated above that it is not necessary to include dispersive effects in the matrix formalism for the scaled equation. The reason is that in reverting to **x**, one may use α (which is a function of δ, as seen from equation (3.51)) rather than α_o, to obtain the non-dispersive matrix elements as a function of δ. This allows one to calculate the dispersive matrix elements from non-dispersive matrix elements, provided one compensates for the fact that the scaling changes the coordinate system at the vertical midplane as well as the momentum of the particle under consideration. That is, different values of α correspond to different values of $\hat{q}_1$, which enters the definition of the coordinates at the vertical midplane via equation (3.70) so that one cannot simply take derivatives of the non-dispersive matrix elements.

Let $\bar{c}_i(\delta)$, $\bar{r}_{ij}(\delta)$, $\bar{t}_{ijk}(\delta)$, and $\bar{u}_{ijkl}(\delta)$ be the matrices obtained by scaling with α, where δ is defined with respect to α_o by equation (3.51). All chromatic terms $\bar{r}_{i6}$, $\bar{t}_{i6k}$, $\bar{u}_{i6kl}$, and $\bar{u}_{i66l}$ are zero, since the chromatic dependence is now taken care of by the functional form of $\bar{c}_i(\delta)$, $\bar{r}_{ij}(\delta)$, $\bar{t}_{ijk}(\delta)$, and $\bar{u}_{ijkl}(\delta)$. The expression for transformation of a vector **x** into a vector $\bar{\mathbf{x}}$ is now

$$\bar{x}_i = \bar{c}_i(\delta) + \sum_{4\geq j} \bar{r}_{ij}(\delta)x_j + \sum_{4\geq j\geq k} \bar{t}_{ijk}(\delta)x_jx_k + \sum_{4\geq j\geq k\geq l} \bar{u}_{ijkl}(\delta)x_jx_kx_l. \tag{3.100}$$

If the matrices are for the transformation from the entrance to the exit, then there is no modification of the coordinate system with scaling, and no qualifications of this expression are needed. If the matrices are for the transformation from the entrance to the vertical midplane, then the coordinate system with respect to which $\bar{x}_i$ is defined is a function of δ also, and this must be taken into account in interpreting the results, as will be done below. If the matrices are for the transformation from the vertical midplane to the exit, then the coordinate system with respect to which x_i is defined is a function of δ; this case will not be pursued here.

Assuming, then, that the initial coordinates are not dependent (through their

coordinate system or otherwise) on δ, then upon expanding $\tilde{r}_{ij}(\delta)$, $\tilde{t}_{ijk}(\delta)$, and $\tilde{u}_{ijkl}(\delta)$ in δ, one obtains

$$\begin{aligned}\bar{x}_i &= (\tilde{c}_i)_{\delta=0} + \delta\left(\frac{\partial\tilde{c}_i}{\partial\delta}\right)_{\delta=0} + \frac{1}{2!}\delta^2\left(\frac{\partial^2\tilde{c}_i}{\partial\delta^2}\right)_{\delta=0} + \frac{1}{3!}\delta^3\left(\frac{\partial^3\tilde{c}_i}{\partial\delta^3}\right)_{\delta=0} \\ &+ \sum_{4\geq j}\left\{(\tilde{r}_{ij})_{\delta=0} + \delta\left(\frac{\partial\tilde{r}_{ij}}{\partial\delta}\right)_{\delta=0} + \frac{1}{2!}\delta^2\left(\frac{\partial^2\tilde{r}_{ij}}{\partial\delta^2}\right)_{\delta=0}\right\}x_j \\ &- \sum_{4\geq j\geq k}\left\{(\tilde{t}_{ijk})_{\delta=0} + \delta\left(\frac{\partial\tilde{t}_{ijk}}{\partial\delta}\right)_{\delta=0}\right\}x_jx_k - \sum_{4\geq j\geq k\geq l}(\tilde{u}_{ijkl})_{\delta=0}x_jx_kx_l.\end{aligned} \tag{3.101}$$

where I work to third order and where $\bar{x}_i$ may contain effects of coordinate system changes with δ. For transformations from the entrance to the exit, the $\bar{x}_i$ are unaffected by coordinate system changes. For transformations to the vertical midplane, it is only $\bar{x}_1$ that is affected by coordinate system scaling through scaling of $\hat{q}_1$. Hence, I shall momentarily ignore coordinate system dependencies and equate $\bar{x}_i$ with the true coordinates in the proper reference frame. I shall then return to treat the case of $\bar{x}_1$ for transport to the vertical midplane separately.

Taking equation (3.102) literally, then, one can identify the chromatic matrix elements as

$$r_{ij} = \left(\frac{\partial}{\partial\delta}\tilde{c}_i(\alpha)\right)_{\delta=0} \tag{3.102}$$

$$t_{i66} = \frac{1}{2!}\left(\frac{\partial^2}{\partial\delta^2}\tilde{c}_i(\alpha)\right)_{\delta=0} \tag{3.103}$$

$$u_{i666} = \frac{1}{3!}\left(\frac{\partial^3}{\partial\delta^3}\tilde{c}_i(\alpha)\right)_{\delta=0} \tag{3.104}$$

$$t_{ij} = \left(\frac{\partial}{\partial\delta}\tilde{r}_{ij}(\alpha)\right)_{\delta=0} \tag{3.105}$$

$$u_{i66j} = \frac{1}{2!}\left(\frac{\partial^2}{\partial\delta^2}\tilde{r}_{ij}(\alpha)\right)_{\delta=0} \tag{3.106}$$

$$u_{i6kj} = \left(\frac{\partial}{\partial\delta}\tilde{t}_{ijk}(\alpha)\right)_{\delta=0} \tag{3.107}$$

To treat the case of $\bar{x}_1$ for transport to the vertical midplane, I rewrite equation (3.100) for $i = 1$ as

$$\bar{x}_1 = \hat{q}_1(\alpha) - q_1^{\text{midplane}} = \tilde{c}_1(\delta) - \sum_{6>j}\tilde{r}_{1j}(\delta)x_j - \sum_{6>j\geq k}\tilde{t}_{1jk}(\delta)x_jx_k - \sum_{6>j\geq k\geq l}\tilde{u}_{1jkl}(\delta)x_jx_kx_l. \tag{3.108}$$

The actual coordinate of interest is not $\bar{x}_1$ but rather $\bar{x}_1 \equiv \hat{q}_1(\alpha_o) - q_1^{\text{midplane}}$, where q_1^{midplane} is, of course, a function of α. Adding $\hat{q}_1(\alpha_o) - \hat{q}_1(\alpha)$ to both sides of equation (3.108), I obtain

$$\bar{x}_1 = \hat{q}_1(\alpha_o) - \hat{q}_1(\alpha) + \tilde{c}_1(\delta) + \sum_{6>j}\tilde{r}_{1j}(\delta)x_j + \sum_{6>j\geq k}\tilde{t}_{1jk}(\delta)x_jx_k + \sum_{6>j\geq k\geq l}\tilde{u}_{1jkl}(\delta)x_jx_kx_l. \tag{3.109}$$

which, when expanded in δ, yields additional terms not listed in equations (3.102) through (3.107), without modifying those that are listed. These additional matrix elements are none other than those resulting from the expansion of $-\hat{q}_1(\alpha)$, which have already been exhibited in the last section, as equations (3.53) through (3.57).

So far, these results would seem to apply only to the matrices (c, r, t, u) and not to (C, R, T, U). However, if one takes $\alpha_o = 1$, one sees that the matrices (c, r, t, u) are numerically equal to (C, R, T, U), from which it follows that the numerical relationships between the chromatic and non-chromatic elements are the same for (C, R, T, U) as for the (c, r, t, u). Another way of realizing that this is so is to notice that c_i, r_{i6}, t_{i66}, and u_{i666} all have the same scaling with α_o, as do r_{ij}, t_{i6j}, and u_{i66j}, and also t_{ijk} and u_{i6jk}. This can be seen from equations (3.96) through (3.99).

An example may make all this clearer. Consider the element t_{162}, which is given by

$$t_{162}(\alpha_o) = \left(\frac{\partial}{\partial\delta}r_{12}(\alpha)\right)_{(\delta=0)} \tag{3.110}$$

$$= \left(\frac{\partial}{\partial\delta}(r_{12}(\alpha_o)\sqrt{1+\delta})\right)_{(\delta=0)} \tag{3.111}$$

$$= \frac{1}{2}r_{12}(\alpha_o). \tag{3.112}$$

Since

$$T_{162} = \frac{t_{162}(\alpha_o)}{\alpha_o} \tag{3.113}$$

and

$$R_{12} = \frac{r_{12}(\alpha_o)}{\alpha_o}, \tag{3.114}$$

it also follows that

$$T_{162} = \frac{1}{2}R_{12}. \tag{3.115}$$

For rapid checks on calculated matrices, or for inclusion in a computer program, it is convenient to work out the consequences of these relations for the matrices (C, R, T, U). I have done this, and the results are

$$R_{56} = \frac{1}{2}C_5, \tag{3.116}$$

$$T_{566} = -\frac{1}{8}C_5, \tag{3.117}$$

$$U_{5666} = \frac{1}{16}C_5, \tag{3.118}$$

$$T_{16J} = \frac{1}{2}R_{1J}(O_1\bar{O}_J - \bar{O}_1O_J), \tag{3.119}$$

$$U_{166J} = \frac{1}{8}R_{1J}(3\bar{O}_1O_J - O_1\bar{O}_J), \tag{3.120}$$

and

$$U_{16JK} = \frac{1}{2}T_{1JK}\left[C_1\bar{O}_J\bar{O}_K - O_1O_JO_K - \bar{O}_1(O_J\bar{O}_K + \bar{O}_JO_K) - 2\bar{O}_1O_JO_K\right], \tag{3.121}$$

where O_I ($\bar{O}_I$) is 1 if I is odd (even) and zero otherwise, and where $6 > J \geq K$. (I emphasize again that these results are invalid for transport from the vertical midplane to the exit, which is a case I have not treated here.)

3.4 Transport Matrices from Numerical Integration

The scaled transport matrices (C, R, T, U) for the alpha-magnet can be found from numerical integration of the scaled equation of motion (equation (3.83)) and fitting. The technique I have used is not confined in its application to the alpha-magnet, though it is most appropriate for elements for which there exists an equivalent of the scaled equation of motion for the alpha magnet. Essentially, an ensemble of N initial vectors, labeled $\mathbf{X}^{(i)}$, $i = 1, 2, \ldots N$ is mapped into an ensemble of final vectors, labeled $\mathbf{Y}^{(i)}$, by numerical integration starting and ending at the appropriate reference planes. These vectors are then required to satisfy

$$Y_I^{(i)} = C_I + \sum_J R_{IJ}X_J^{(i)} + \sum_{J\geq K} T_{IJK}X_J^{(i)}X_K^{(i)} + \sum_{J\geq K\geq L} U_{IJKL}X_J^{(i)}X_K^{(i)}X_L^{(i)} + \sum_{J\geq K\geq L\geq M} V_{IJKLM}X_J^{(i)}X_K^{(i)}X_L^{(i)}X_M^{(i)} + \mathcal{O}((\mathbf{X}^{(i)})^5), \tag{3.122}$$

which is essentially the definition of the transport matrices, where I've included a fourth-order matrix V. I emphasize that the $\mathbf{Y}^{(i)}$ are not *calculated from* this matrix expression, but are rather being approximated by it, having been calculated by numerical integration of the equation of motion with initial condition $\mathbf{X}^{(i)}$. I am including the fourth-order terms explicitly in order to show how to prevent fourth-order influences from corrupting the computation of (C, R, T, U). The fifth-order terms will be assumed to be negligible.

3.4.1 One-Variable Terms

In principle, one could fit this by finding the (C, R, T, U) that minimized the sum of the squared deviations of the right-hand-side from the left-hand-side. In practice, this is computationally difficult and also extremely inefficient. To see a more efficient procedure, imagine that one was only interested in calculating C_I. Clearly, one would only need to track the fiducial particle.

At first sight, one might think that one could then go on to find R_{1J} by finding $\mathbf{Y}$ for each vector of an ensemble, $\mathbf{X}^{(J)}$, of initial vectors, each of which had only a

non-zero J^{th} component:

$$\mathbf{X}^{(J)} = a_J \mathbf{e}_J, \tag{3.123}$$

with J = 1...6, a_J a constant, and $\mathbf{e}_J$ the unit vector for the J^{th} component of X. In fact, the $Y_I^{(J)}$ values thus obtained would include the influence of not only R_{IJ}, but also of all non-zero T_{IJJ}, U_{IJJJ}, and V_{IJJJJ} matrix elements:

$$Y_I^{(J)} = C_I + R_{IJ}a_J + T_{IJJ}a_J^2 + U_{IJJJ}a_J^3 + V_{IJJJJ}a_J^4 + \mathcal{O}(a_J^5). \tag{3.124}$$

Obviously, one can extract C_I, R_{IJ}, T_{IJJ}, and U_{IJJJ} by fitting a fourth-order polynomial to this form (assuming that terms of fifth-order and higher can be ignored), if one takes a sufficient number of values of a_J for each J. A minimum of five initial vectors are needed for each value of J. Since I consider only static systems, J=5 (i.e., path-length dependent) terms are all zero, so a minimum of twenty-five vectors needs to be integrated. As I will discuss below, I use N vectors per component J, with N odd and $N \geq 5$:

$$\mathbf{X}^{(Jj)} = (j - \frac{N+1}{2})a_J \mathbf{e}_J, \quad j = 1 \ldots 5. \tag{3.125}$$

The reason for this particular choice of $\mathbf{X}^{(Jj)}$, which is symmetric about and includes the origin, will become apparent below. a_J is chosen sufficiently small so as to avoid large contributions to Y from terms higher than fifth order, while obtaining reasonable influence from third order terms, so that fitting will yield sufficiently precise values for the third-order coefficients. This step gives all elements C_I, R_{IJ}, T_{IJJ}, U_{IJJJ}, and, as a useful bonus, V_{IJJJJ}. It remains to find $T_{IJK} = T_{IKJ}$, $U_{IJKK} = U_{IKJK} = U_{IKKJ}$, and U_{IJKL}, for J > K and K > L.

3.4.2 Two-Variable Terms

To find T_{IJK} and U_{IJKK}, I integrate the equations of motion for a new ensemble of initial vectors for each (J,K) pair with J > K, described by

$$\mathbf{X}^{(Jj,Kk)} = (j - \frac{N+1}{2})a_J \mathbf{e}_J + (-1)^k a_K \mathbf{e}_K, \tag{3.126}$$

where j = 1...N, k = 1 or 2, N is an odd integer, and a_J and a_K are constants.

I now construct a residual final vector, $\Delta\mathbf{Y}^{(Jj,Kk)}$ for each $\mathbf{X}^{(Jj,Kk)}$ by subtracting off the contributions of the known matrix elements.

$$\begin{aligned}
\Delta Y_I^{(Jj,Kk)} &= Y_I^{(Jj,Kk)} - C_I - \sum_M \left\{ R_{IM}X_M + T_{IMM}X_M^2 + U_{IMMM}X_M^3 + V_{IMMM}X_M^4 \right\} \\
&= T_{IJK}X_JX_K + U_{IJJK}(X_J)^2X_K + U_{IJKK}X_J(X_K)^2 \\
&\quad + V_{IJJJK}(X_J)^3X_K + V_{IJJKK}(X_JX_K)^2 + V_{IJKKK}X_J(X_K)^3 \\
&\quad + \mathcal{O}(\mathbf{X}^5),
\end{aligned} \tag{3.127}$$

where for brevity $\mathbf{X} \equiv \mathbf{X}^{(Jj,Kk)}$ in this equation. Using equation (3.126) and dropping terms of fifth order and higher, this becomes

$$\begin{aligned}
\Delta Y_I^{(Jj,Kk)} &= T_{IJK}(j - \frac{N+1}{2})a_J(-1)^k a_K + U_{IJJK}((j - \frac{N+1}{2})a_J)^2(-1)^k a_K + \\
&\quad U_{IJKK}(j - \frac{N+1}{2})a_J a_K^2 + V_{IJKKK}(j - \frac{N+1}{2})a_J(-1)^k a_K^3 + \\
&\quad V_{IJJKK}((j - \frac{N+1}{2})a_J)^2 a_K^2 + V_{IJJJK}((j - \frac{N+1}{2})a_J)^3(-1)^k a_K
\end{aligned} \tag{3.128}$$

I define the sum and difference of the residuals for k=1 and k=2 as

$$\Delta\mathbf{Y}^{(Jj,K,S)} \equiv \Delta\mathbf{Y}^{(Jj,K,2)} + \Delta\mathbf{Y}^{(Jj,K,1)} \tag{3.129}$$

and

$$\Delta\mathbf{Y}^{(Jj,K,D)} \equiv \Delta\mathbf{Y}^{(Jj,K,2)} - \Delta\mathbf{Y}^{(Jj,K,1)}. \tag{3.130}$$

Using equation (3.128), these evaluate to

$$\Delta\mathbf{Y}^{(Jj,K,S)} = 2U_{IJKK}(j - \frac{N+1}{2})a_J a_K^2 + 2V_{IJJKK}((j - \frac{N+1}{2})a_J)^2 a_K^2 \tag{3.131}$$

and

$$\begin{aligned}
\Delta\mathbf{Y}^{(Jj,K,D)} &= 2T_{IJK}(j - \frac{N+1}{2})a_J a_K + 2U_{IJJK}((j - \frac{N+1}{2})a_J)^2 a_K + \\
&\quad 2V_{IJKKK}(j - \frac{N+1}{2})a_J a_K^3 + 2V_{IJJJK}((j - \frac{N+1}{2})a_J)^3 a_K
\end{aligned} \tag{3.132}$$

From equation (3.131), one can find U_{IJKK} and V_{IJJKK} from the linear and quadratic terms, respectively, of a fit that is quadratic in $(j - \frac{N+1}{2})a_J$. Similarly, equation (3.132) indicates that one can find $T_{IJK} - a_K^2 V_{IJKKK}$ and U_{IJJK} from the linear and

quadratic terms of a fit that is cubic in $(j - \frac{N+1}{2})a_J$. By doing the analysis for two different values of a_K, one can separate T_{IJK} from $T_{IJK} + a_K^2 V_{IJKKK}$. For a general element, then, at the very least one needs twenty integrations (i.e., N=5, two values of a_K, j=1,2) for every pair (J, K), for J > K, or $20 * 15 = 300$ integrations. (The twenty is the number of integrations per pair; the fifteen is the number of (J, K) pairs such that $6 \geq J > K \geq 1$.) Since the elements with J=5 or K=5 are known beforehand to be zero for a static element, this is reduced to a minimum of $20 * 10 = 200$ integrations for the alpha-magnet.

3.4.3 Three-Variable Terms

Having completed this step, only the elements U_{IJKL} with J > K and K > L remain to be found. To obtain these, new initial vectors are chosen for each triplet (J, K, L) with J > K > L:

$$\mathbf{X}^{(J,K,L,i)} = (-1)^i(a_J\mathbf{e}_J + a_K\mathbf{e}_K + a_L\mathbf{e}_L), \tag{3.133}$$

where i is 1 or 2, and a_J, a_K, and a_L are constants.

Again I compute residual final vectors $\Delta\mathbf{Y}^{(J,K,L,i)}$ by subtracting off the contributions of all R, T, and U matrix elements calculated so far:

$$\begin{aligned}\Delta Y_I^{(J,K,L,i)} &\equiv Y_I^{(J,K,L,i)} - C_I - \sum_N \left\{R_{IN}X_N - T_{INN}X_N^2 - U_{INNN}X_N^3 - V_{INNNN}X_N^4\right\} \\ &\quad - \sum_{N>M}\left\{T_{INM}X_NX_M - (U_{INMM}X_N(X_M)^2 - U_{INNM}(X_N)^2X_M)\right\} \\ &= \sum_{N>M>P} U_{INMP}X_NX_MX_P - \sum_{N>M} V_{INNMM}(X_NX_M)^2 \\ &\quad - \sum_{N>M>P}\left\{V_{INMPP}X_NX_MX_P^2 - V_{INMMP}X_NX_M^2X_P + V_{INNMP}X_N^2X_MX_P\right\} \\ &\quad + \mathcal{O}(\mathbf{X}^5),\end{aligned} \tag{3.134}$$

where for brevity $\mathbf{X} \equiv \mathbf{X}^{(J,K,L,i)}$ in this equation. Using equation (3.133) and dropping terms of fifth order and higher, this becomes

$$\begin{aligned}\Delta Y_I^{(i)} &= U_{IJKL}a_Ja_Ka_L(-1)^i + \\ &\quad (V_{IJJKK}a_J^2a_K^2 - V_{IJJLL}a_J^2a_L^2 + V_{IKKLL}a_K^2a_L^2) - \\ &\quad (V_{IJKLL}a_Ja_Ka_L^2 + 12V_{IJKKL}a_Ja_K^2a_L^2 - V_{IJJKL}a_J^2a_Ka_L).\end{aligned} \tag{3.135}$$

I now form the difference of the residuals for i=2 and i=1, obtaining

$$\Delta Y_I^{(D)} \equiv Y_I^{(2)} - Y_I^{(1)} = 2U_{IJKL}a_Ja_Ka_L. \tag{3.136}$$

Thus, one can obtain the U_{IJKL} with J > K and K > L by integrating a two additional vectors for each triplet J > K > L, requiring 40 additional integrations for a general element. For a static element, $U_{IJKL} = 0$ for J=5, K=5, or L=5, which reduces the number of additional integrations to 20.

3.4.4 Initial-Vector Ensemble

The reader may have noticed that the ensembles specified by equations (3.125), (3.126), and (3.133) overlap. Because of this, it is possible to use the ensemble of vectors defined by

$$\mathbf{X}^{(j,j,k,k,l,i)} = (\frac{N_J - 1}{2} - j)a_J\mathbf{e}_J - (\frac{N_K + 1}{2} - k)a_K\mathbf{e}_K + (\frac{N_L + 1}{2} - l)a_L\mathbf{e}_L, \tag{3.137}$$

with N_J odd, $6 \geq J > K > L \geq 1$, and j, k, and l taking integer values between $-(N-1)/2$ and $(N-1)/2$ (where $N = N_J$, N_K, or N_L, for j, k, and l, respectively) except that $j = k = l = 0$ (the null vector) appears only once for all triplets (J, K, L). The maximum amplitude of the J^{th} vector component is

$$M_J = \frac{N_J - 1}{2}a_J. \tag{3.138}$$

Since for a static element, X_5 is irrelevant, one can choose $N_5 = 1$ and $a_5 = 0$. It is also convenient to choose $N_J = N$ for all $J \neq 5$. Given both of these choices, the number of vectors in the ensemble is

$$10(N^3 - 1) + 10(N^2 - 1) + 1, \tag{3.139}$$

where I count ten (J, K, L) triplets with neither J, K, nor L equal to 5, contributing $N^3 - 1$ vectors each, exclusive of the null vector; ten (J, K, L) triplets with one of J, K, or L equal to 5, contributing $N^2 - 1$ vectors each, exclusive of the null vector; plus one null vector.

For N = 5 this ensemble contains about 6 times as many vectors as the minimum needed, but using it has the advantage of simplicity of coding and also of providing additional data to improve the accuracy of some of the elements by averaging.

Doubtless a more computationally efficient ensemble could be coded than I have used in my codes.

Having assembled this ensemble, one integrates each initial vector to obtain the corresponding final vector. One then selects out the necessary sub-ensembles corresponding to equation (3.125), (3.126), or (3.133)) for each stage of the analysis.

3.4.5 Accuracy Considerations and Limits

I have taken pains in the above analysis to eliminate the influence of fourth-order terms in order to increase the accuracy of the third-order matrix. In addition, suitable choice of the constants a_J can ensure that the effects of fourth and higher-order terms are negligible. "Suitable" must be determined empirically, or by reference to the magnitude of the matrix elements once they are roughly known. A starting point is to assume that the dominant fourth-order matrix elements have magnitudes similar to those of dominant matrix elements of third-order, from which one would conclude that that $M_J = 10^{-3}$ would be suitable to obtain 0.1% accuracy of the third-order results, even without the corrections for the V matrix that are included in the equations. Further, fifth order terms would be expected have an influence of one part in a million relative to the third-order. If similar results are obtained for a wide range of values of M_J, then one can conclude that the influence of higher-order terms is indeed negligible. In addition, if the contributions of the first, second, and third-order matrices to the final coordinates $\mathbf{Y}$ are seen to be different by several orders of magnitude between successive orders, then one can reasonably conclude that higher-order effects are several orders of magnitude below the third-order effects.

Invariably the above procedure will yield some small, non-zero matrix elements which may or may not be genuine, due to the accumulation of inaccuracies in the integration, subtraction of higher-order terms, and fitting. If one knows that the accuracy of any integration is of order 10^{-p}, where p is an integer, then one can conclude that a computed matrix element is spurious if it fails to satisfy the appropriate criterion (depending on the order of the matrix element) from the following list:

$$R_{IJ}M_J > 10^{-p} \tag{3.140}$$

$$T_{IJK}M_JM_K > 10^{-p} \tag{3.141}$$

$$U_{IJKL}M_JM_KM_L > 10^{-p}. \tag{3.142}$$

One can also use these same relations to estimate the resolution with which genuine non-zero matrix elements could be calculated.

$$\Delta R_{IJ} > \frac{10^{-p}}{M_J} \tag{3.143}$$

$$\Delta T_{IJK} > \frac{10^{-p}}{M_JM_K} \tag{3.144}$$

$$\Delta U_{IJKL} > \frac{10^{-p}}{M_JM_KM_L}. \tag{3.145}$$

One expects that this resolution will not be achieved, since it does not consider the inaccuracies in fitting and subtraction to obtain residuals. Nevertheless, these criteria do provide a solid lower bound on the precision of the matrix elements. In the case of the alpha-magnet, I have shown above that the accuracy of integrations is 10^{-14}.

3.5 Transport Matrices for the Alpha Magnet

I have written a computer program, `salpha_matrix`, that implements the ideas of the previous two sections. Matrices up to third order have been computed for transport from the entrance of the alpha-magnet to the exit, from the entrance to the vertical midplane, and from the vertical midplane to the exit.

3.5.1 Program Tests and Choice of Initial Amplitudes

For purposes of testing the coding and the method of obtaining the matrix, the program has the option of generating C, R, T, and U matrices with all components given by random numbers between -1 and 1, and then tracking vectors through these matrices instead of integrating the equations of motion for the alpha magnet. It then attempts to recover the random matrices by analyzing the initial and final vectors only, just as would be done for initial and final vectors obtained by integration. This tests the ability of the program to separate various orders, but does not test its ability to suppress the effects of orders higher than third. Table 3.2 summarizes the results of this test. As will be seen, the errors are below those that are encountered in fitting matrices for the alpha-magnet, as would be expected. The errors from this procedure can be considered to place the ultimate limit on the accuracy with which matrices for the alpha-magnet can be calculated.

Table 3.2: **Accuracy of Recovery of a Randomly Generated Matrix**

matrix	maximum error of fit for any matrix element	average deviation of fit for all matrix elements
C	$1.38 \cdot 10^{-13}$	$7.59 \cdot 10^{-15}$
R	$8.87 \cdot 10^{-14}$	$2.71 \cdot 10^{-14}$
T	$1.13 \cdot 10^{-10}$	$2.48 \cdot 10^{-11}$
U	$1.47 \cdot 10^{-7}$	$1.77 \cdot 10^{-8}$

An initial round of computations for the alpha-magnet were done with N = 5 and with all M_J values equal, for a series of different values from 10^{-2} to 10^{-5}. After the matrix was obtained, the average of the absolute values of the residuals of the final

vectors for all initial vectors were computed to assess the degree to which the fits contained sufficiently high-order terms to match the actual final vectors. Residuals were computed by successively adding linear, second-order, and third-order terms to assess the effect of each order. The average absolute residuals for n^{th} order are simply

$$\mathcal{R}_I^{(n)} = \frac{1}{M}\sum_i \left| Y_I^{(i)} - \left\{ C_I - \sum_J R_{IJ} X_J^{(i)} - (n > 1?) \sum_{J\geq K} T_{IJK} X_J^{(i)} X_K^{(i)} - \right. \right. \quad (3.146)$$

$$\left. \left. (n > 2?) \sum_{J\geq K\geq L} U_{IJKL} X_I^{(i)} X_J^{(i)} X_K^{(i)} \right\} \right|, \quad (3.147)$$

where the index i runs over all M initial vectors in the ensemble specified by equation (3.137), and $(n > m?)$ represents a function that returns 1 if $n > m$ and 0 otherwise. Table 3.3 summarizes some of the results. ($\mathcal{R}_6^{(n)}$ is identically zero, since the momentum is not changed by the magnet, and hence is not listed.) It is no coincidence that for any particular I and n, $\mathcal{R}_I^{(n)}$ varies with M_J according to M_J^{n-1}, since for valid fits (i.e., those that don't err by compromising lower order coefficients in order to match higher-order contributions) $\mathcal{R}_I^{(n)}$ is simply the average contribution of the $(n-1)^{th}$ order terms.

Table 3.3: **Residuals from Matrix Fits**

n	M_J	$\mathcal{R}_1^{(n)}$	$\mathcal{R}_2^{(n)}$	$\mathcal{R}_3^{(n)}$	$\mathcal{R}_4^{(n)}$	$\mathcal{R}_5^{(n)}$
1	10^{-2}	$1.84 \cdot 10^{-3}$	$1.89 \cdot 10^{-3}$	$4.51 \cdot 10^{-4}$	$3.79 \cdot 10^{-4}$	$7.63 \cdot 10^{-5}$
1	10^{-3}	$1.84 \cdot 10^{-5}$	$1.89 \cdot 10^{-5}$	$4.49 \cdot 10^{-6}$	$3.53 \cdot 10^{-6}$	$7.76 \cdot 10^{-7}$
1	10^{-4}	$1.84 \cdot 10^{-7}$	$1.90 \cdot 10^{-7}$	$4.49 \cdot 10^{-8}$	$3.51 \cdot 10^{-8}$	$7.76 \cdot 10^{-9}$
2	10^{-2}	$3.44 \cdot 10^{-5}$	$2.39 \cdot 10^{-5}$	$4.59 \cdot 10^{-5}$	$1.43 \cdot 10^{-4}$	$6.38 \cdot 10^{-6}$
2	10^{-3}	$3.44 \cdot 10^{-8}$	$2.39 \cdot 10^{-8}$	$4.61 \cdot 10^{-8}$	$1.43 \cdot 10^{-7}$	$6.34 \cdot 10^{-9}$
2	10^{-4}	$3.44 \cdot 10^{-11}$	$2.39 \cdot 10^{-11}$	$4.61 \cdot 10^{-11}$	$1.42 \cdot 10^{-10}$	$6.34 \cdot 10^{-12}$
3	10^{-2}	$9.40 \cdot 10^{-7}$	$8.78 \cdot 10^{-7}$	$2.08 \cdot 10^{-6}$	$2.37 \cdot 10^{-6}$	$7.19 \cdot 10^{-7}$
3	10^{-3}	$8.98 \cdot 10^{-11}$	$8.43 \cdot 10^{-11}$	$2.08 \cdot 10^{-10}$	$2.36 \cdot 10^{-10}$	$7.21 \cdot 10^{-11}$
3	10^{-4}	$1.14 \cdot 10^{-14}$	$1.01 \cdot 10^{-14}$	$2.07 \cdot 10^{-14}$	$2.36 \cdot 10^{-14}$	$7.40 \cdot 10^{-15}$

Table 3.3 shows that for $M_J = 10^{-4}$, the third-order residuals are of order 10^{-14}, which is the accuracy limit of the integrations. Hence, fourth-order contributions are "in the noise", and third-order contributions are three orders of magnitude above it. I find that for such small M_J values, the chromatic terms do not follow equations (3.102)

through (3.107) as well as for $M_J = 10^{-3}$. For this reason, I choose the matrices computed with $M_J = 10^{-3}$ as the most accurate. Analysis of the chromatic terms indicates that the T matrix elements are accurate to about 10^{-6}, indicating that p in equations (3.143) through (3.145) is 12 (rather than 14 as would be thought from the accuracy of the integration). I use this value of p in order to "filter" small T_{IJK} and U_{IJKL} values for significance, as per equations (3.140) through (3.142). That is, T_{IJK} values smaller than 10^{-6} and U_{IJKL} values smaller than 10^{-3} are taken to be zero.

3.5.2 Final Results

Having verified the program's matrix-fitting algorithm and found the limits of its accuracy, I computed the matrices for the alpha magnet using $N = 7$. I used an accuracy limit of 5×10^{-13} to filter out spurious non-zero matrix elements. This limit is a compromise between one that is somewhat too large for the T matrix elements, and somewhat too small for the U matrix elements. Hence, some small, dubious U matrix elements will appear in the results that follow.

Checks of the Results

A number of checks have been made on these matrices. The determinants of the first order matrices for entrance-to-exit, entrance-to-vertical midplane, and vertical midplane-to-exit have been found to be 1 to within 2×10^{-12}. (This accuracy is not fully reflected in the results given below, since I have not quoted a sufficient number of significant figures. Also, the reader should beware of checking this claim with a hand calculator, since many use only 10 or 11 digits.)

The relationships between the non-chromatic and chromatic terms were used to evaluate the accuracy of the method, as discussed above; the reader is invited to use equations (3.102) through (3.107) and (3.53) through (3.60) to verify for himself that the results do indeed satisfy the expected numerical ratios. As a sample, for the matrix from the entrance to the vertical midplane, I find that

$$\frac{R_{56}}{C_5} = \frac{1}{2} - 2 \cdot 10^{-12} \tag{3.148}$$

$$\frac{T_{566}}{C_5} = -\frac{1}{8} - 6 \cdot 10^{-10} \tag{3.149}$$

$$\frac{U_{5666}}{C_5} = \frac{1}{16} - 2 \cdot 10^{-8} \tag{3.150}$$

$$\frac{T_{162}}{R_{12}} = \frac{1}{2} - 2 \cdot 10^{-9} \tag{3.151}$$

$$\frac{U_{1662}}{R_{12}} = -\frac{1}{8} - 6 \cdot 10^{-6}. \tag{3.152}$$

The reader will see below that the computed entrance-to-vertical midplane and vertical midplane-to-exit R matrices satisfy the expected relationship for two elements that are the reverse of each other[6], namely

$$\tilde{R} = \begin{pmatrix} R_{22} & R_{12} \\ R_{21} & R_{11} \end{pmatrix}, \tag{3.153}$$

where $\tilde{R}$ is the matrix for system that is reversed in order relative to the system for which R is the matrix.

The matrices for the first and second parts of the alpha-magnet were concatenated (using a third-order matrix concatenation program written by the author) and compared to the matrices computed for the full magnet. No significant discrepancies were found for the R matrix. The only discrepancies found in the T matrix were fractional variations of 10^{-5} in the two smallest elements; all other T matrix elements either showed no discrepancy, or discrepancies only in the last decimal place. For most U matrix elements, the discrepancy was less than .1 %, while for a few of the smallest U matrix elements, the error was between 1 and 10 %.

In order to ensure that there were no transcription errors made, `salpha_matrix` provided output in LaTeX format, which was included in this document with only minor editing to properly columnate the data.

Entrance-to-Exit Transport

For transport from the entrance to the exit, tables 3.4 and 3.5 list non-zero T and U matrix elements, respectively. The following are the centroid and R matrix elements (unlisted elements are zero):

$$C_5 = 4.642099465061 \tag{3.154}$$

$$R_{56} = 2.321049733 \tag{3.155}$$

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} -1.000000000 & -2.321049733 \\ 0.000000000 & -1.000000000 \end{pmatrix} \quad (3.156)$$

$$\begin{pmatrix} R_{33} & R_{34} \\ R_{43} & R_{44} \end{pmatrix} = \begin{pmatrix} -0.7371140937 & 7.618204274 \\ -0.05994362928 & -0.7371140937 \end{pmatrix} \quad (3.157)$$

Table 3.4: **Non-zero T Matrix Elements from Entrance to Exit**

$T_{122} = -9.985582 \cdot 10^{-1}$	$T_{133} = -6.047097 \cdot 10^{-1}$	$T_{143} = -6.415746$
$T_{144} = 3.782911 \cdot 10^{1}$	$T_{162} = -1.160525$	$T_{233} = -2.996743 \cdot 10^{-1}$
$T_{243} = -7.370063$	$T_{244} = 3.808545 \cdot 10^{1}$	$T_{331} = -5.157770 \cdot 10^{-2}$
$T_{332} = 9.264364$	$T_{342} = 6.892273$	$T_{364} = 3.809102$
$T_{432} = -1.750135$	$T_{441} = 5.157770 \cdot 10^{-2}$	$T_{442} = 9.384079$
$T_{463} = 2.997181 \cdot 10^{-2}$	$T_{522} = 5.802624 \cdot 10^{-1}$	$T_{533} = 1.104632 \cdot 10^{-2}$
$T_{543} = -2.283314 \cdot 10^{-1}$	$T_{544} = -1.403871$	$T_{566} = -5.802624 \cdot 10^{-1}$

Table 3.5: **Non-zero U Matrix Elements from Entrance to Exit**

$U_{1222} = 9.634 \cdot 10^{-1}$	$U_{1331} = -2.579 \cdot 10^{-1}$	$U_{1332} = -8.573$
$U_{1431} = -5.301$	$U_{1432} = 8.263 \cdot 10^{1}$	$U_{1441} = 3.829 \cdot 10^{1}$
$U_{1442} = 9.435 \cdot 10^{1}$	$U_{1622} = -4.993 \cdot 10^{-1}$	$U_{1633} = 3.023 \cdot 10^{-1}$
$U_{1644} = 1.891 \cdot 10^{1}$	$U_{1662} = 2.901 \cdot 10^{-1}$	$U_{2332} = -8.779$
$U_{2431} = 5.157 \cdot 10^{-1}$	$U_{2432} = 9.310 \cdot 10^{1}$	$U_{2441} = 6.341$
$U_{2442} = 4.557 \cdot 10^{1}$	$U_{2633} = 2.997 \cdot 10^{-1}$	$U_{2643} = 3.685$
$U_{3321} = -1.462$	$U_{3322} = -4.979 \cdot 10^{-1}$	$U_{3333} = 1.047$
$U_{3411} = 4.431 \cdot 10^{-2}$	$U_{3421} = 1.033 \cdot 10^{-1}$	$U_{3422} = 1.101 \cdot 10^{1}$
$U_{3432} = -2.705 \cdot 10^{-3}$	$U_{3433} = -2.310$	$U_{3443} = 2.220 \cdot 10^{1}$
$U_{3444} = -1.075 \cdot 10^{2}$	$U_{3631} = 2.579 \cdot 10^{-2}$	$U_{3642} = 3.446$
$U_{3664} = -9.523 \cdot 10^{-1}$	$U_{4321} = 1.285 \cdot 10^{1}$	$U_{4332} = -4.323 \cdot 10^{-1}$
$U_{4421} = 1.463$	$U_{4422} = 2.946$	$U_{4432} = -1.636 \cdot 10^{-3}$
$U_{4433} = -1.441$	$U_{4443} = 8.688 \cdot 10^{1}$	$U_{4444} = -3.156 \cdot 10^{2}$
$U_{4622} = 8.751 \cdot 10^{-1}$	$U_{4641} = -2.575 \cdot 10^{-2}$	$U_{4663} = -2.248 \cdot 10^{-2}$
$U_{5222} = 4.993 \cdot 10^{-1}$	$U_{5331} = 1.546 \cdot 10^{-3}$	$U_{5332} = 2.469 \cdot 10^{-2}$
$U_{5421} = 1.873 \cdot 10^{-2}$	$U_{5432} = -4.126 \cdot 10^{-1}$	$U_{5441} = 9.242 \cdot 10^{-4}$
$U_{5442} = -2.145 \cdot 10^{1}$	$U_{5622} = 2.901 \cdot 10^{-1}$	$U_{5633} = -5.524 \cdot 10^{-3}$
$U_{5644} = -7.019 \cdot 10^{-1}$	$U_{5666} = 2.901 \cdot 10^{-1}$	

Entrance-to-Midplane Transport

For transport from the entrance to the vertical midplane, tables 3.6 and 3.7 list non-zero T and U matrix elements, respectively. The following are the centroid and R matrix elements (unlisted elements are zero):

$$C_5 = 2.321049732530 \quad (3.158)$$

$$R_{51} = -2.179660432 \quad (3.159)$$

$$R_{52} = -2.529550131 \quad (3.160)$$

$$R_{16} = -9.089085575 \cdot 10^{-1} \quad (3.161)$$

$$R_{56} = 1.160524866 \quad (3.162)$$

$$R_{66} = 1.000000000 \quad (3.163)$$

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} 0.000000000 & 0.4169954844 \\ -2.398107503 & -2.783063390 \end{pmatrix} \quad (3.164)$$

$$\begin{pmatrix} R_{33} & R_{34} \\ R_{43} & R_{44} \end{pmatrix} = \begin{pmatrix} 0.07531765053 & 2.182639820 \\ -0.3979387890 & 1.745181272 \end{pmatrix} \quad (3.165)$$

Table 3.6: **Non-zero T Matrix Elements from Entrance to Vertical Midplane**

$T_{111} = 1.581820$	$T_{121} = 3.671483$	$T_{122} = 2.357651$
$T_{133} = 3.170513 \cdot 10^{-1}$	$T_{143} = -4.724714 \cdot 10^{-1}$	$T_{144} = -2.932169 \cdot 10^{-1}$
$T_{162} = 2.084977 \cdot 10^{-1}$	$T_{166} = 2.272271 \cdot 10^{-1}$	$T_{221} = 2.613530$
$T_{222} = 1.835742$	$T_{233} = -3.739056 \cdot 10^{-1}$	$T_{243} = 9.437959 \cdot 10^{-1}$
$T_{244} = 7.286366 \cdot 10^{-1}$	$T_{261} = 1.199054$	$T_{331} = 5.249703 \cdot 10^{-1}$
$T_{332} = 3.240582$	$T_{341} = -2.367091$	$T_{342} = -1.933840$
$T_{364} = 1.091320$	$T_{431} = -9.936063 \cdot 10^{-2}$	$T_{432} = 2.168953$
$T_{441} = -2.536989$	$T_{442} = -2.015590$	$T_{463} = 1.989694 \cdot 10^{-1}$
$T_{521} = 1.875460$	$T_{522} = 1.378390$	$T_{533} = -3.473390 \cdot 10^{-1}$
$T_{543} = 4.235456 \cdot 10^{-1}$	$T_{544} = 1.614540$	$T_{562} = -1.264775$
$T_{566} = -2.901312 \cdot 10^{-1}$		

Table 3.7: **Non-zero U Matrix Elements from Entrance to Vertical Midplane**

$U_{1211} = -3.085$	$U_{1221} = -5.581$	$U_{1222} = -2.478$
$U_{1331} = 4.933 \cdot 10^{-1}$	$U_{1332} = 3.610 \cdot 10^{-2}$	$U_{1431} = -1.791$
$U_{1432} = -2.841$	$U_{1441} = -5.547 \cdot 10^{-1}$	$U_{1442} = -1.004$
$U_{1611} = -7.909 \cdot 10^{-1}$	$U_{1 \ 2} = 1.179$	$U_{1633} = -1.585 \cdot 10^{-1}$
$U_{1644} = -1.466 \cdot 10^{-1}$	$U_{1 \ 52} = -5.213 \cdot 10^{-2}$	$U_{1666} = -1.136 \cdot 10^{-1}$
$U_{2111} = -6.200$	$U_{2 \ 11} = -2.159 \cdot 10^{1}$	$U_{2221} = -2.700 \cdot 10^{1}$
$U_{2222} = -9.493$	$U_{2 \ 31} = 2.284 \cdot 10^{-1}$	$U_{2332} = 1.145$
$U_{2421} = -1.025 \cdot 10^{-3}$	$U_{2431} = 1.686$	$U_{2432} = 1.572$
$U_{2441} = -4.851$	$U_{2442} = -4.688$	$U_{2621} = -1.307$
$U_{2632} = 3.739 \cdot 10^{-1}$	$U_{2643} = -4.719 \cdot 10^{-1}$	$U_{2661} = -8.993 \cdot 10^{-1}$
$U_{3311} = 6.554 \cdot 10^{-2}$	$U_{3321} = -3.313$	$U_{3322} = -4.769$
$U_{3333} = 3.965 \cdot 10^{-1}$	$U_{3411} = 1.448$	$U_{3421} = 1.908$
$U_{3422} = 3.934 \cdot 10^{-1}$	$U_{3433} = -1.195$	$U_{3443} = -8.190 \cdot 10^{-1}$
$U_{3444} = 1.457$	$U_{3631} = -2.625 \cdot 10^{-1}$	$U_{3642} = -9.669 \cdot 10^{-1}$
$U_{3664} = -2.728 \cdot 10^{-1}$	$U_{4311} = -3.463 \cdot 10^{-1}$	$U_{4321} = -4.190$
$U_{4322} = -4.777$	$U_{4333} = 2.704 \cdot 10^{-1}$	$U_{4411} = 1.604$
$U_{4421} = 2.868$	$U_{4422} = 1.015$	$U_{4433} = -9.890 \cdot 10^{-1}$
$U_{4443} = 1.350 \cdot 10^{-1}$	$U_{4444} = 5.265 \cdot 10^{-1}$	$U_{4631} = 9.936 \cdot 10^{-2}$
$U_{4632} = -1.084$	$U_{4641} = 1.268$	$U_{4663} = -1.492 \cdot 10^{-1}$
$U_{5111} = -1.264$	$U_{5211} = -4.402$	$U_{5221} = -7.153$
$U_{5222} = -2.112$	$U_{5331} = -1.045 \cdot 10^{-1}$	$U_{5332} = 1.548 \cdot 10^{-1}$
$U_{5431} = 1.514$	$U_{5432} = 3.542$	$U_{5441} = -2.804$
$U_{5442} = -1.689$	$U_{5622} = 6.892 \cdot 10^{-1}$	$U_{5633} = 1.737 \cdot 10^{-1}$
$U_{5644} = 8.073 \cdot 10^{-1}$	$U_{5662} = 3.162 \cdot 10^{-1}$	$U_{5666} = 1.451 \cdot 10^{-1}$

Midplane to Exit Transport

For transport from the vertical midplane to the exit, tables 3.8 and 3.9 list non-zero T and U matrix elements, respectively. The following are the centroid and R matrix elements (unlisted elements are zero):

$$C_5 = 2.321049732530 \tag{3.166}$$

$$R_{16} = -2.529550131 \tag{3.167}$$

$$R_{26} = -2.179660432 \tag{3.168}$$

$$R_{52} = -0.9089085575 \tag{3.169}$$

$$R_{56} = 1.160524866 \tag{3.170}$$

$$R_{66} = 1.000000000 \tag{3.171}$$

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} -2.783063390 & 0.4169954844 \\ -2.398107503 & 0.0000000000 \end{pmatrix} \tag{3.172}$$

$$\begin{pmatrix} R_{33} & R_{34} \\ R_{43} & R_{44} \end{pmatrix} = \begin{pmatrix} 1.745181272 & 2.182639820 \\ -0.3979387890 & 0.07531765053 \end{pmatrix} \tag{3.173}$$

Table 3.8: **Non-zero T Matrix Elements from Vertical Midplane to Exit**

$T_{111} = 7.654960 \cdot 10^{-1}$	$T_{121} = 1.089830$	$T_{122} = 7.654960 \cdot 10^{-1}$
$T_{133} = 1.491568$	$T_{143} = 4.885853$	$T_{144} = 3.609338$
$T_{161} = 1.391532$	$T_{162} = 1.199054$	$T_{166} = 1.264775$
$T_{211} = 3.133763$	$T_{222} = 6.596119 \cdot 10^{-1}$	$T_{233} = 1.417467$
$T_{243} = 4.999269$	$T_{244} = 3.804388$	$T_{261} = 6.895662$
$T_{266} = 4.223593$	$T_{331} = 2.227000$	$T_{332} = -1.057913$
$T_{341} = 1.950208$	$T_{342} = -9.870661 \cdot 10^{-1}$	$T_{363} = 2.024140$
$T_{364} = 2.863881$	$T_{431} = 5.477911$	$T_{432} = -4.143302 \cdot 10^{-2}$
$T_{441} = 6.310239$	$T_{442} = 2.189102 \cdot 10^{-1}$	$T_{463} = 5.177889$
$T_{464} = 5.735430$	$T_{511} = 1.668521$	$T_{521} = -5.000000 \cdot 10^{-1}$
$T_{533} = -1.736188 \cdot 10^{-1}$	$T_{543} = -4.342785 \cdot 10^{-1}$	$T_{544} = 4.109782 \cdot 10^{-2}$
$T_{561} = 3.033067$	$T_{562} = -9.089086 \cdot 10^{-1}$	$T_{566} = 1.088259$

Table 3.9: Non-zero U Matrix Elements from Vertical Midplane to Exit

$U_{1111} = -1.360$	$U_{1211} = -5.249 \cdot 10^{-1}$	$U_{1222} = -4.872 \cdot 10^{-1}$
$U_{1331} = 2.309$	$U_{1332} = -3.196$	$U_{1431} = 4.658$
$U_{1432} = -7.406$	$U_{1441} = 2.017$	$U_{1442} = -4.261$
$U_{1611} = -4.089$	$U_{1621} = -9.537 \cdot 10^{-1}$	$U_{1622} = 3.828 \cdot 10^{-1}$
$U_{1631} = 5.964 \cdot 10^{-4}$	$U_{1633} = 1.353$	$U_{1641} = 6.949 \cdot 10^{-4}$
$U_{1643} = 4.234$	$U_{1644} = 3.638$	$U_{1661} = -4.413$
$U_{1662} = -7.334 \cdot 10^{-1}$	$U_{1666} = -1.969$	$U_{2111} = -1.336 \cdot 10^{1}$
$U_{2221} = -1.361$	$U_{2331} = 2.331$	$U_{2332} = -2.750$
$U_{2431} = 1.181$	$U_{2432} = -5.745$	$U_{2441} = -2.352$
$U_{2442} = -2.750$	$U_{2611} = -3.957 \cdot 10^{1}$	$U_{2621} = -3.419 \cdot 10^{-3}$
$U_{2622} = -1.237$	$U_{2631} = -3.495 \cdot 10^{-3}$	$U_{2633} = 7.009 \cdot 10^{-1}$
$U_{2641} = -4.345 \cdot 10^{-3}$	$U_{2643} = -1.426$	$U_{2644} = -2.138$
$U_{2661} = -4.114 \cdot 10^{1}$	$U_{2666} = -1.528 \cdot 10^{1}$	$U_{3311} = 3.497$
$U_{3321} = -3.620$	$U_{3322} = -3.336 \cdot 10^{-1}$	$U_{3333} = -1.167$
$U_{3411} = 4.993$	$U_{3421} = -4.032$	$U_{3422} = -2.847 \cdot 10^{-1}$
$U_{3431} = 5.003 \cdot 10^{-4}$	$U_{3433} = -4.874$	$U_{3443} = -6.236$
$U_{3444} = -2.195$	$U_{3631} = 5.245$	$U_{3632} = -3.290$
$U_{3641} = 9.078$	$U_{3642} = -4.158$	$U_{3663} = 1.371$
$U_{3664} = 3.409$	$U_{4311} = -9.066$	$U_{4321} = -3.494$
$U_{4322} = -1.567$	$U_{4333} = -2.991$	$U_{4411} = -1.456 \cdot 10^{1}$
$U_{4421} = -2.893$	$U_{4422} = -1.724$	$U_{4431} = -5.345 \cdot 10^{-4}$
$U_{4433} = -1.323 \cdot 10^{1}$	$U_{4443} = -1.849 \cdot 10^{1}$	$U_{4444} = -8.388$
$U_{4631} = -2.196 \cdot 10^{1}$	$U_{4632} = -3.155$	$U_{4641} = -2.962 \cdot 10^{1}$
$U_{4642} = -2.630$	$U_{4643} = -5.769 \cdot 10^{-4}$	$U_{4663} = -1.386 \cdot 10^{1}$
$U_{4664} = -1.633 \cdot 10^{1}$	$U_{5111} = -9.178 \cdot 10^{-1}$	$U_{5211} = -1.307$
$U_{5221} = -9.179 \cdot 10^{-1}$	$U_{5222} = 4.545 \cdot 10^{-1}$	$U_{5331} = -2.232$
$U_{5332} = 2.105 \cdot 10^{-1}$	$U_{5431} = -6.162$	$U_{5432} = 1.565 \cdot 10^{-1}$
$U_{5441} = -4.254$	$U_{5442} = 4.173 \cdot 10^{-1}$	$U_{5611} = -3.337$
$U_{5621} = -2.375$	$U_{5622} = -8.343 \cdot 10^{-1}$	$U_{5633} = -1.941$
$U_{5643} = -5.601$	$U_{5644} = -3.846$	$U_{5661} = -4.550$
$U_{5662} = -8.523 \cdot 10^{-1}$	$U_{5666} = -1.923$	

3.6 Effects of Field Errors

All of the above analysis of the alpha-magnet assumes that the functional form of the magnetic field is that of a perfect quadrupole. In reality, no magnet is ideal. A review of the derivation of the scaled equation of motion shows that non-linear terms in the magnetic field will, strictly speaking, invalidate the scaling. In other words, the magnet will not be strictly achromatic, as a perfect alpha-magnet would be. One result of this is that the nominal ideal trajectory (i.e., the trajectory injected at incidence angle θ_o) will no longer exit the magnet at the same location that it entered at.

Magnetic field errors are a fact of life in accelerator physics. The favored approach to dealing with them is to evaluate the effect of specific types of errors (e.g., higher-order multipoles) with an eye toward what level of error one's application can tolerate. In accordance with this, I have studied the effect of certain types of field errors, such as sextupole terms, to find what effect they have on the performance of the alpha-magnet. (Similar, less complete work on this problem is reported in [32].) It has been found from computer studies that for a variety of errors, the residual dispersion after the alpha magnet can be reduced to acceptable levels by modifying the injection angle, θ_i, in such a way as to cause the ideal trajectory to once again exit at the entrance point. If the magnet retains reflection symmetry about the plane $q_2 = 0$, it is always possible to find such a value of θ_i, which I will call θ_m, or the "mirror angle". The reader can convince himself of this by reviewing the argument by which I proved that the perfect alpha-magnet has such an injection angle, namely θ_o.

The field in the imperfect alpha-magnet can be expressed as

$$\begin{aligned} \mathbf{B}(\mathbf{q}) &= g(q_3, 0, q_1) + \Delta\mathbf{B}(\mathbf{q}) \\ &= \frac{g}{\alpha}\left\{(Q_3, 0, Q_1) + \Delta\mathbf{B}(\frac{\mathbf{Q}}{\alpha})\frac{\alpha}{g}\right\}, \end{aligned} \tag{3.174}$$

where $\Delta\mathbf{B}(\mathbf{q})$ is the departure of the field from a true, uniform quadrupole field and, as before, $\mathbf{Q} \equiv \mathbf{q}\alpha$.

Comparison with equation (3.83) shows that the scaled equation of motion with

field errors is

$$\mathbf{Q}'' = -\frac{1}{1-\delta}\mathbf{Q}' \cdot \mathbf{B}(\frac{\mathbf{Q}}{\alpha})\frac{\alpha}{g} \tag{3.175}$$

$$= -\frac{1}{1-\delta}\mathbf{Q}' \cdot \left\{(Q_3, 0, Q_1) - \Delta\mathbf{B}(\frac{\mathbf{Q}}{\alpha})\frac{\alpha}{g}\right\} \tag{3.176}$$

3.6.1 Multipole Errors

With this equation in hand, it is possible to evaluate the effect of various field errors. Note that since α appears only as a multiplicative factor for the field error, it is still possible to find results with some universality. In particular, if $\Delta\mathbf{B}$ is a pure multipole error, then the effect of the field error in the equation of motion will have a well-defined scaling with α and the multipole coefficient.

Multipole fields can be classified as upright or rotated[6], depending on whether the magnetic fields are changed in sign upon reflection of the magnet through the $q_3 = 0$ plane or not, respectively. Upright multipoles have field lines that cross the $q_3 = 0$ plane with normal incidence. For rotated multipoles, field lines do not cross the $q_3 = 0$ plane. Clearly the alpha-magnet has upright symmetry, and if one confines oneself to errors that do not alter this symmetry, then one can express errors in the alpha-magnet in terms of the upright multipoles. For example, any deviation of the poles from a hyperbola will produce only upright multipole errors, as will displacement of the mirror plane, since neither of these errors changes the fact that the field lines cross $q_3 = 0$ with normal incidence.

The field due to a pure upright multipole is[6]:

$$\Delta B_n = A_n \sum_{m=1}^{\lfloor n/2 \rfloor} (-1)^{n-1} \frac{q_1^{n-2m}}{(n-2m)!}\frac{q_3^{2m-1}}{(2m-1)!}\dot{q}_1 - A_n \sum_{m=1}^{\lfloor (n-1)/2 \rfloor} (-1)^{m-1} \frac{q_1^{n-2m-1}}{(n-2m-1)!}\frac{q_3^{2m-2}}{(2m-2)!}\dot{q}_3, \tag{3.177}$$

where $n \geq 1$ is an integer, the "order" of the multipole. $n = 1$ is a dipole, $n = 2$ a quadrupole, and so forth.

For insertion into equation (3.176), this must be rewritten in terms of scaled

coordinates, as

$$\frac{\alpha}{g}\Delta B_n = M_n \sum_{m=1}^{\lfloor n/2 \rfloor} (-1)^{m-1} \frac{Q_1^{n-2m}}{(n-2m)!}\frac{Q_3^{2m-1}}{(2m-1)!}\dot{Q}_1 - M_n \sum_{m=1}^{\lfloor (n+1)/2 \rfloor} (-1)^{m-1} \frac{Q_1^{n-2m+1}}{(n-2m+1)!}\frac{Q_3^{2m-2}}{(2m-2)!}\dot{Q}_3, \tag{3.178}$$

where I have defined the dimensionless normalized multipole strength

$$M_n \equiv \frac{A_n}{g\alpha^{n-2}}. \tag{3.179}$$

Even without integration one can conclude that for the same fractional multipole error, $\frac{A_n}{g}$, the perturbation is stronger for smaller α, i.e., for alpha-magnets operated so as to obtain larger values of $\dot{q}_1$. This is as expected, since the multipole field grows as q_1^n. As expected, in the limit of very large α, i.e., very small $\dot{q}_1$, multipole errors have no effect.

It is of course possible to compute the matrices for equation (3.176) with $\Delta\mathbf{B}$ as given by (3.179) as was done for the equation of motion without field errors. The matrices thus obtained are to be considered functions of M_n, with M_n ultimately a function of α. Hence, if the matrices are found for some particular value of M_n for some particular n, then if the matrices are scaled to some particular value of α according to equation (3.93), the result is appropriate to a multipole strength of

$$A_n = M_n g \alpha^{n-2}. \tag{3.180}$$

I have written a computer program, **serrors**, which computes third-order scaled alpha-magnet matrices in the presence of various types of field errors, including multipole errors. Figures 3.10 through 3.12 show the effect of sextupole errors on the mirror-angle (θ_m), $\dot{Q}_1$, and the non-zero elements of the matrix R. Note the particularly strong effect in the vertical plane.

3.6.2 Entrance-Hole-Induced Errors

I performed magnetic measurements on the SSRL alpha-magnet to assess the deviation of the field from an ideal quadrupolar field. Figure 3.13 shows the measured

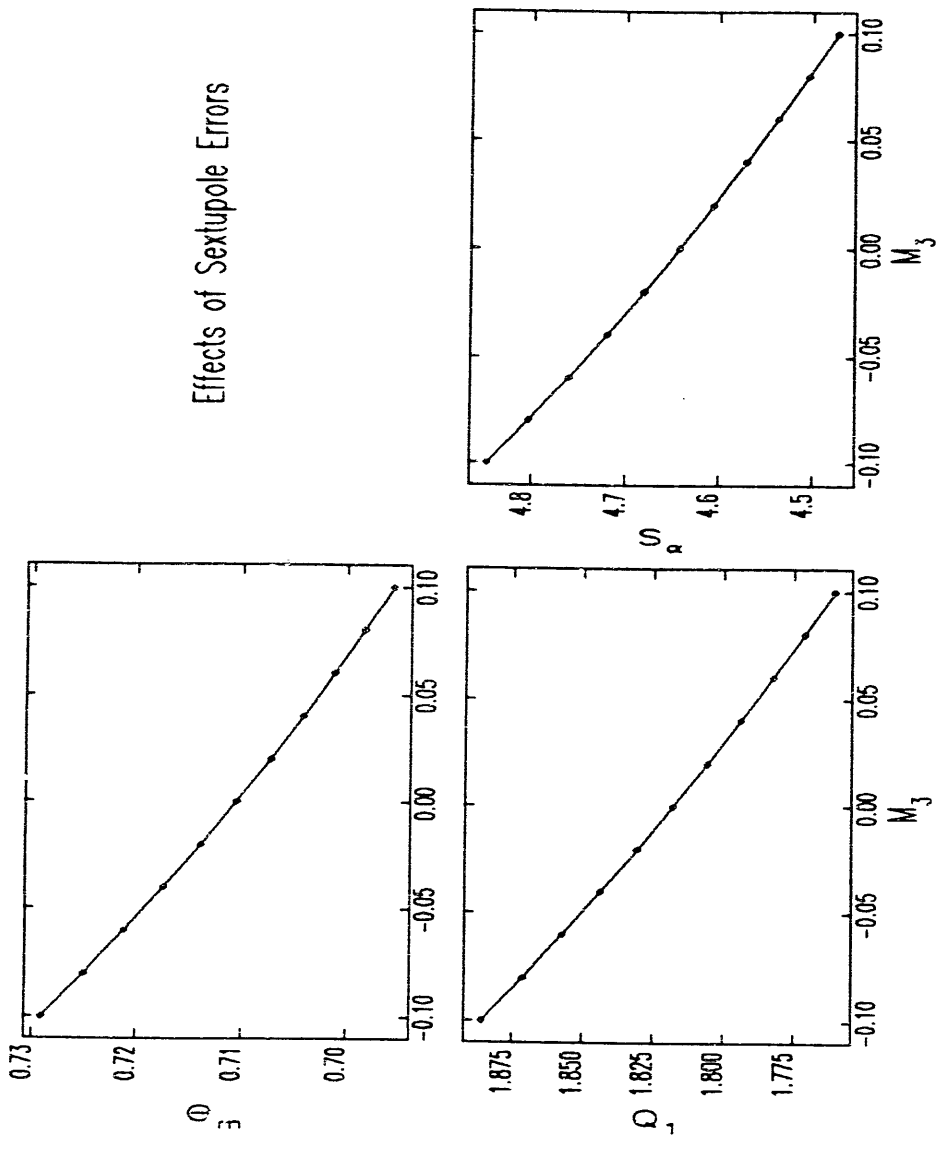

Figure 3.10: Effects of Sextupole Errors—Part 1

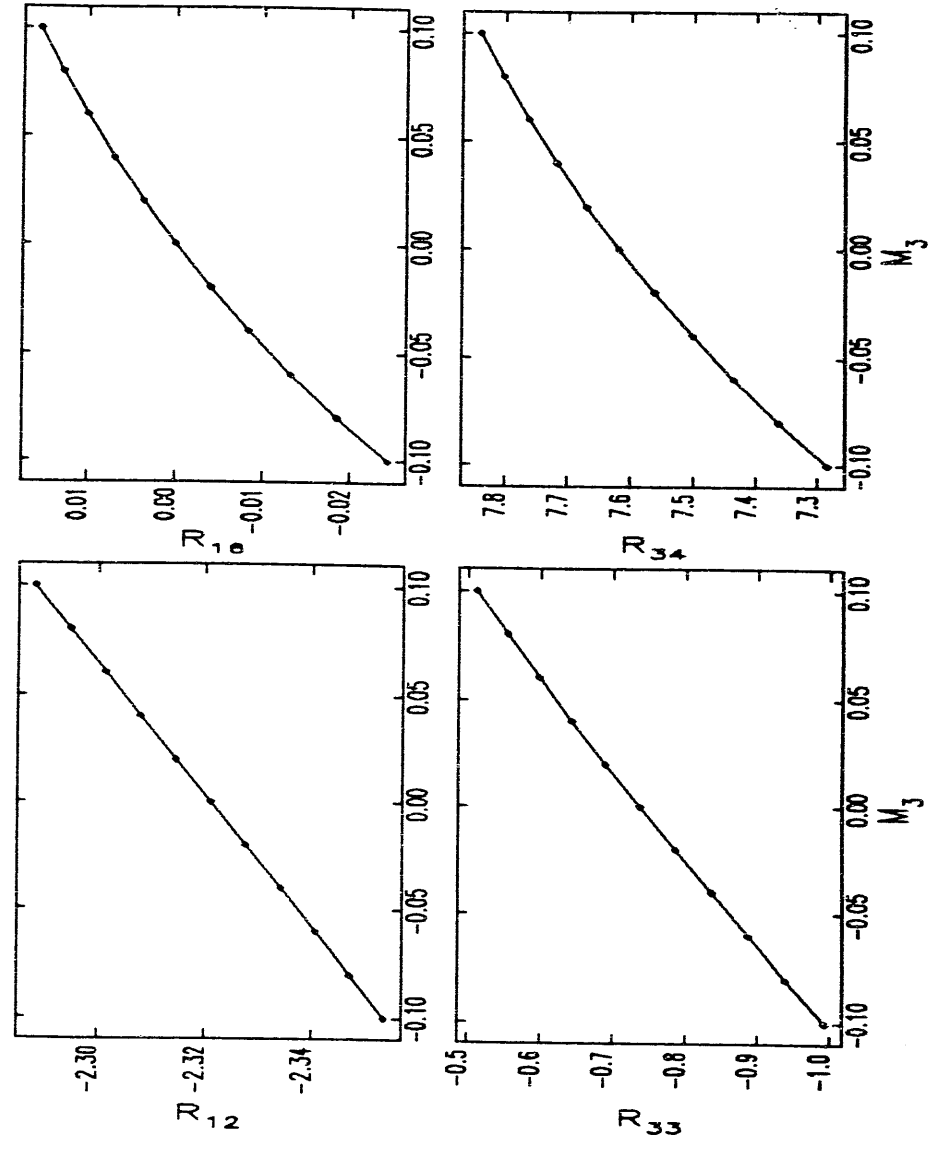

Figure 3.11: Effects of Sextupole Errors—Part 2

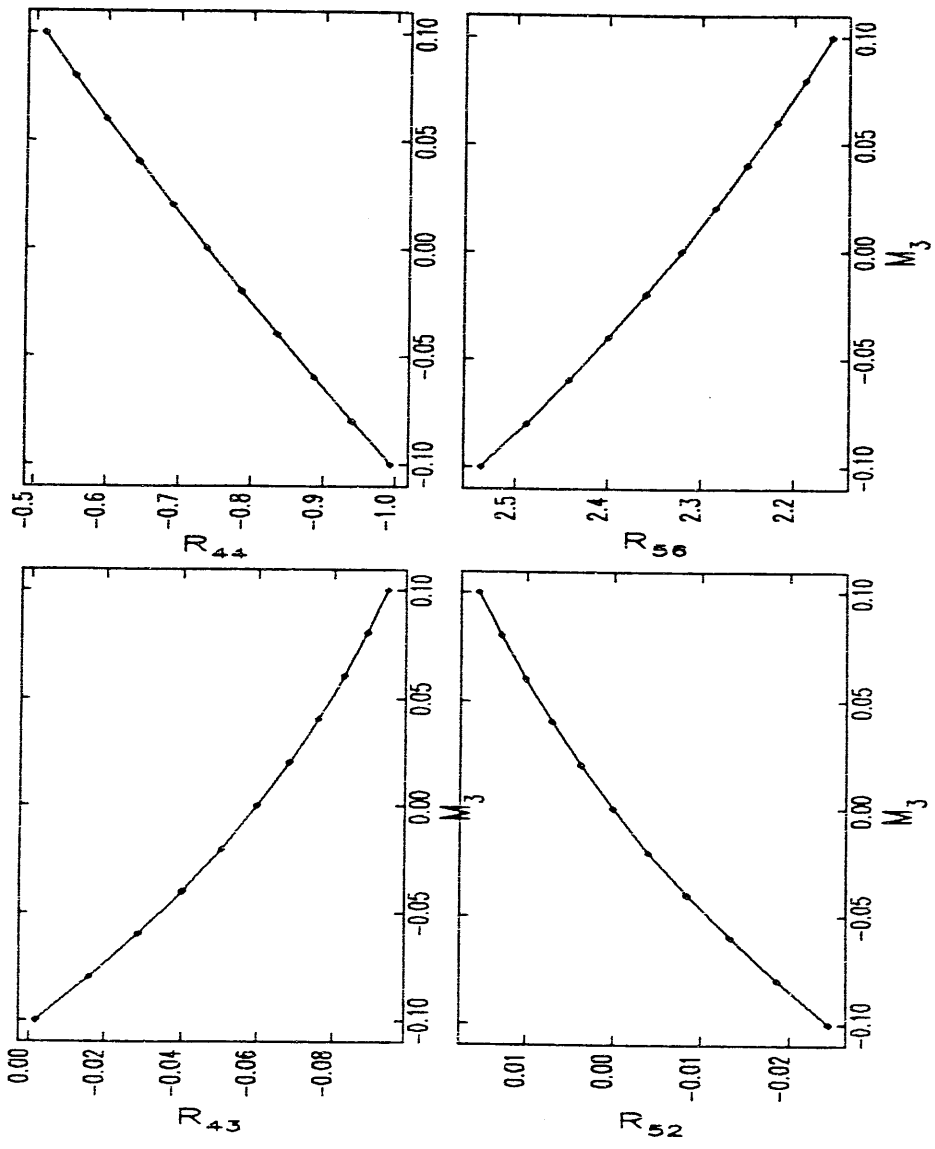

Figure 3.12: Effects of Sextupole Errors—Part 3

gradient vs q_1 for $q_2 = 0$, while figure 3.14 shows the measured gradient vs q_2 for $q_1 \approx 10$mm. The deviation of the field from a perfect quadrupole was dominated by perturbations from the beam aperture or "hole" cut in the magnetic mirror-plate of the alpha magnet. This can be appreciated by comparing Figure 3.13 with Figure 3.2, which shows the gradient before the hole was cut. This hole is, of course, necessary in order to get the beam into and out of the magnet. I have found that the field error in the $q_3 = 0$ plane is well approximated by

$$\Delta B_3 = g(K - Ee^{-q_1/d})F(q_2), \tag{3.181}$$

where K and E are positive constants, d is a decay constant for the field error, and F is the function

$$F(q_2) = \begin{cases} 1 & |q_2| < W_1 \\ \frac{W_2 - q_2}{W_2 - W_1} & W_1 \le |q_2| < W_2. \\ 0 & W_2 \le |q_2| \end{cases} \tag{3.182}$$

W_1 and W_2 are constants characterizing the width of the field perturbation in q_2. W_1 is roughly equal to the width of the hole in the midplane.

Fits to the data in the Figures give

$$\begin{aligned} K &= 0.054\text{cm} \quad E = 0.39\text{cm} \quad d = 0.72\text{cm} \\ W_1 &= 2.0\text{cm} \qquad\qquad W_2 = 3.6\text{cm}. \end{aligned} \tag{3.183}$$

Using Maxwell's curl equation, one can find an approximation to the full error field:

$$\frac{\alpha}{g}\Delta\mathbf{B} = E\alpha\left\{-\hat{\mathbf{q}}_1\frac{Q_3}{\alpha d}e^{-\frac{q_1}{\alpha d}}F(\frac{Q_2}{\alpha}) - -\hat{\mathbf{q}}_2\frac{Q_3}{\alpha d}e^{-\frac{q_1}{\alpha d}}F'(\frac{Q_2}{\alpha}) + \hat{\mathbf{q}}_3(1 - \frac{K\alpha}{E\alpha})e^{-\frac{q_1}{\alpha d}}F(\frac{Q_2}{\alpha})\right\} \tag{3.184}$$

(The possibility of dipole fields in the $\mathbf{q}_2$ and $\mathbf{q}_1$ directions can be eliminated by symmetry and by assuming that there are no rotated multipole fields present, respectively.) The constants K, E, and d occur in equation (3.184) only when multiplied by α. Similarly, the constants W_1 and W_2 occur only in the combinations $W_1\alpha$ and $W_2\alpha$, as seen from the definition of F. Any given magnet has fixed values for K, E, d, W_1, and W_2, while α will vary was the gradient of the magnet and the beam momentum are varied.

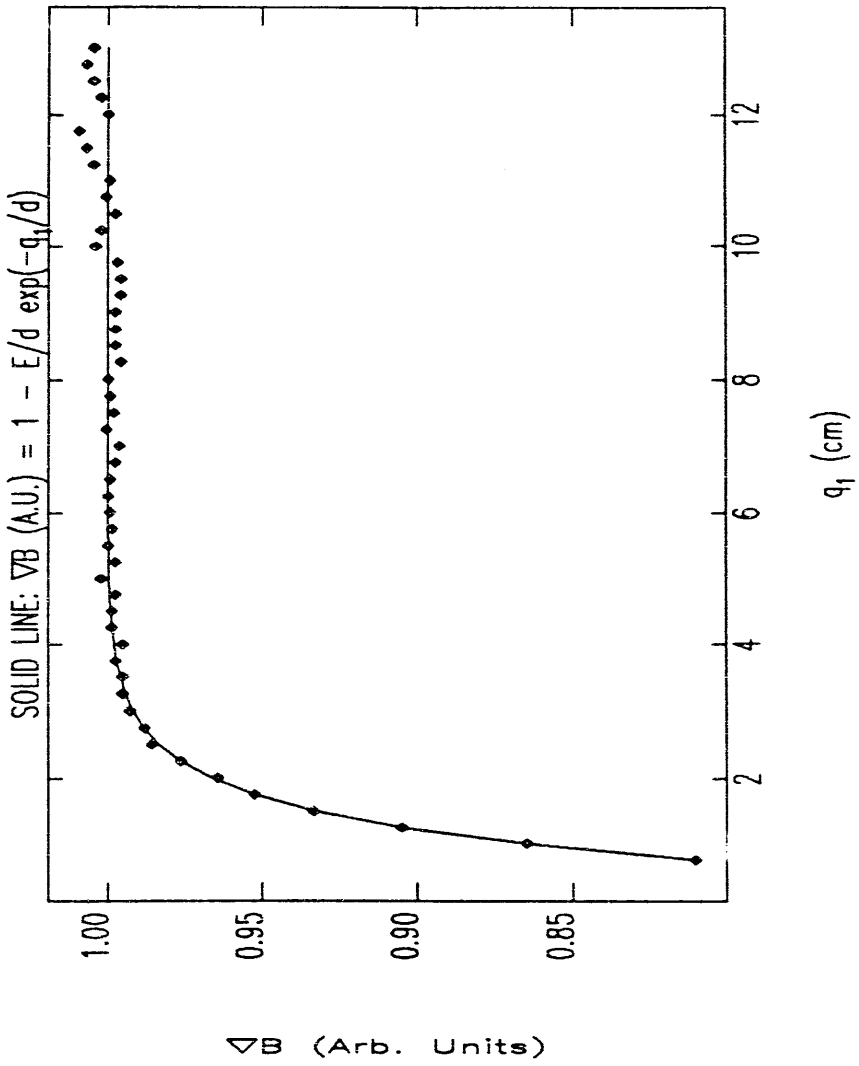

Figure 3.13: Hole-Induced Gradient Errors vs q_1

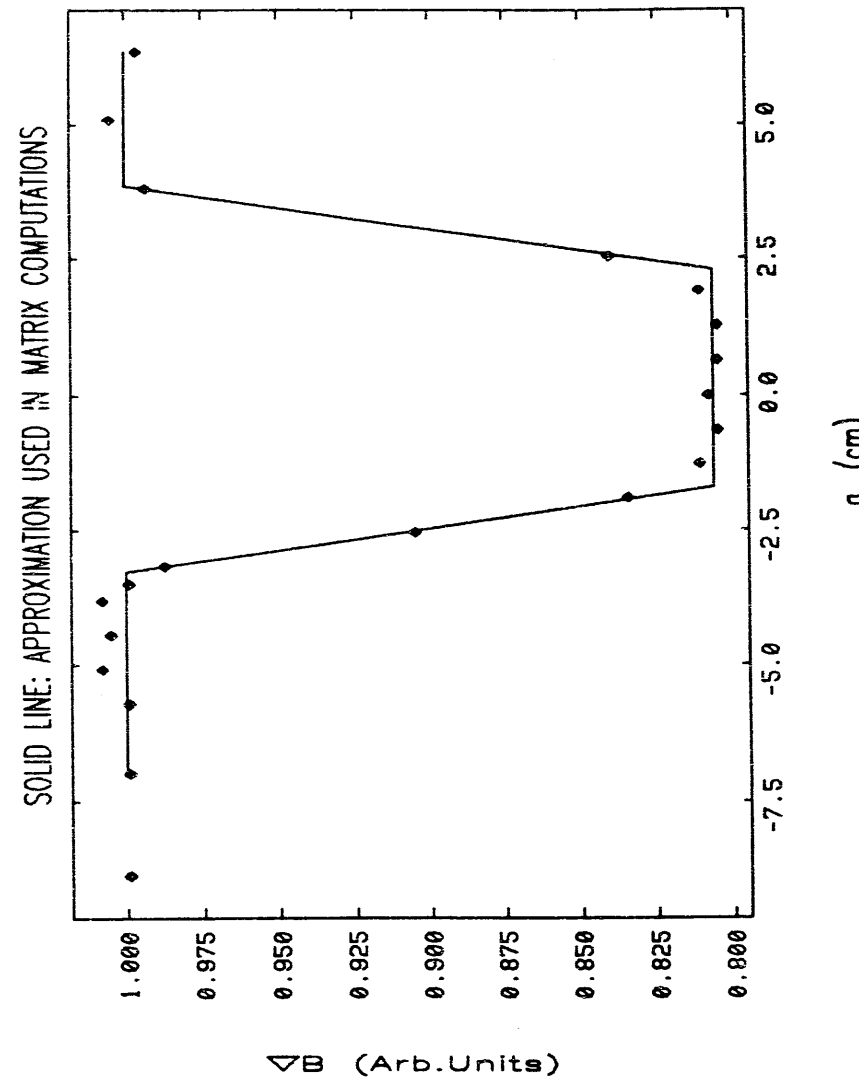

Figure 3.14: Hole-Induced Gradient Errors vs q_2

To show how it is possible to find matrices for the hole-induced error in a given alpha-magnet as a function of α it is convenient to define dimensionless error field parameters

$$\begin{aligned} \tilde{K} &\equiv \alpha K \qquad \tilde{E} \equiv \alpha E \qquad \tilde{d} \equiv \alpha d \\ \tilde{W}_1 &\equiv \alpha W_1 \qquad\qquad \tilde{W}_2 \equiv \alpha W_2. \end{aligned} \tag{3.185}$$

and a function corresponding to F

$$\tilde{F}(Q_2) = \begin{cases} 1 & Q_2 < \tilde{W}_1 \\ \frac{\tilde{W}_2 - Q_2}{\tilde{W}_2 - \tilde{W}_1} & \tilde{W}_1 \le Q_2 < \tilde{W}_2. \\ 0 & \tilde{W}_2 \le Q_2 \end{cases} \tag{3.186}$$

The hole-induced error field (equation (3.184)) is expressible as

$$\frac{\alpha}{g}\Delta\mathbf{B} = \tilde{E}\left\{-\hat{Q}_1\frac{Q_3}{\tilde{d}}e^{-\frac{Q_1}{\tilde{d}}}\tilde{F}(Q_2) + -\hat{Q}_2\frac{Q_3}{\tilde{d}}e^{-\frac{Q_1}{\tilde{d}}}\tilde{F}'(Q_2) - \hat{Q}_3(1 - \frac{\tilde{K}}{\tilde{E}})e^{-\frac{Q_1}{\tilde{d}}}\tilde{F}(Q_2)\right\}, \tag{3.187}$$

which is formally independent of α, as desired for insertion into equation (3.176).

serrors takes E, K, d, W_1, and W_2 as input, and computes the matrices as a function of a variable M, where $\tilde{E} = M * E$, $\tilde{K} = M * K$, etc. Clearly, by choosing the scaled matrices for $M = \alpha$ and scaling them according to equation (3.93) with $\alpha_o = \alpha$, one obtains the matrices for the magnet with errors for a given value of α.

While one chooses the value of M is equal the α value of interest, the reader should not make the mistake of concluding that **serrors** is varying α, or calculating matrices at a given value of α. **serrors** is scaling the spatial extent and magnitude of the error field in *scaled* coordinates, and calculating the matrices for the scaled equation of motion in the presence of these error fields. By choosing $M = \alpha$, one obtains matrices that correspond to a certain trajectory size relative to the fixed spatial extent of the error fields. Another way to use these **serrors** results is to view M as a quantity related to the size of the beam-hole, in which case $M \neq \alpha$.

Figures 3.15 through 3.17 show the effects of hole-induced errors on $\hat{Q}_1$, S_α, θ_m, and strongly-affected R-matrix elements, as calculated by **serrors**. Typical values of α for the SSRL magnet and RF gun are between 0.12cm and 0.18cm. Note the large effect on the vertical plane, similar to that seen for sextupole errors. Experiments show that the vertical plane R matrix deviates significantly from that for an ideal

magnet, a subject to which I will return in the next section. Running experience shows that injection angle corrections of 10-20 mrad are needed, with the sign such as to make θ smaller. It is unclear, however, what part of this is due to field errors and what part is required by alignment errors. The real value of these calculations is to evaluate the magnitude of the effects of such errors, to see whether the injection angle correction required for realistic error levels is feasible or not.

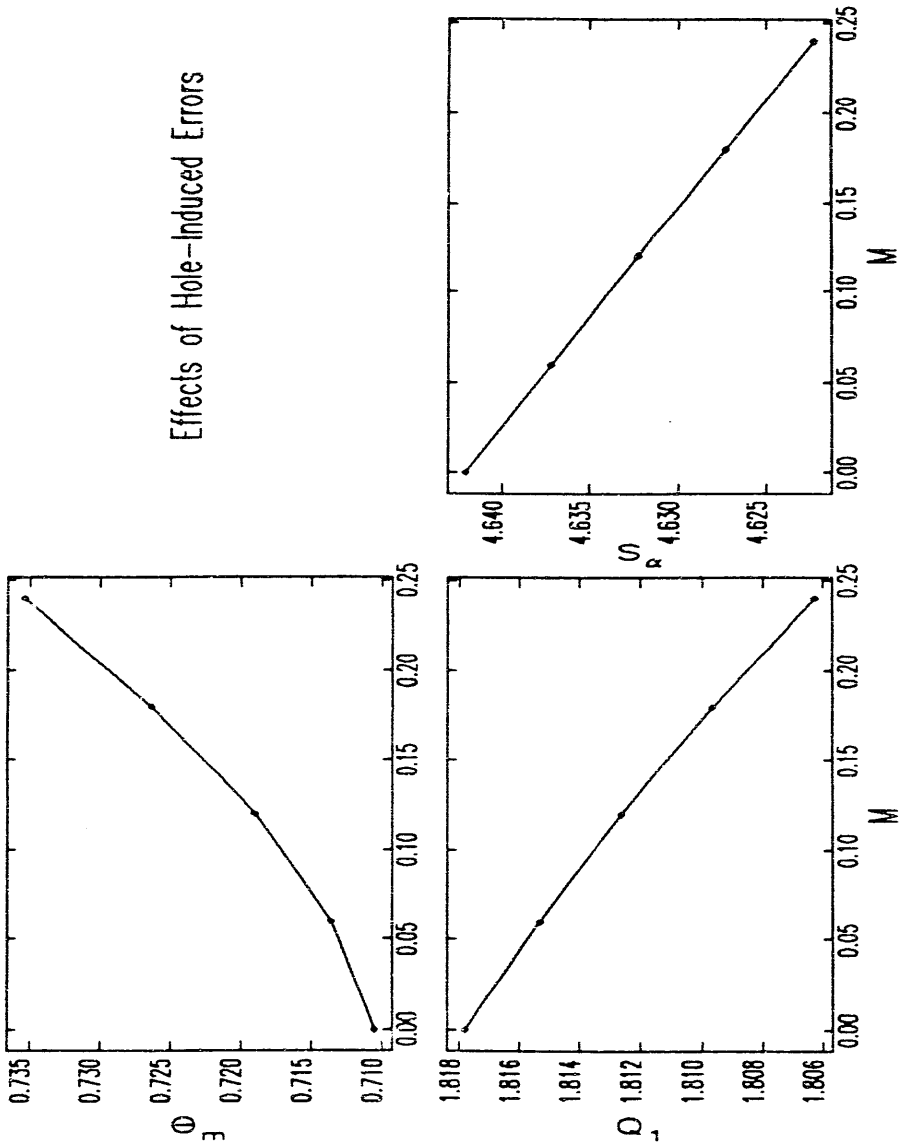

Figure 3.15: Effects of Hole-Induced Gradient Errors—Part 1

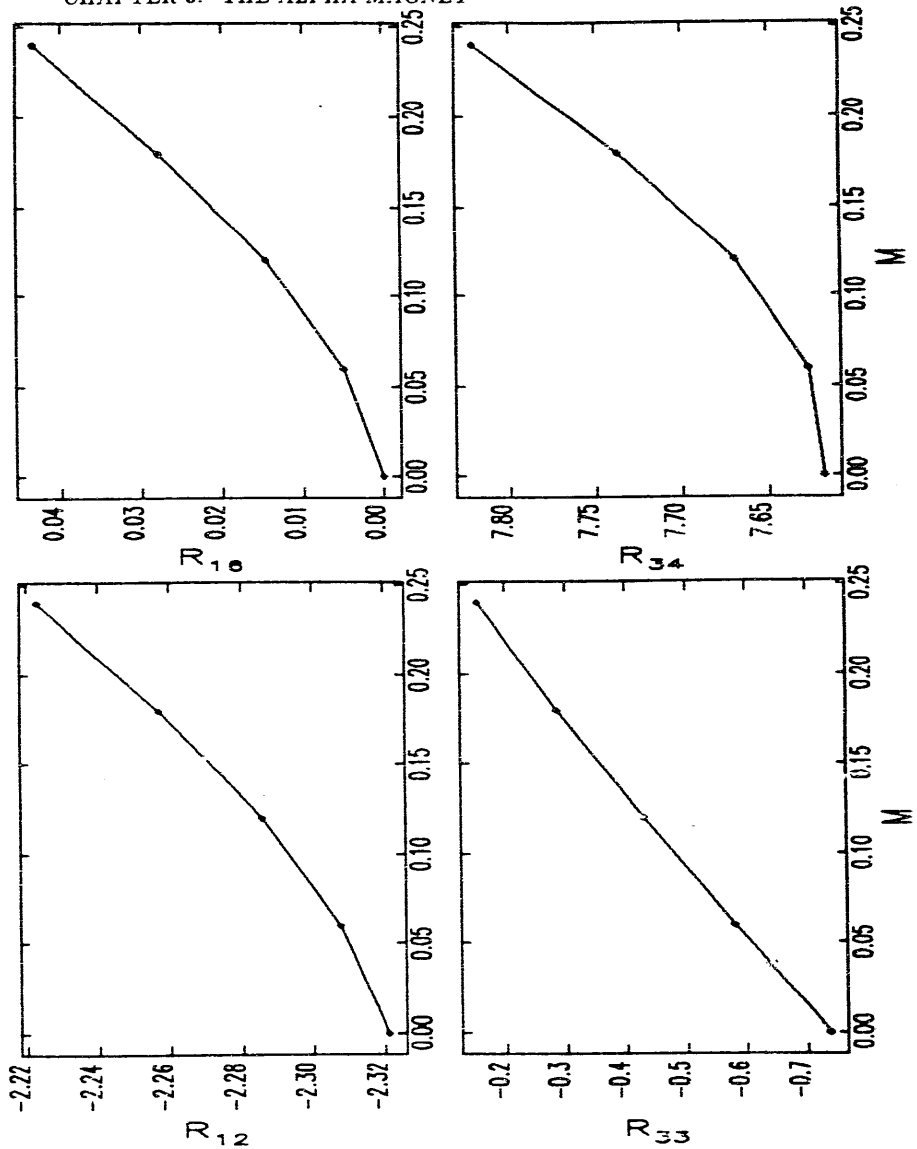

Figure 3.16: Effects of Hole-Induced Gradient Errors—Part 2

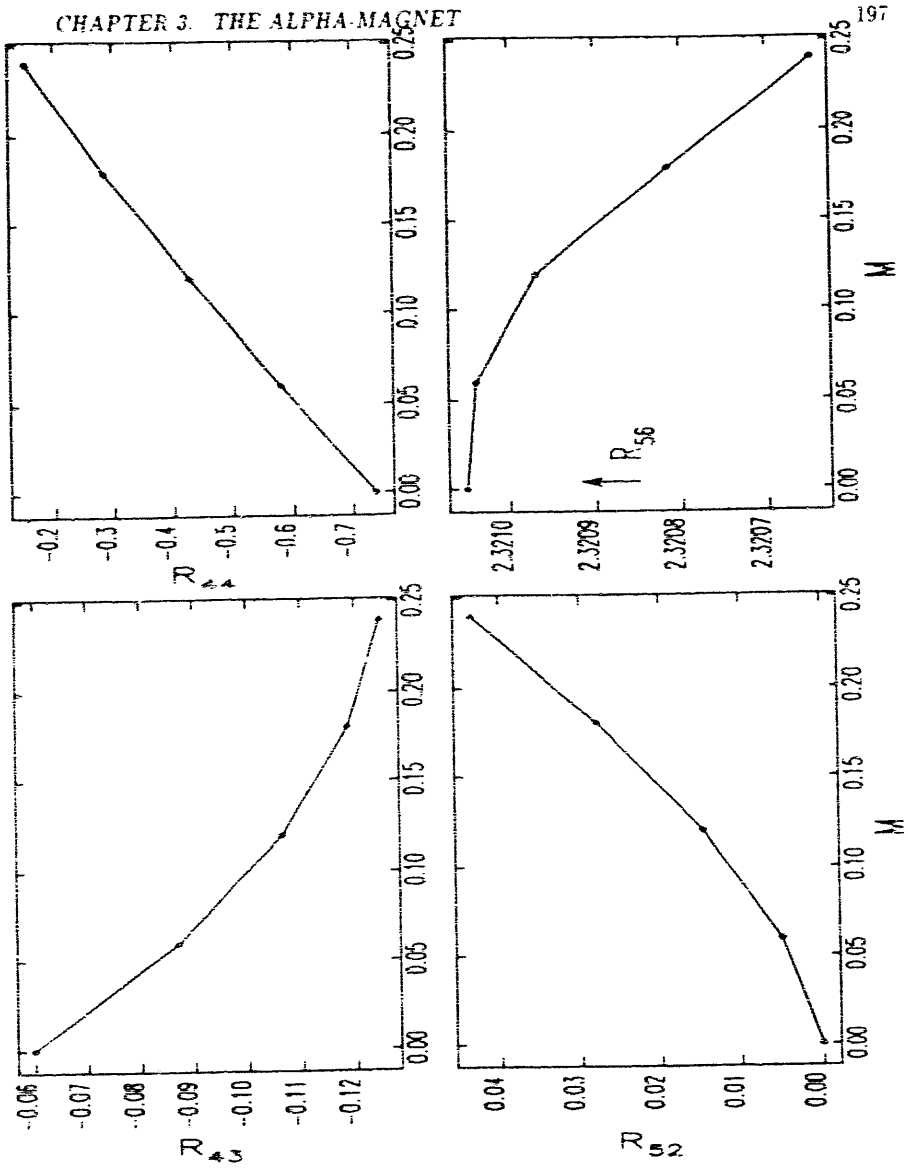

Figure 3.17: Effects of Hole-Induced Gradient Errors—Part 3

3.7 Alpha-Magnet Beam-Optics Experiments

A commonly-made measurement on beam-transport systems is that of determining the transverse centroid offset of the beam downstream from a steering magnet as a function of the angular kick imparted by that steering magnet. The linear term of a fit to the offset vs kick angle gives the r_{12} (or r_{34} for a vertically steering magnet) matrix element for transport from the steering magnet to the place where the centroid position is measured.

The Gun-to-Linac transport line (see Chapter 5) has horizontal and vertical steering just before the alpha-magnet, and a phosphorescent screen downstream of the alpha-magnet (the "chopper-screen", since it is part of the chopper tank). There is also a phosphorescent screen inside the alpha-magnet (the "alpha-magnet screen") that intercepts the beam when the alpha-magnet is turned off. These phosphorescent screens are viewed via closed-circuit TV. In addition, a Lecroy 9450 digital oscilloscope is available to digitize the TV scan, permitting accurate measurement of both horizontal and vertical beam positions. All that is required is to calibrate the TV sweep using features on the screens for which the positions are known (e.g., the edges of the screen).

I will let L_1 denote the distance from the center of this steering magnet, known as GTL_CORR2, to the "crossing-point" of the alpha-magnet ($q_1 = q_2 = q_3 = 0$). Also, L_2 and L_3 denote, respectively the distance from the crossing-point to the alpha-magnet screen, and from the crossing-point to the screen after the alpha-magnet. L_1 is found to be 117 mm, and L_3 to be 459 mm, where I use values from updated engineering drawings, checked by my own measurements with a ruler. The distance from the crossing-point to the screen in the alpha-magnet 200 ± 10mm, with the large uncertainty being due to the way the screen is held inside the alpha-magnet on long, easily-bent copper tubes.

3.7.1 Characterization of the Steering Magnet

I performed magnetic measurements on GTL_CORR2 with a quadrupole and an alpha-magnet-simulating iron plate in the proper positions relative to GTL_CORR2. The magnetic field as a function of longitudinal position z is shown in figure 3.18.

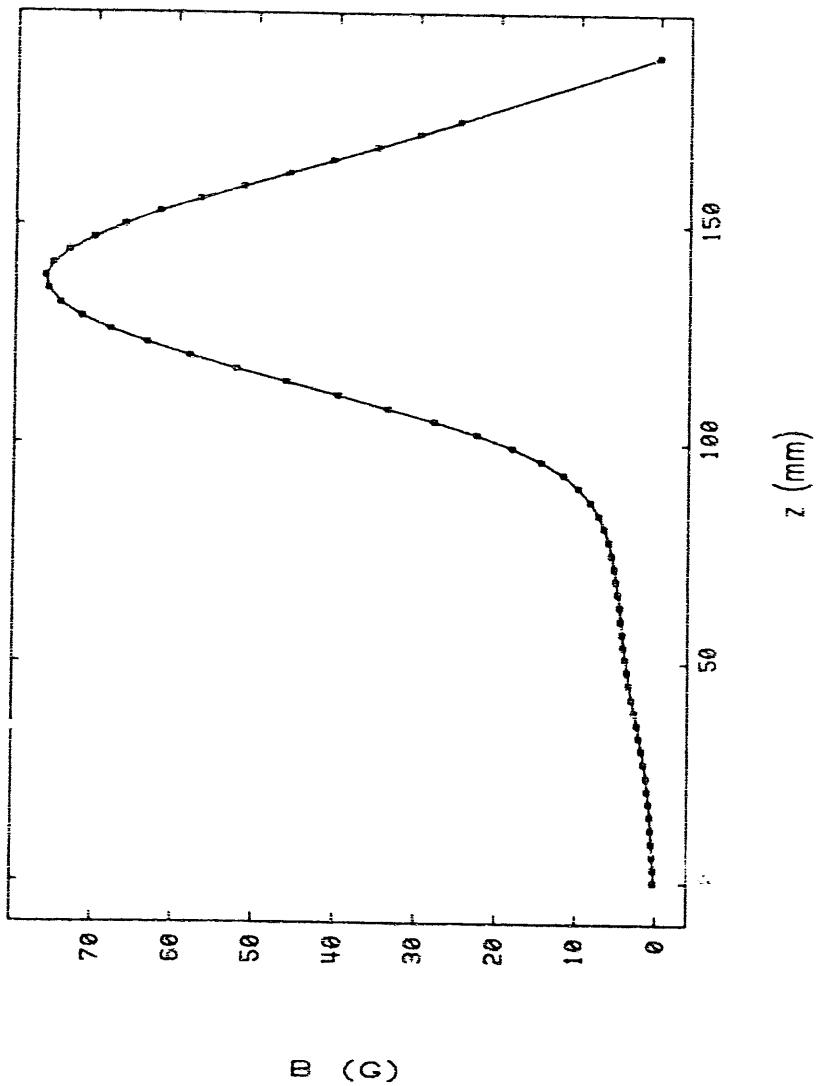

Figure 3.18: Magnetic Measurements for GTL_CORR2

and indicates that the equivalent angular kick for a zero-length steering magnet is 18mm ahead of the geometric center of the magnet (which coincides with the peak of the magnetic field vs z). Hence, GTL_CORR2 can be simulated by a zero-length steering magnet that is $L_1 + 18$mm from the alpha-magnet crossing point. I will thus let $L_1 \rightarrow L_1 + 18$mm, and treat GTL_CORR2 as a zero-length deflector.

Because GTL_CORR2 is in close proximity to both the alpha-magnet and the immediately preceding quadrupole, it is advisable to check that the calibration of angular deflection vs driving current (obtained from magnetic measurements) is correct. This was done using the alpha-magnet screen, since the transport to this screen from GTL_CORR2 is described by a simple drift-space matrix:

$$r^{(2)} = \begin{pmatrix} 1 & L_1 - L_2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & L_1 - L_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.188}$$

In this section, I will use $r^{(i)}$ to represent the r-matrix from GTL_CORR2 to point i, where i is 1, 2, or 3 for the crossing-point, alpha-magnet screen, or chopper-screen, respectively. I leave off the dispersive and path length elements to shorten the notation.

This check was carried out using the magnetic measurements to set GTL_CORR2 to a series of nominal horizontal (or vertical) deflection angles, $\theta_{nom,i}$, and measuring the resulting horizontal (or vertical) displacement, x_i, at the chopper screen. A linear fit to x_i vs $\theta_{nom,i}$ gives the nominal value of $r_{12}^{(nom)}$, uncorrected for errors in the deflection angle. Since it is that $r_{12}^{(2)} = r_{34}^{(2)} = 0.337 \pm 0.010$mm/mrad (this is just $L_1 - L_2$) the *actual* angular deflection is readily calculated, giving

$$\theta_{act} = \theta_{nom} \frac{r_{12}^{(nom)}(\text{mm})}{0.337 \pm 0.010\text{mm/mrad}} \tag{3.189}$$

Linear least-squares fits to the data from these experiments gave $r_{12}^{(nom)} = 0.305 \pm 0.006$mm/mrad and $r_{34}^{(nom)} = 0.322 \pm 0.015$mm/mrad, from which I conclude that

$$\theta_{x,act} = \theta_{x,nom} 0.91 \pm .03, \tag{3.190}$$

and

$$\theta_{y,act} = \theta_{y,nom} 0.96 \pm .05. \quad (3.191)$$

3.7.2 Comparison of Experimental Results and Theory

Having corrected the calibration of GTL_CORR2, I then did a series of measurements of $r_{12}^{(3)}$ and $r_{34}^{(3)}$ for various alpha-magnet gradients for constant beam momentum. For these experiments, the low-energy scraper inside the alpha-magnet was set to allow only about $\pm 5\%$ momentum spread through, to lessen any possible ambiguity about what the momentum of the particles seen on the chopper screen was. Spectrum measurements allow the determination of the median momentum of the particles let through, and this quantity was used as the effective momentum of the beam centroid. Table 3.10 summarizes the results.

In order to compare these results to theory, it is necessary to use **serrors**-calculated scaled matrices for the appropriate values of the error parameter M (i.e., M = α, for hole-induced errors), to scale these matrices to the values of α listed in the table, and to finally concatenate these matrices with drift space matrices:

$$r = d(L_3)A^{-1}(\alpha)R(M)A(\alpha)d(L_1). \quad (3.192)$$

where d(L) represents the matrix for a drift space of length L. Table 3.11 gives the results of this procedure, listing the r_{12} and r_{34} values corresponding to each of the cases in Table 3.10. Also listed for comparison are the values for a perfect alpha-magnet. These results are displayed in figures 3.19 and 3.20.

As seen from Figure 3.20, the $r_{34}^{(3)}$'s are very sensitive to errors, hence the agreement seen here may be fortuitous. In the same vein, some disagreement is hardly unexpected.

With the exception of the anomalous point at $\alpha = 0.166$, all of the measured r_{12}'s are 5–10% smaller than the theoretical values for the alpha-magnet with errors. The first explanation of the discrepancies in the horizontal plane one might entertain is that the momentum (or, equivalently, the alpha-magnet gradient calibration) is in error by 5–10%. This, however, would not explain the discrepancies observed.

First, the correction of the calibration of GTL_CORR2 would eliminate any effects of momentum errors on deflection angle. Hence, any momentum errors would come into play only through the alpha-magnet. However, as seen from the slope of r_{12} vs α for the theoretical data in Figure 3.19, a very large momentum error would be required to explain the observed discrepancy. For the point with $\alpha = .121$, for example, one would need to postulate a momentum error of about a factor of two, which I do not consider remotely possible. A remaining possibility is that the calibration of GTL_CORR2 is in error, due to inaccurate knowledge of the position of the alpha-magnet screen. A 5–10% error in this calibration would require a 15–30mm error in the position of the screen. This is not inconceivable, given the lack of precision in the screen assembly.

Table 3.10: **Alpha Magnet r_{12} and r_{34} Measurements**

gradient (G/cm)	$< p >$ MeV/c	α 1/cm	$r_{12}^{(3)}$ mm/mrad	$r_{34}^{(3)}$ mm/mrad
255.1 ± 1.3	3.00 ± 0.04	0.160 ± 0.001	-0.743 ± 0.014	0.006 ± 0.005
202.9 ± 1.0	2.71 ± 0.04	0.150 ± 0.001	-0.672 ± 0.012	0.070 ± 0.002
172.8 ± 0.9	2.83 ± 0.04	0.135 ± 0.001	-0.716 ± 0.012	0.120 ± 0.004
149.1 ± 0.7	2.81 ± 0.04	0.126 ± 0.001	-0.730 ± 0.016	0.149 ± 0.004
129.9 ± 0.6	2.80 ± 0.04	0.118 ± 0.001	-0.739 ± 0.018	0.180 ± 0.005

Table 3.11: **Calculated Alpha-Magnet r_{12} and r_{34}**

	alpha-magnet with errors		perfect alpha-magnet	
α 1/cm	$r_{12}^{(3)}$ mm/mrad	$r_{34}^{(3)}$ mm/mrad	$r_{12}^{(3)}$ mm/mrad	$r_{34}^{(3)}$ mm/mrad
0.160	-0.736	0.171	-0.739	-0.021
0.150	-0.745	0.197	-0.749	0.014
0.135	-0.763	0.244	-0.766	0.076
0.126	-0.775	0.279	-0.778	0.120
0.118	-0.788	0.315	-0.790	0.164

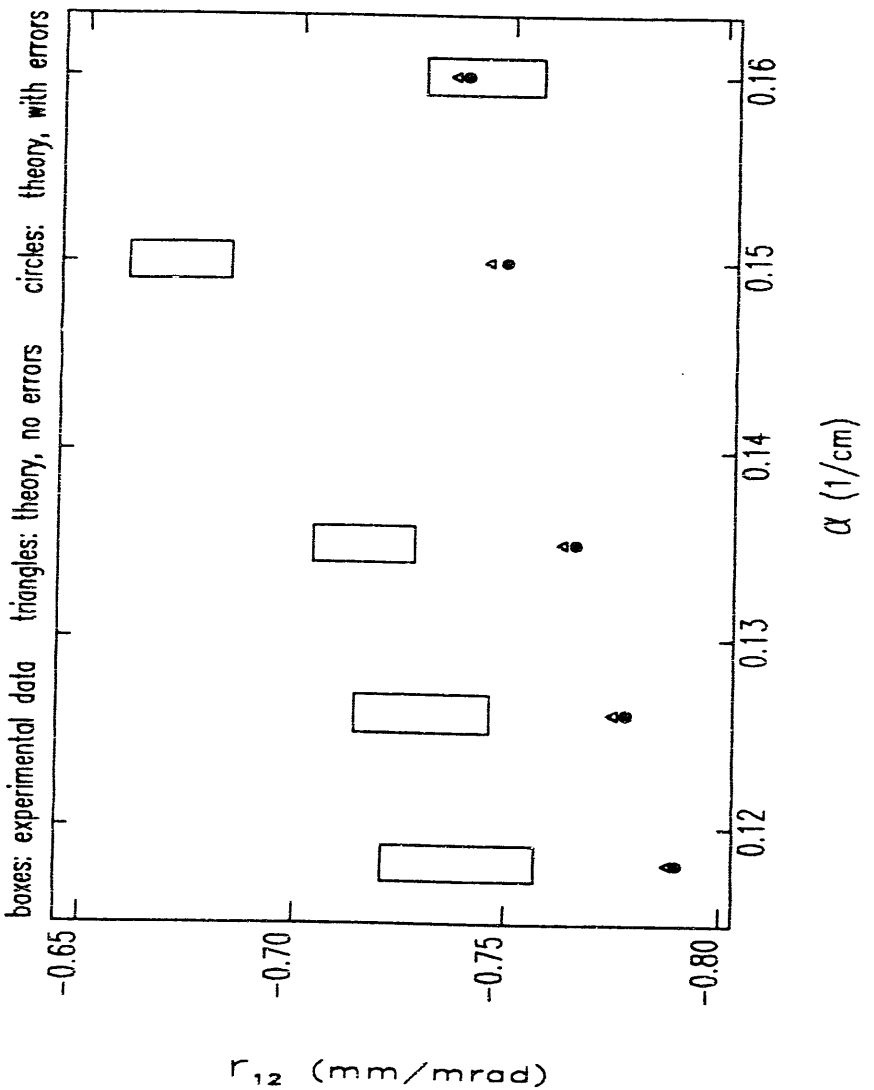

Figure 3.19: Measured and Theoretical Alpha-Magnet r_{12}'s

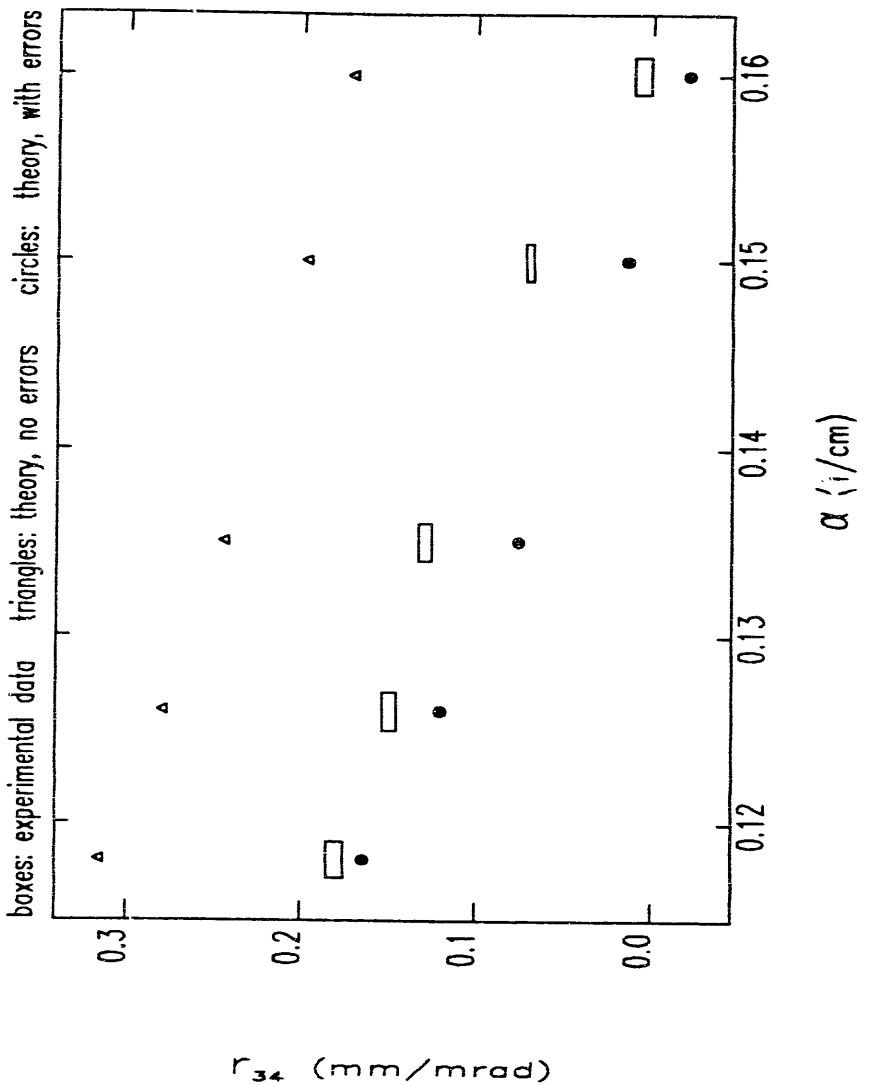

Figure 3.20: Measured and Theoretical Alpha-Magnet r_{34}'s

Chapter 4

Longitudinal Dynamics

In Chapter 2, I described the longitudinal phase-space distribution of the RF gun beam, and indicated that this phase-space ill-suits the beam to direct injection into a S-Band linear accelerator section (here-after referred to simply as "the linac"). In this chapter, I will show why this is so, and how the gun longitudinal phase-space may be transformed into something that is amenable to further acceleration. At issue is the need for a small fractional energy spread, which is required for efficient transport through a subsequent beamline, use as the drive for FELs, and other applications. I shall also show how the rather long (25 ps or so) bunch at the end of the gun can—at least in the absence of excessive errors and space-charge effects—be compressed to a very short 1-2 ps bunch, thus promising the potential of very high peak currents, something that is desirable in FEL applications, among others.

Discussion of the transformation of the gun longitudinal phase-space cannot take place without an understanding of the longitudinal dynamics of electrons in magnetic systems and linear accelerators. I will first discuss longitudinal dynamics in linear accelerators, and in particular how one can predict the longitudinal phase-space at the end of a finite-length accelerator when starting with a beam that is not fully relativistic. This discussion will show why the RF gun beam is unsuited to direct injection into the linac.

I will then discuss how magnetic beamline elements can be used to alter a beam's longitudinal phase-space. Using results from Chapter 3, I will demonstrate that an alpha-magnet has advantages in such an application. I will present the results of

optimized alpha-magnet-based bunch compression, with inclusion of detailed longitudinal dynamics calculations, and show how this achieves significantly better results compared to the first-order method of simply injecting the shortest possible bunch into the linac.

Finally, I will present results that include consideration of aberrations in the gun-to-linac transport line, and use these results to compare the SSRL preinjector to other projects.

4.1 Longitudinal Dynamics in Linear Accelerators

There is extensive literature on longitudinal dynamics in linear accelerators [56, 55, 41, 69]. Rather than attempting to duplicate that work here, I shall simply make use of some of the results. In particular, I shall use the commonly-made assumption [41, 55] that the longitudinal electric field of a traveling-wave linear accelerator may be approximated by the first space-harmonic.

$$E_z(z,t) = E \cdot \cos(kz - \omega t), \quad 0 \le z \le L, \tag{4.1}$$

where k is the propagation constant, $\omega/(2\pi)$ is the RF frequency, L is the length of the structure (which starts at z=0), and E is a constant. For a velocity-of-light structure such as the SLAC constant-gradient structure used for the SSRL linac, $\omega = kc$, c being the speed of light. While the actual field contains components with propagation constants $k_n = k - \frac{2\pi n}{p}$, where n is an integer and p is the periodic length of the structure [56], these components have phase-velocities ω/k_n less than the speed of light, and hence a relativistic particle will not remain in phase with any but the first space-harmonic. Because of this, the higher space-harmonics impart no net energy to a relativistic beam. Traveling wave accelerators are specifically designed to have small amplitudes in the non-synchronous space-harmonics, since these carry away RF power without contributing to acceleration [56].

The equations of motion for an electron in the presence of this field are found from the Lorentz force. For $0 \le z \le L$,

$$\frac{dp}{dt} = -\frac{eE}{m_e c}\cos(kz - \omega(t - t_o) + \phi_o) \tag{4.2}$$

$$\frac{dz}{dt} = \frac{pc}{\sqrt{p^2+1}}. \tag{4.3}$$

where $p \equiv \beta\gamma$ is the momentum and ϕ_o is the initial RF phase for the fiducial particle, which enters the accelerator at $t = t_o$. I will assume that $-eE \ge 0$, so that the fiducial particle is accelerated for $\phi_o = 0$. If $\phi_o < 0$, then the fiducial particle is "behind the crest", meaning that if it is sufficiently non-relativistic, it may fall further behind; if it falls back far enough (or if $\phi_o < -\pi/2$) the particle will be decelerated. Similarly,

if $\phi_o > 0$, the fiducial particle is "ahead of the crest", meaning that if it is not fully relativistic, it may fall back to be nearer the maximum accelerating phase.

4.1.1 Approximate Treatment for Highly-Relativistic Particles

In general, these equations are unsolvable by analytic means, being coupled and non-linear. However, for $p^2 \gg 1$, the electrons are fully relativistic, and $z = c(t - t_i)$, where t_i is the time at which the particle enters the accelerator. In this case, the acceleration experienced by any particle is constant:

$$\frac{dp}{dt} = -\frac{eE}{m_e c}\cos(-\omega(t_i - t_o) + \phi_o). \tag{4.4}$$

The solution, for $t_i \le t \le t_i + L/c$,

$$p(t) = p(t_i) - (t - t_i)\frac{eE}{m_e c}\cos(-\omega(t_i - t_o) + \phi_o), \tag{4.5}$$

may be used to determine the final momentum after an accelerating section of length L:

$$p_f = p_i - \frac{eEL}{m_e c^2}\cos(-\omega(t_i - t_o) + \phi_o), \tag{4.6}$$

where $p_i = p(t_i)$ and $p_f = p(t_f)$.

Since the "useful" electrons out of the RF gun typically have $p \ge 4$, giving $\beta \ge 0.97$, this result is of more than academic interest. Though it is far from exact, it is a useful approximation, and aids the understanding of the detailed results.

The quantity $\phi_i = -\omega(t_i - t_o) + \phi_o$ is the initial phase for some particular particle, even if the particle is non-relativistic. However, for a particle that is initially fully relativistic, the initial phase is also the RF phase throughout the accelerator section. Hence, if it is desired to accelerate a bunch of relativistic particles with small final momentum spread, it is necessary to inject these particles into the accelerator over a sufficiently small time-interval. Suppose that the bunch initially has no momentum spread, and that $\phi_o = 0$ to obtain maximum acceleration of the fiducial particle. Assume further that the fiducial particle is at the center of the bunch, which has length δt. Then the spread in final momenta will be

$$\delta p = (p_f - p_i)(1 - \cos(\omega\delta t/2)), \tag{4.7}$$

where $p_f = p_i - \frac{eEL}{m_e c^2}$ is the final momentum of the fiducial particle. For $\omega\delta t \ll 1$, this implies a fractional momentum spread of

$$\frac{\delta p}{p_f} = \frac{p_f - p_i}{p_f}\frac{\delta\phi^2}{8}. \tag{4.8}$$

where $\delta\phi = \omega\delta t$ is the phase spread of the incoming beam. If the beam is accelerated to very high momentum relative to p_i, then this becomes

$$\frac{\delta p}{p_f} = \frac{\delta\phi^2}{8}. \tag{4.9}$$

and one sees that the final fractional momentum spread is, to first order, quadratic in the initial phase spread if one injects the fiducial (and central) particle at the crest. Hence, in order to obtain a small final momentum spread, one must inject a sufficiently short bunch into the accelerator:

$$\delta\phi \le \sqrt{8\left(\frac{\delta p}{p_f}\right)_{\text{desired}}}. \tag{4.10}$$

For the SSRL Injector, a fractional momentum spread of less than 0.5% was needed, to accommodate the acceptance of the synchrotron[26]. Hence, from this analysis one would conclude that an initial total phase-spread of less than about 12° is required, *if* one ignores the initial momentum spread in the RF gun beam. I shall show below that, however, that one cannot ignore the initial momentum spread, if one really desires such low final momentum spread. Note that injection with the central particle off the crest will only increase the final momentum spread for a beam with no initial momentum spread, while for a beam that has some initial, time-correlated momentum spread, injection off the crest can be used to compensate the initial momentum spread, as will be seen below.

4.1.2 Numerical Solution and the Contour Approach

Computer methods can easily solve equations (4.2) and (4.3) to high precision, so it is not necessary to attempt to find a solution that is valid for non-relativistic electrons. For the current project in particular, the input longitudinal phase-space

distribution is itself not amenable to analytical treatment, but is rather obtained from numerical simulations. Hence, I will move on to discuss computer-aided treatment of this problem.

For some of my computer studies, I employed another pair of equations[41], which are useful if longitudinal motion is one's only interest. Rather than start with the Lorentz equation, one starts with[31]

$$\frac{dm_e c^2\gamma}{dt} = -e\mathbf{v}\cdot\mathbf{E}. \tag{4.11}$$

and assumes the velocity to be parallel to electric field. One form of the resultant equations is (reference [41] gives these equations and a detailed discussion of them)

$$\frac{d\gamma}{d\zeta} = \mathcal{E}\cos 2\pi(\zeta - \tau) \tag{4.12}$$

$$\frac{d\tau}{d\zeta} = \frac{1}{\beta}, \tag{4.13}$$

where $\zeta = \frac{z}{\lambda}$, $\lambda = \frac{2\pi}{k}$, $\mathcal{E} = -\frac{eE\lambda}{m_e c^2}$, and $\tau = \frac{\omega t}{2\pi}$. The RF phase for any particle at z=0 is particle is $\phi = -2\pi\tau$, which is consistent with the convention I used above. In terms of normalized electric-field $\mathcal{E}$, the change in γ in a section of length L for an initially relativistic particle is

$$\Delta\gamma = \mathcal{E}\Delta\zeta = \mathcal{E}\frac{L}{\lambda} \tag{4.14}$$

There is no particular advantage to these equations over a similarly-scaled form of equations (4.2) and (4.3) for numerical work—I state them because I happened to use them in some of my computations. Specifically, I have written a computer program (`linac_cg`, where "CG" stands for "constant-gradient") that integrates equations (4.12) and (4.13) for a set of particles distributed on a grid over some region of initial (ϕ, p) space. The program computes the final momentum and phase for each particle, and displays the results in contour-plot form. From these, one can deduce the resultant momentum spread and phase spread for any particular injected bunch simply by finding which contours are intersected when the phase-space distribution for the injected bunch is overlayed on the contour graphs. Note that a different plot must be generated for each value of E. For the SLAC-type constant-gradient sections

used for the SSRL pre-injector[70], the nominal energy gain per section is given by[6]

$$\Delta\gamma = 20.4\sqrt{P(\mathrm{MW})}, \tag{4.15}$$

where P is the RF power to the section. Combining this with the previous equation, using L = 3.048m and λ = 0.105m, I obtain

$$\mathcal{E} = 0.703\sqrt{P(\mathrm{MW})}. \tag{4.16}$$

While nominal RF power per section for the SSRL Pre-injector Linac is 30 MW, the RF power to the first section is limited operationally to 20 to 25 MW. Since the energy gain scales only as the square-root of the momentum, the differences among these are relatively minor. For this reason, and for brevity, I present only the results for 20 MW RF power, and display these in Figures 4.1 and 4.2.

The horizontal axis for these graphs is the phase, ϕ_i, at which the particle is injected, while the vertical axis is the initial momentum, p_i. As before, $\phi_i > 0$ indicates injection ahead of the RF crest. The contours show lines in (ϕ_i, p_i) space of constant final momentum, p_f, or final phase, ϕ_f. The momentum contours are spaced by $\Delta p_f = 4.2$ and the phase contours by $\Delta\phi_f = 10°$. The labels for the contours are positioned so that the contour closest to the lower left corner of the first letter in the label is the one to which the label applies.

A bunch with an initial longitudinal phase-space distribution that matches a constant final momentum contour will be accelerated to zero momentum spread, and similarly for a bunch that matches a constant final phase contour. Regions where many lines occur in a small area indicate rapidly changing final parameters as a function of initial parameters. Regions where the contour lines are widely spaced indicate slowly changing final parameters as a function of initial parameters.

Examination of Figure 4.1 shows that, as expected from the above analysis, the final-momentum contours are most widely spaced for ϕ_i near zero. The region of widest contour spacing moves to positive ϕ_i as the initial momentum decreases because, for a bunch of non-relativistic electrons, injection at slightly positive ϕ_i results in the bunch center falling back toward $\phi_i = 0$ as the electrons gain energy. If such a bunch were injected at $\phi_i = 0$, it would fall back to $\phi_i < 0$ before reaching relativistic velocities, and as a result the bunch momentum would be decreased while its momentum spread would increase.

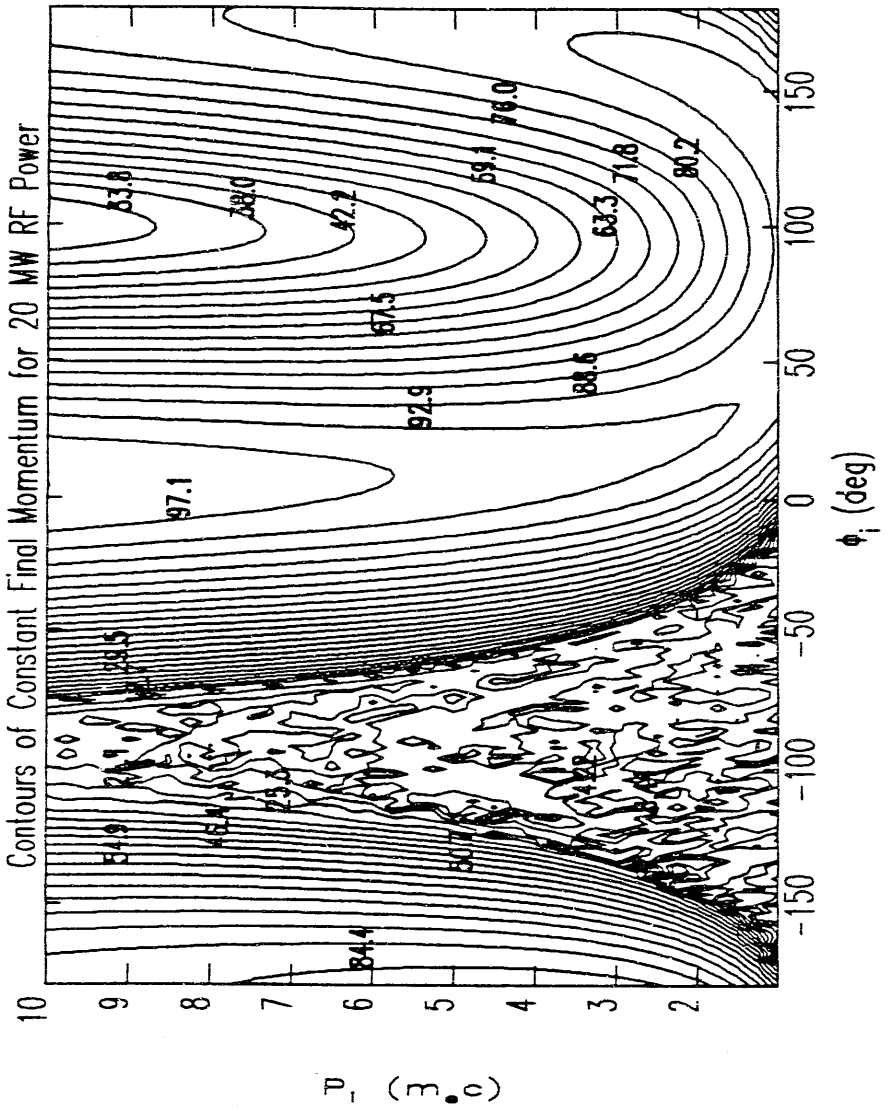

Figure 4.1: Constant Final Momentum Contours

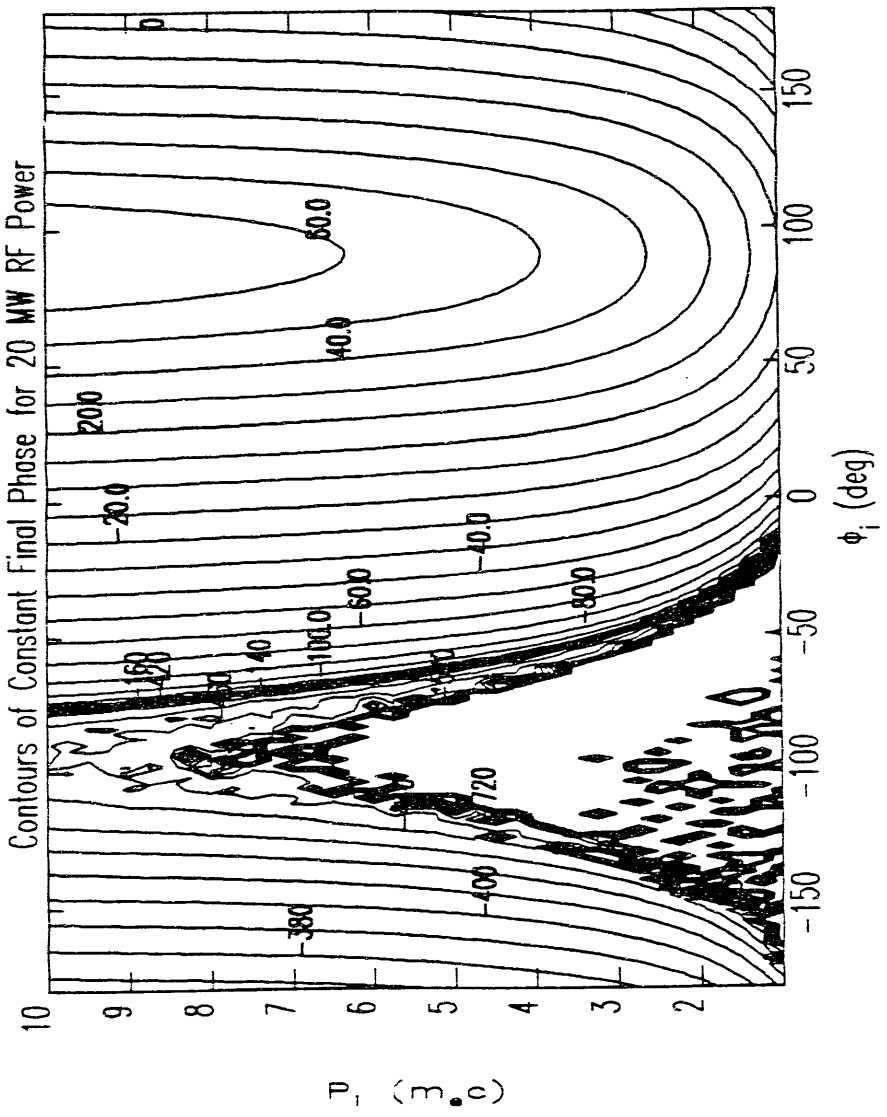

Figure 4.2: Constant Final Phase Contours

For a bunch centered at some phase $-55° \lesssim \phi_i < 0$, the smallest final momentum spread will be achieved when the higher-momentum particles come in behind the lower-momentum particles. Similarly, for a sufficiently high-momentum bunch centered on some phase $0 \lesssim \phi_i \lesssim 30°$, the smallest final momentum spread will be achieved when the higher-momentum particles come in ahead of the lower-momentum particles. In both cases, one can understand this by imagining that a sinusoidally-varying momentum change is simply being added to the initial momenta, as illustrated in Figure (4.3).

For the second of these regions, as the bunch center is moved to smaller momentum and/or larger ϕ_i, one sees another effect come into play. The slope of the constant final-momentum contours changes so that higher-momentum particles must be injected behind lower-momentum particles. In this regime, velocity variation is important. It is necessary to inject the higher-momentum particles so that they will catch up to the lower-momentum particles as the bunch travels down the accelerator. This bunching can contribute to small momentum spread, since once bunched the particles will travel the remainder of the accelerator section at the same phase (provided they are all relativistic by the time they are bunched), thus experiencing the same energy gain in the remainder of the section. (In the jargon of the field, one says that members of such a group of particles all have the same "asymptotic phase".)

The same velocity effect also occurs in the first of the regions mentioned in the paragraph before last, it simply does not cause a change in the slope of the contours, since the slope is required to be the same from both considerations of sinusoidal field variation and velocity variation in the bunch.

For $\phi_i \gtrsim 90°$, the slope of the constant final momentum contours changes again. In this region, higher-momentum particles must be injected first so that they are decelerated more than lower-momentum particles.

Centered around $\phi_i = -90°$ is a "chaotic" region, where the final momentum and phase of an injected particle depends strongly on the initial momentum and phase. Particles injected into this region are first decelerated, then accelerated again as they fall back relative to the traveling wave. Some of the particles injected here are back-accelerated, exiting the accelerator section at $z = 0$, while others finally exit at $z = L$ only after many cycles of acceleration and deceleration. As expected, the width of this

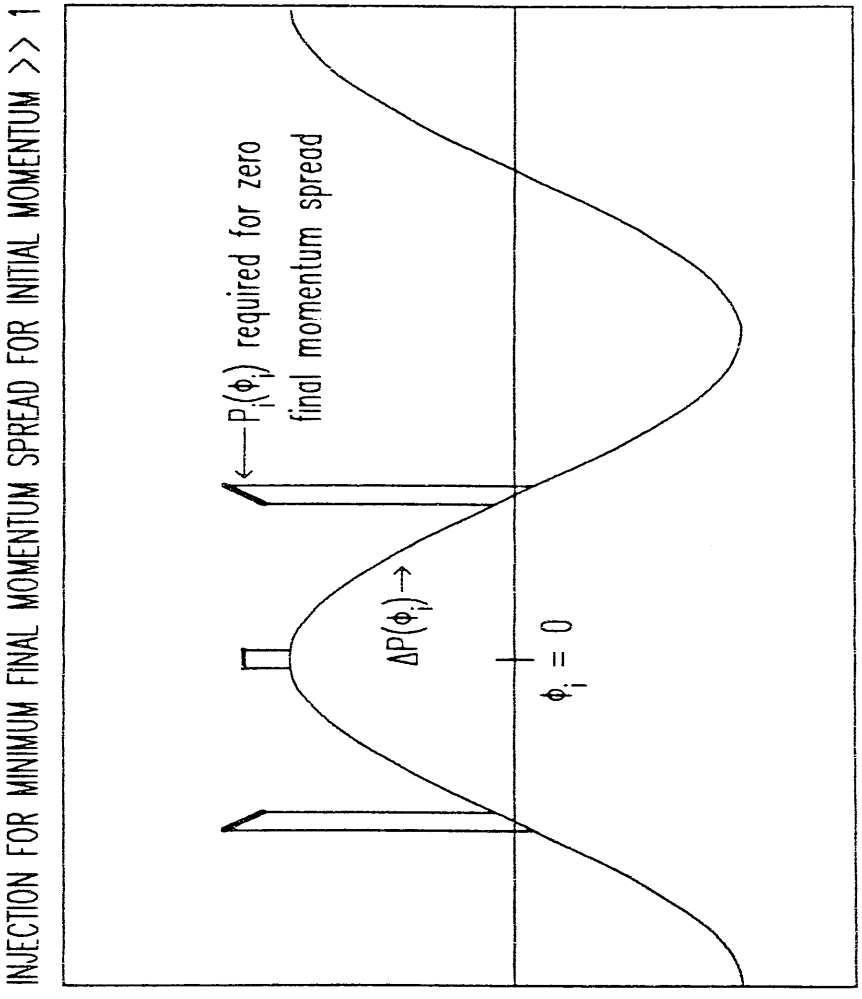

Figure 4.3: Explanation of Slopes of Constant Final-Momentum Contours

region decreases at the initial momentum increases, since particles of higher initial momentum loose less velocity when decelerated by the same fields, and eventually become captured at an accelerating phase.

Figure 4.2 gives additional insight into the longitudinal dynamics. One sees that the contour $\phi_f = 0$ lies in the region $\phi_i \geq 0$, approaching $\phi_i = 0$ as p_i increases. This is because in order for a slow-moving particle to end up at $\phi_f = 0$, it must be injected ahead of the crest so that the velocity-of-light RF wave catches up to it as it becomes relativistic. This effect is less important when the particle is initially highly-relativistic, which is why the contour approaches $\phi_i = 0$ as p_i increases.

This Figure shows that, by and large, in order to obtain a short bunch, one must first *have* a relatively short bunch. The slight slope to the constant ϕ_f contours around $\phi_i = 0$ indicates that it is best to inject the lower momentum particles ahead of the higher-momentum particles in OP bunch, so that the former will fall back to the same phase as the later before the entire bunch becomes relativistic.

In order to get both a short bunch and a small final momentum spread, it is necessary that one inject the bunch along a constant ϕ_f contour in a region where the constant p_f contours are widely spaced. Ideally, one would find two contours, one for constant ϕ_f and one for constant p_f, that coincide over the required interval of p_i, and inject one's bunch with the required phase-space distribution, ϕ_i vs p_i.

Typical operating conditions for the RF gun produce a peak momentum of $p = 5$, with momenta down to $p = 4$ accepted (giving approximately $\pm 10\%$ momentum spread about $p = 4.5$ for the "particles of interest"). As was demonstrated in Chapter 2, the higher momentum particles exit the gun first, with the particles of interest occupying roughly 25 ps, or roughly 25° of S-Band phase. From the above discussion, it is clear that this longitudinal phase-space distribution must be altered so that the higher-momentum particles enter the linac after the lower-momentum particles. (The region $\phi_i > 90$ is ruled out because the particles are decelerated before being accelerated, which is undesirable as it would lead to increased space-charge effects.) This can be accomplished by means of magnetic compression, as will be shown in the next section of this chapter. For present purposes, I shall assume that the magnetic compression system can supply the desired momentum-time correlation, and attempt to locate the optimum phase for injection in order to get the smallest final momentum

spread and bunch length.

I used `linac_cg` to compute the constant final-momentum and final-phase contours for the region $-50° \leq \phi_i \leq 90°$ and $3 \leq p_i \leq 6$, with $\Delta\phi_i = 1°$ and $\Delta p_i = 0.1$. For reasons discussed below, the region $-10 \leq \phi_i \leq 20$ is of particular interest. This region is shown in Figures 4.4 and 4.5.

From the previous set of Figures, one sees that for $\phi_i \lesssim -20°$, the contours of constant p_f and those of constant ϕ_f are most nearly parallel. This indicates that if the initial phase-space distribution could be shaped to match the contours in this region, then this might be the best place to inject. The problem with injection in this region is that since the contours of constant p_f are equispaced in Δp_f, the fractional momentum spread between the contours in this region is larger than for those just ahead of the crest. In addition, these contours are much more closely-spaced in ϕ_i than those nearer the crest. Hence, injection in this region is unlikely to yield good results in practice, since it is unlikely that the initial bunch phase-space could be tailored to the contours sufficiently well to obtain low final momentum spread.

From the Figures, I conclude that injecting closer to the crest, but still behind it by a few degrees looks promising, as does injection ahead of the crest by perhaps 15-20 degrees. The latter region suffers more from crossing of the contours of constant ϕ_f and constant p_f. Clearly, some compromise will have to be made between minimum final phase-spread and minimum final momentum spread. How one makes this compromise depends on one's application. For example, if additional accelerator sections follow, then it is probably best to inject into the first section so as to minimize the phase-spread at exit, so that all particles have, as much as possible, the same phase in all subsequent accelerator sections. This will ensure that the absolute momentum spread does not grow, in additional to giving the shortest bunch. As the bunch goes through subsequent sections, the fractional momentum spread will be further decreased.

The SSRL preinjector has a total of three accelerator sections. Hence, I will attempt to optimize injection into the first section primarily in order to obtain a short bunch. Having narrowed down the range of initial phase to be considered, it is next necessary to include details of the initial bunch phase-space distribution. This requires discussion of magnetic compression, which I go into immediately in the next section, returning to the combined problem in the section after next.

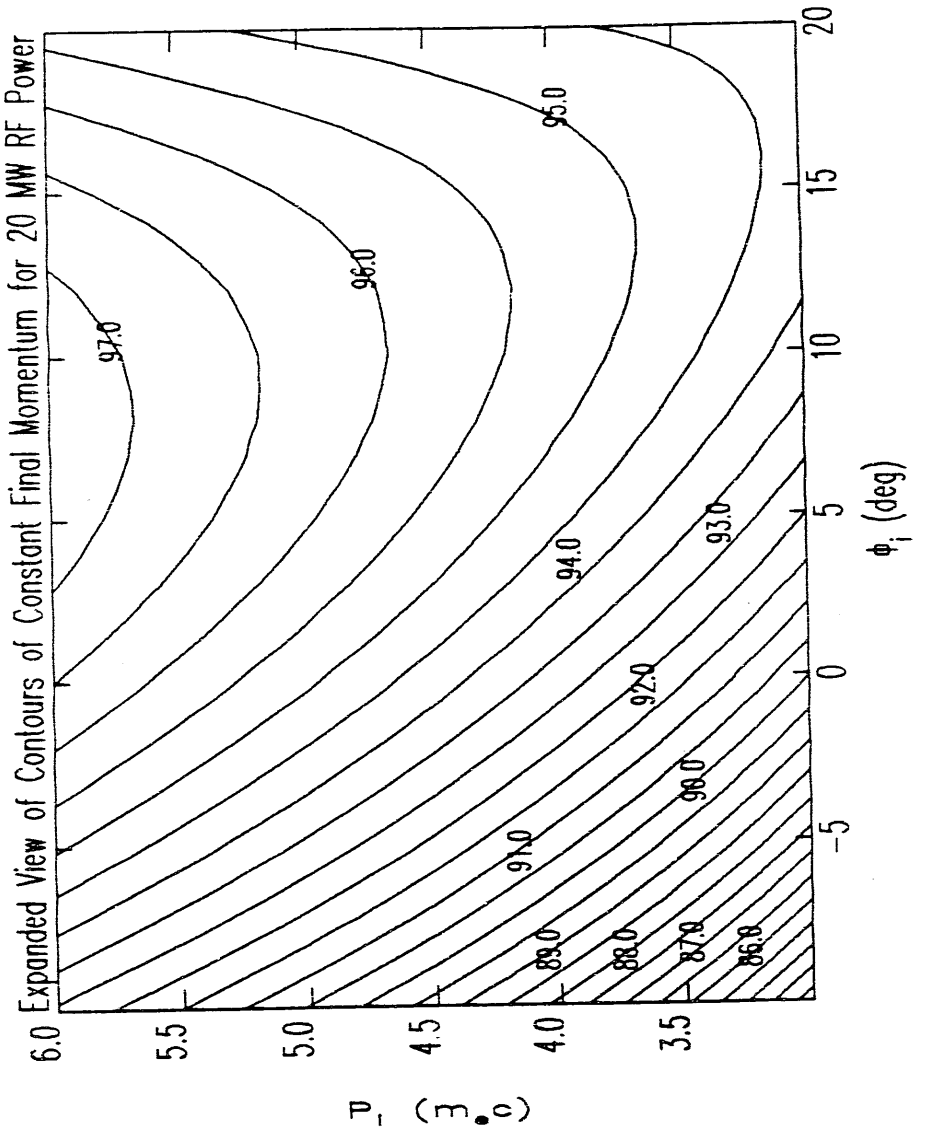

Figure 4.4: Expanded View of Constant Final Momentum Contours

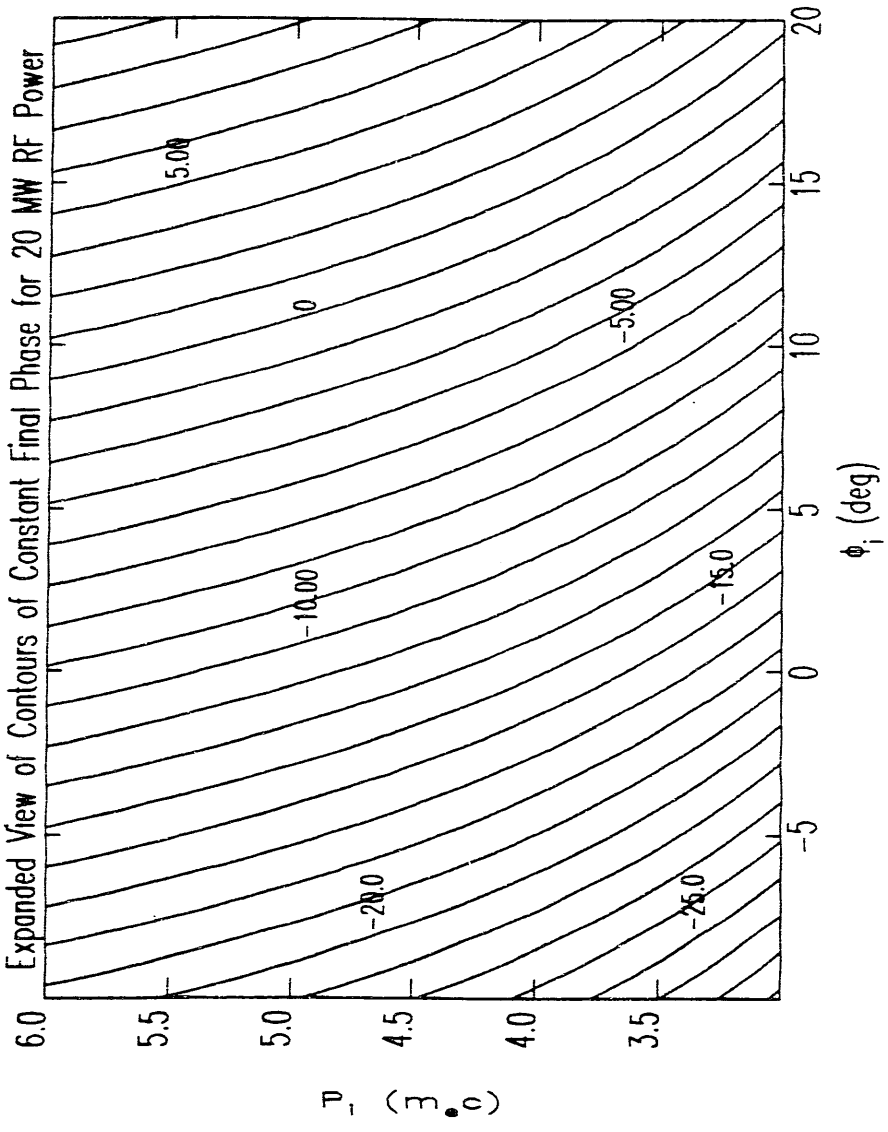

Figure 4.5: Expanded View of Constant Final Phase Contours

4.2 Magnetic Bunch Compression

In the previous section I have shown that, roughly speaking, what one wants is to deliver to the linac the shortest possible bunch. More precisely, one wants a bunch with a longitudinal phase-space distribution at the beginning of the linac that will be compressed further during initial acceleration. This depends upon being able to reverse the time-order of electrons within each RF gun bunch, so that the lowest momentum particles enter the linac first and subsequently fall back to the same phase as the highest momentum particles. Hence, a magnetic bunching system will have to be able to accomplish some degree of time-order reversal of particles in the bunch, over a sufficiently large momentum interval. However, this is typically only slightly more difficult that producing a very short bunch at the entrance of the linac, and so it is convenient to think in terms of how to produce a very short bunch. This allows a separation of the problem of magnetic bunching from the details of longitudinal dynamics in the linac, and hence prevents the issues from being obscured by too much detail at the outset. Once the mechanism of magnetic bunching has become clear, it is then possible to go back and consider the effects of longitudinal dynamics in the first part of the linac.

4.2.1 First Order Solution for Bunch Compression

Consider, then, that it is desired to produce a very short bunch at the entrance of the linac. It is known that the bunches from the gun have a particular momentum vs. exit-time characteristic, namely, that higher-momentum electrons exit the gun ahead of lower-momentum electrons. It is convenient to use the momentum deviation, $\delta \equiv (p - p_o)/p_o$, in terms of which, for a sufficiently small momentum interval about the central momentum $p_o \equiv (\beta\gamma)_o$,

$$t_{exit}(\delta) \approx t_o + \left(\frac{dt_{exit}}{d\delta}\right)_o \delta, \tag{4.17}$$

where t_{exit} refers to the time of a particle's exit from the gun and

$$\left(\frac{dt_{exit}}{d\delta}\right)_o < 0. \tag{4.18}$$

(Examination of the longitudinal phase-space distributions shown in previous chapters shows that, for the $\pm 10\%$ momentum spread that will be used, this linear approximation is not exact. However, it is again not my purpose now to deal with this detail, but rather to explain the principle.) What is desired in order to have a very short bunch at the end of the bunching system is that

$$t_{arrival}(\delta) = t_{exit}(\delta) + \Delta t_{flight}(\delta) = \text{constant}, \tag{4.19}$$

where the subscript "arrival" refers to arrival at the end of the bunching system (i.e., the entrance to the linac) and where I use Δt_{flight} to indicate that the time-of-flight is an interval rather than the time of some event. Combining these and using the linear approximation of equation (4.17), I obtain

$$t_{arrival}(\delta) = t_o + \left(\frac{dt_{exit}}{d\delta}\right)_o \delta + \Delta t_{flight}(\delta) = \text{constant}, \tag{4.20}$$

and hence

$$\left(\frac{dt_{arrival}}{d\delta}\right)_o = \left(\frac{dt_{exit}}{d\delta}\right)_o + \left(\frac{d\Delta t_{flight}}{d\delta}\right)_o = 0. \tag{4.21}$$

From this last equation and equation (4.18), one can see that

$$\left(\frac{d\Delta t_{flight}}{d\delta}\right)_o = -\left(\frac{dt_{exit}}{d\delta}\right)_o > 0 \tag{4.22}$$

must be obtained at the end of the bunching system. In words, since higher momentum particles come out ahead of lower momentum particles, they must be put through a system in which the time of flight is longer for high-momentum particles in order for all particles to arrive at the end at the same time.

For particles that are not fully relativistic, time-of-flight depends upon both velocity and the length of the path taken

$$\Delta t_{flight} = \frac{s(\delta)}{\beta c} \tag{4.23}$$

where $s(\delta)$ represents the length of the path taken by a particle with momentum deviation δ. Since $\beta = p/\sqrt{(p^2+1)}$, it follows, to first order in δ, that

$$\frac{1}{\beta} = \frac{1}{\beta_o} - \left(\beta_o - \frac{1}{\beta_o}\right)\delta. \tag{4.24}$$

Expanding $s(\delta)$ to first order as well, one obtains

$$\left(\frac{d\Delta t_{flight}}{d\delta}\right)_o = \frac{s_o}{c}\left(\beta_o - \frac{1}{\beta_o}\right) + \frac{1}{\beta_o c}\left(\frac{ds}{d\delta}\right)_o. \tag{4.25}$$

The first term of this expression shows that for highly relativistic particles, for which $\beta_o \to 1$, the effect of velocity variation on time-of-flight disappears, as would be expected. For non-relativistic particles, this term is negative, indicating that velocity effects will fight the bunching process. This is again expected, since the higher velocity of higher-momentum particles will help them to "pull ahead" even further. Clearly, if *bunching* is to occur, it will come from the variation of path-length with momentum.

4.2.2 Achieving Momentum-Dependent Path Length

Until now, I have said nothing about how one achieves a momentum-dependent path-length, although the name of this section is an indication. If particles of different momenta are to have different path-lengths in going through a transport line, they must of course first be made to take different paths through that system. In addition, these different paths must have different lengths. For example, merely sending particles of different momentum through a drift space at different transverse positions will not produce the desired effect. To see what is needed, consider first the expression for the path length in a transport line without bending magnets, where the central particle travels a straight-line path:

$$s(\delta) = \int_0^{s_o} \sqrt{1 - x'^2(s_o, \delta)}\, ds_o, \tag{4.26}$$

where the integration is with respect to the path length for the central particle. To first order there is no variation of path-length with momentum in such a beamline.

Now allow the central particle to traverse a section of a wedge bending magnet that bends it through an angle $\Delta\theta_o$, as illustrated in Figure 4.6. The path length for the central particle is given by $\Delta s_o = \rho_o \Delta\theta_o$, where ρ_o is the bending radius for the central particle. For an arbitrary particle, the path length is given by $\Delta s = \rho\Delta\theta$, where

$$\rho = \rho_o(1+\delta), \tag{4.27}$$

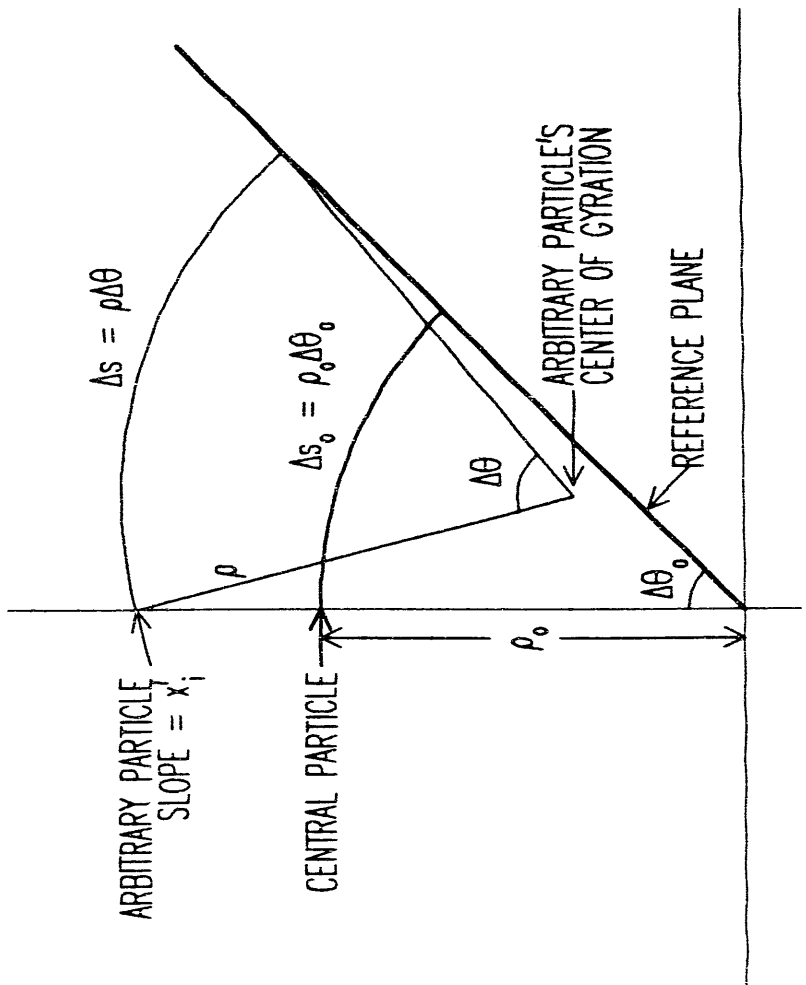

Figure 4.6: Particle Motion in a Wedge Bending Magnet

and where θ is the angle the particle is bent through in reaching the reference plane, as illustrated in Figure 4.6.

Particles other than the central particle will in general enter the bending magnet with different momenta, positions, and slopes relative to the central particle. Let x_i and x_i' be the initial position and slope of a particular particle, respectively. Some trigonometry reveals that the angle through which an arbitrary particle is bent is

$$\Delta\theta = \Delta\theta_o + \mathrm{atan}(x_i') - \mathrm{asin}\left\{\frac{\rho_o + x_i}{\rho}\sin(\Delta\theta_o) - \sin\left[\Delta\theta_o + \mathrm{atan}(x_i')\right]\right\}. \quad (4.28)$$

So far, no approximations have been made beyond assuming an ideal, hard-edge magnetic field. In order to get a first-order expression for the differential path-length in an infinitesimal section of a bending magnet, I expand to first-order in $\Delta\theta_o$, x_i, and x_i', obtaining

$$\Delta s = \rho\Delta\theta = \Delta\theta_o(\rho_o + x_i). \quad (4.29)$$

The initial coordinate x_i is at this point arbitrary. I am interested, however, only in momentum-dependent effects, and hence I will assume that the position of any particle at the entrance is a function only of its momentum, through the dispersion function D, defined by

$$x_i(\delta) = D\delta + \mathcal{O}(\delta^2). \quad (4.30)$$

Hence, the differential path-length is

$$\Delta s = \Delta s_o(1 + \frac{D}{\rho_o}\delta), \quad (4.31)$$

from which I conclude that

$$\frac{ds}{d\delta} = \int_0^{s_o} \frac{D(s_o)}{\rho_o} ds_o, \quad (4.32)$$

(I use the total derivative because x_i and x_i' are assumed to depend on δ, i.e., this quantity is not necessarily the matrix element r_{56} (it is equal to r_{56} only when the integration starts from a point where $D=0$).)

Referring back to equations (4.22) and (4.25), one sees that positive $\frac{ds}{d\delta}$ is required for bunch compression. This is obtained when D and ρ have the same sign, which is always the case for dispersion generated within the magnet itself (otherwise $\frac{ds}{d\delta}$ could be negative for a lone bending magnet, which is absurd). I define the sign of D

with respect to a right handed coordinate system (x, y, z), with z along the direction of motion and y along the upward vertical, so that for positive dispersion a larger momentum deviation implies a larger x coordinate. Hence, positive dispersion is generated by a bend to the right. The sign of ρ, as well as the sign of θ, for a bending magnet is then required to be the same as the sign of the dispersion it generates. This ensures that Δs_α is positive and that a lone bending magnet produces positive $\frac{ds}{d\delta}$. This is consistent with the conventions used by the beam line program MAD[71] and the tracking program elegant[49].

4.2.3 Options for Implementing Magnetic Compression

From this discussion it is clear that a single bending magnet could be used to provide bunch compression. However, there is inevitably dispersion at the end of a system with a single bending magnet, which is undesirable as it increases beam-size, effectively increasing the beam emittance. The next obvious step is to use two bending magnets of the same sign, with a focussing quadrupole between them to match the dispersion to zero at the end of the second bend (see Steffen[67] for examples of such systems). Such a system has a number of advantages, a principle one being that chromatic aberrations can be corrected through the addition of sextupoles between the bending magnets. However, there is the disadvantage that, since the bending angles of the magnets are fixed by the requirement of steering the central momentum down the center of the beamline, the bending radius ρ is fixed for each magnet, and hence D and $\frac{ds}{d\delta}$ are also fixed. Such a system is thus unsuitable for situations requiring variable compression, such as is needed for the RF gun, where $\frac{dt_{exit}}{d\delta}$ varies with p_o (i.e., as a function of the RF field level in the gun). Since the RF gun was still under development when the bunch compression system was being designed, it was not known before-hand what the operating momentum would be, and hence a system with variable compression was desirable.

For this reason it was decided to use a different type of magnetic-bunch compression scheme, namely one employing an alpha-magnet. The properties of this magnet are covered in detail in a Chapter 3. For present purposes, I will simply state that it

is first-order achromatic but has momentum-dependent path-length described by

$$s_\alpha(\delta) = K_\alpha \sqrt{\frac{p_o(1+\delta)}{g}}, \tag{4.33}$$

where s is in meters, g is the gradient in G/cm, $K_\alpha = 1.91655$, and, as above, $p_o = 1.2\,\beta_o\gamma_o$. From this, it is seen that for the alpha-magnet

$$\left(\frac{ds_\alpha}{d\delta}\right)_o = \frac{1}{2} K_\alpha \sqrt{\frac{p_o}{g}}. \tag{4.34}$$

Like bending-magnet-based schemes, an alpha-magnet provides momentum-dependent path-length because of bending and the resultant dispersion. However since the alpha-magnet is a gradient magnet, the bending radius varies with position along the central trajectory. The alpha-magnet has the advantage that the gradient, and hence $\frac{ds_\alpha}{d\delta}$, can be varied without changing the central trajectory outside of the alpha-magnet. While there are other systems with this property[72], the alpha-magnet is probably the simplest. It also has the advantage of relatively small aberrations, but has the disadvantage that there is no simple way to incorporate sextupoles for correction of chromatic aberrations in external quadrupoles that might be required as part of the beamline.

Evaluating (4.33) for $\delta = 0$, and inserting the result along with equation (4.34) into (4.25), and thence into (4.22), I obtain the requirement for bunching

$$\frac{1}{c}\left(K_\alpha\sqrt{\frac{p_o}{g}} - L_{drift}\right)\left(\beta_o - \frac{1}{\beta_o}\right) + \frac{K_\alpha}{2\beta_o c}\sqrt{\frac{p_o}{g}} = -\frac{dt_{exit}}{d\delta}, \tag{4.35}$$

where I have used

$$s(\delta) = L_{drift} + s_\alpha(\delta) \tag{4.36}$$

to incorporate the effects of any drift spaces between the gun and alpha-magnet and between the alpha-magnet and linac. It will prove useful to group the alpha-magnet terms together, as in

$$\frac{L_{drift}}{c}\left(\beta_o - \frac{1}{\beta_o}\right) + \frac{K_\alpha}{c}\sqrt{\frac{p_o}{g}}\left(\beta_o - \frac{1}{2\beta_o}\right) + \frac{dt_{exit}}{d\delta} = 0. \tag{4.37}$$

Solving for the gradient, I obtain

$$g = \frac{p_o K_\alpha^2\left(\beta_o - \frac{1}{2\beta_o}\right)^2}{\left[c\frac{dt_{exit}}{d\delta} + L_{drift}\left(\beta_o - \frac{1}{\beta_o}\right)\right]^2}. \tag{4.38}$$

This equation reveals a number of aspects of bunch-compression with an alpha-magnet. Since $\frac{dt_{exit}}{d\delta} < 0$ and $\beta_o - \frac{1}{\beta_o} < 0$, the denominator will be zero only if $L_{drift} = 0$ and the initial bunch has zero length, i.e., only if bunching is not needed. One also sees that the longer the drift spaces, the lower the alpha-magnet gradient must be, in order to compensate for the debunching.

The term $(\beta_o - \frac{1}{\beta_o})^2$ in the numerator combines the effects of debunching in the alpha-magnet due to differential velocity with that of bunching in the alpha-magnet due to differential path-length. Note that the solution (4.38) is not valid for $\beta_o < 1/\sqrt{2}$, since then all terms on the left-hand side of (4.37) are negative. Hence, for $\beta_o < 1/\sqrt{2}$, the alpha magnet *cannot* bunch, as the effects of velocity variation will always overcome the effects of path-length variation. (This is false only if $\frac{dt_{exit}}{d\delta} > 0$, a situation that does not apply for the RF gun.)

Taking the limit of equation (4.38) as $\beta_o \to 1$, one obtains

$$\lim_{\beta_o \to 1} g = \frac{p_o K_o^2}{\left(2c\frac{dt_{exit}}{d\delta}\right)^2}, \tag{4.39}$$

which indicates that for constant bunch length and constant fractional momentum spread (implying constant $\frac{dt_{exit}}{d\delta}$), the gradient must increase with increasing central momentum. If, however, the *absolute* momentum spread is kept constant (as happens with acceleration of relativistic particles near the crest of the RF field), then $\frac{dt_{exit}}{d\delta}$ scales as p_o, which indicates that the gradient must scale inversely with momentum. A smaller gradient implies a larger alpha magnet, since the size of the central trajectory scales as $1/\sqrt{g}$ (see Chapter 3). Hence, bunching before acceleration is advantageous in that it decreases the size of the alpha-magnet, at the cost of requiring a higher gradient. Similarly, as $\beta_o \to 1/\sqrt{2}$ from $\beta_o > 1/\sqrt{2}$, the gradient must become vanishingly small, implying a increasingly large alpha-magnet.

4.3 Optimized Bunch Compression for the RF Gun

In the previous two sections, I discussed the principles that must be employed in choosing the injection phase and the alpha-magnet gradient for production of short bunches with low final momentum spread. I attempted to separate the two aspects, for simplicity in the discussion. In this section, I demonstrate how to obtain optimum performance with simultaneous consideration of acceleration and magnetic bunching, along with inclusion of the detailed initial phase-space. Not surprisingly, this optimization is best done numerically.

I have written a program, `alpha_opt`, that accepts initial longitudinal phase-space information in terms of (ϕ_i, p_i) pairs for macro-particles (e.g., from MASK or `rfgun`) and attempts to find the optimum alpha-magnet gradient for a specified accelerator phase and energy gain. It optimizes for either the minimum mean absolute phase deviation or the minimum total phase-length of the final bunch, though I have used the latter criterion exclusively in this work. Equation (4.33) is used without approximation in (4.36) to give the momentum-dependent path-length. Equation (4.1) is used for the traveling wave field. To simulate particle motion in the accelerator, I employ equations (4.2) and (4.3) (scaled for more efficient computation), which I integrate using the so-called "leap-frog" method[61], which is second-order accurate in the time-step. Typically, I find that taking time-steps smaller than 30 ps makes no change in the results (i.e., no change of more than ± 0.001ps in the bunch length).

4.3.1 Use of alpha_opt to Optimize Bunch Compression

The combined distance, L_{drift}, from the gun to the alpha-magnet and from the alpha-magnet to the center of the first linac cell (where the traveling wave begins) was chosen based on simulations of the gun longitudinal phase-space, the anticipated strength and good-field-region of the alpha-magnet, and the need for a sufficiently long drift-space to accommodate the quadrupoles and chopper. Because the gun was still under development at the time the alpha-magnet and GTL were being specified, I chose L_{drift} so that compression would be feasible over a wide range of gun operating

momenta, rather than finding an optimum for any particular beam momentum. If L_{drift} were chosen to be too short, then an excessively strong alpha-magnet would be needed in order to reduce the compression, while if L_{drift} were chosen to be too long, an unreasonably large good-field region would be needed to provide more compression. L_{drift} = 1.5m was initially chosen based on preliminary simulations with `rfgun` and `alpha_opt`, along with knowledge of the (then preliminary) magnetic design of the alpha-magnet. Later, L_{drift} was increased to 1.7m in order to provide more space for other GTL components.

More specifically, there is a 0.6m drift space from the gun to the alpha-magnet crossing point, and a 1.1m drift from the alpha-magnet to the linac. See Chapter 5 for more discussion of the layout of the GTL.

I performed a series of `alpha_opt` runs starting with the MASK-generated longitudinal phase-space distribution for the RF gun operated at E_{p2} = 75MV/m and J = 10A/cm^2. The linac simulation parameters were such that an initially relativistic particle injected at the crest would gain 45 MeV (which corresponds to 20.7 MW RF power). I took the highest-momentum particle as the fiducial particle, and chose to attempt to compress the beam for a variety of momentum spreads, namely $\pm$10%, $\pm$5%, and $\pm$2.5%. That is, I applied a perfect momentum-filter to the MASK-generated beam, accepting only particles such that $p_c(1-f) \le p \le p_c(1+f)$, with $p_c \equiv p_{max}/(1+f)$, where $\pm f$ is the fractional momentum spread accepted. In this way, the selected momentum range always contains the highest momentum particles. (This same capability exists on the actual beamline, where a scraper inside the alpha-magnet can be moved into the beam from the low-momentum side.)

For each value of f, I first found the alpha-magnet gradient which produced the shortest bunch at the entrance to the linac. I then used this gradient and sent the bunch down the linac with the highest-momentum particle injected at the crest, fully expecting that the result would be a less than optimally compressed bunch. The simulations confirmed this expectation, as the data listed in Table 4.1 shows. (In this and all subsequent Tables and Figures, Δt and ΔP are the full spread of the values, e.g., $\Delta t = t_{max} - t_{min}$.) In addition, one sees that the absolute moment spread has increased. The phase-space distributions at the linac entrance and exit are represented graphically in Figure 4.7. Note that these graphs are of *time* and momentum, rather

than phase and momentum, and that particles to the left are ahead of particles to the right. As expected, the initially higher-momentum particles pull ahead of the initially lower-momentum particles, resulting in a longer final bunch.

Table 4.1: **Optimization for a Short Bunch at the Linac Entrance**

$\Delta p/p$ (%)	Q (pC)	g_o (G/cm)	Δt_i (ps)	Δp_i (m_ec)	Δt_f (ps)	Δp_f (m_ec)
$\pm$10	110.9	335.40	1.050	0.978	3.030	1.808
$\pm$5	80.3	338.32	0.724	0.513	1.285	0.818
$\pm$2.5	50.4	321.22	0.493	0.263	0.750	0.436

The conclusion to be gained from this result is that it is not sufficient to design a compression system that will generate a short bunch at the entrance to the linac. It is necessary to take into account the longitudinal dynamics in the linac in order to ascertain whether one can indeed produce a very short bunch at the end of the linac. In the present case, one expects that what is needed is to increase the compression (by using smaller gradients in the alpha-magnet) so that the lower-momentum particles enter the accelerator ahead of the higher-momentum particles. This expectation is confirmed by `alpha_opt`.

I directed `alpha_opt` to optimize the alpha-magnet for the shortest bunch at the end of the linac. The same linac parameters were used as before. The optimum alpha-magnet gradients are smaller than previously found. Table 4.2 lists the results for this optimization. The phase-space distributions at the linac entrance and exit are represented graphically in Figure 4.8. One sees that for this optimization the increase in the absolute momentum spread is significantly smaller than for the previous optimization. The explanation is that the final bunch length is achieved a relatively short distance into the accelerator section (because the particles are already relativistic), and hence in the previous optimization the bunch had a large phase-spread during most of the acceleration, resulting in an increase in momentum spread.

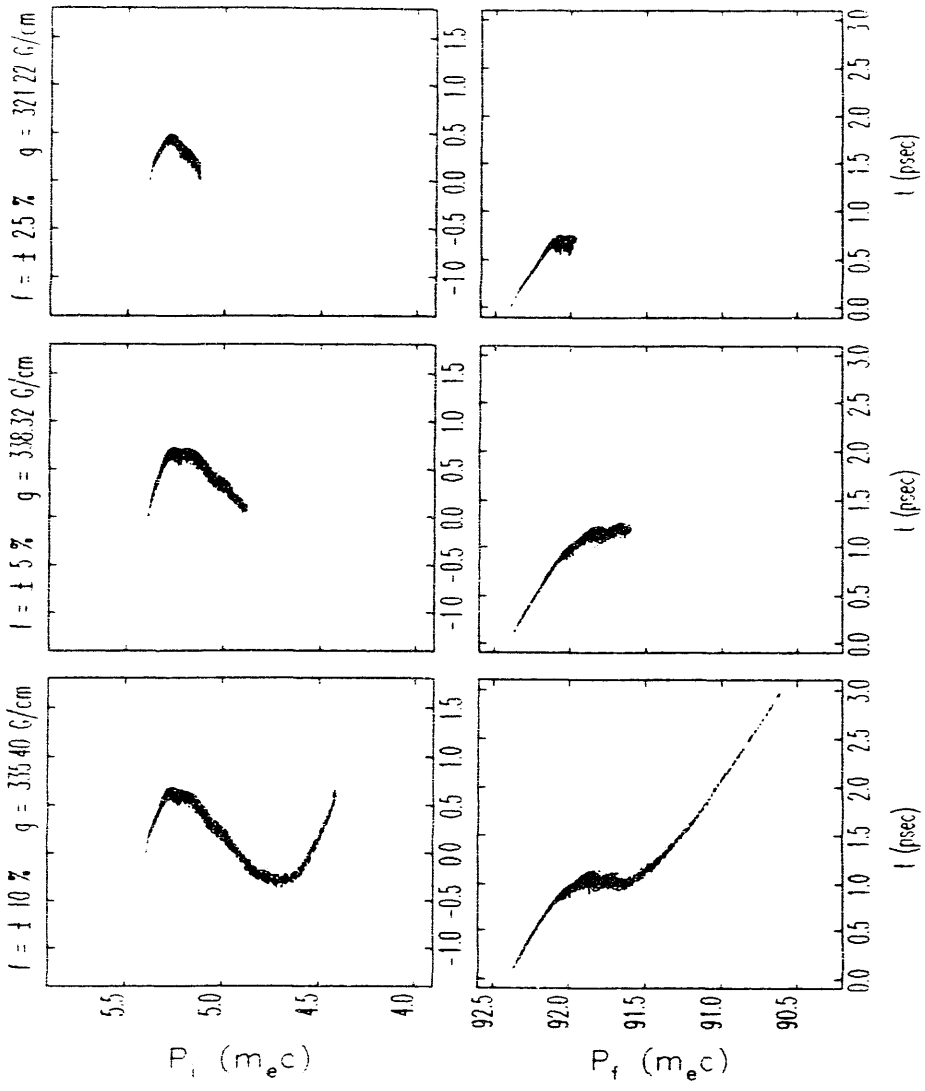

Figure 4.7: Result for Compression Optimized for a Short Bunch at Linac Entrance

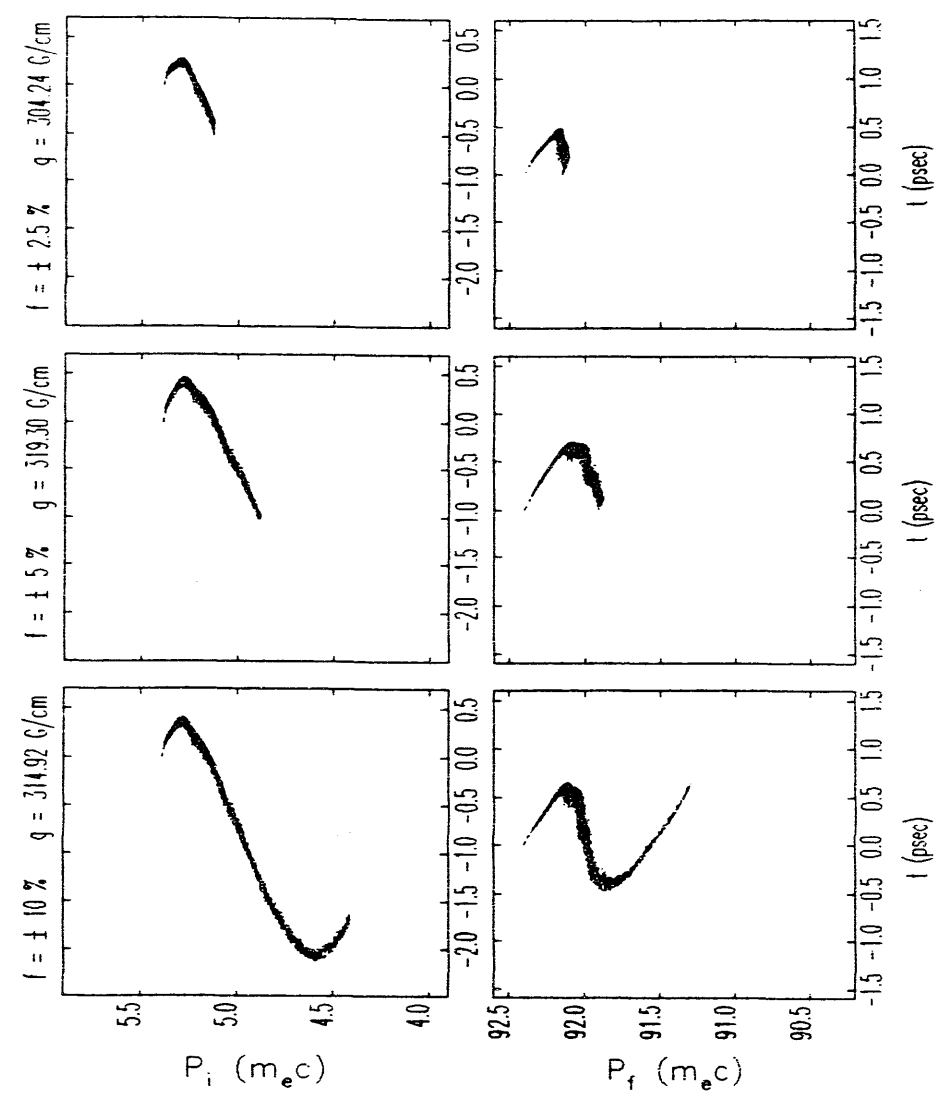

Figure 4.8: Results for Compression Optimized for a Short Bunch at Linac Exit

Table 4.2: **Optimizations for a Short Bunch at the Linac Exit**

$\Delta p\ p$ (%)	g_o (G/cm)	Δt_i (ps)	Δp_i ($m_e c$)	Δt_f (ps)	Δp_f ($m_e c$)
± 10	314.92	2.548	0.978	1.139	1.112
± 5	319.30	1.556	0.513	0.719	0.527
± 2.5	304.24	0.810	0.263	0.496	0.297

4.3.2 Optimization of the Injection Phase

For $\Delta p\ p = \pm 10\%$, I have done an additional series of simulations, designed to investigate the effect of the initial phase of the fiducial particle. In particular, I repeated the optimization for a series of values of the initial phase of fiducial particle. The results are listed in Table 4.3 and displayed graphically in Figure 4.9. Notice that the smallest final momentum spread and the highest average final momentum are achieved by injecting the bunch $10 - 15°$ ahead of the crest, so that it falls back to the crest before becoming fully relativistic. The smallest final bunch length is achieved for $\phi_o = 20°$. As might have been expected from the contour method of the previous section, the optimizations for highest total momentum gain, smallest final momentum spread, and smallest final bunch length are to some extent incompatible, though not grossly so. While some advantage in terms of final bunch length is obtained by accelerating well off the crest, the advantage is small and is obtained at the expense of considerably higher final momentum spread. That this should be so is confirmed by the contour-plots of the first section, where one sees that the contours of constant final phase become more widely spaced as ϕ_i increases from zero up to around 90°. The explanation is, perhaps, that injecting further from the crest allows a longer time for the particles to bunch before they are fully relativistic. Presumably, if this explanation is correct, one would find the optimum injection phase for the shortest bunch becoming smaller as one decreased the rate of acceleration.

The reader may notice that the numbers for $\phi_o = 0$ in Table 4.3 are different from those in Table 4.2. The reason for this is that for the optimizations presented in Table 4.3, I used a sample of the MASK-generated longitudinal phase-space distribution containing only 20% of the macro-particles in order to economize computer resources,

whereas in the previous two Tables I used all the macro-particles (3461 macro-particles for $\pm 10\%$ initial momentum spread). Each set of data is self-consistent in the size of the sample used.

Table 4.3: **Optimizations for $\Delta p/p = \pm 10\%$ for Various Injection Phases**

ϕ_o (degrees)	g_o (G/cm)	Δt_i (ps)	Δt_f (ps)	$\langle p_f \rangle$ ($m_e c$)	Δp_f ($m_e c$)
-25	313.36	2.641	1.230	76.858	1.489
-20	314.08	2.580	1.193	81.339	1.416
-15	314.60	2.536	1.162	85.082	1.340
-10	314.90	2.511	1.137	88.104	1.257
-5	315.11	2.494	1.116	90.403	1.182
0	315.22	2.485	1.096	91.994	1.110
5	315.07	2.497	1.079	92.899	1.034
10	314.80	2.520	1.063	93.127	0.964
15	314.40	2.553	1.046	92.701	0.918
20	313.81	2.603	1.029	91.649	0.973
25	313.01	2.670	1.011	90.000	1.087
30	312.00	2.757	0.991	87.800	1.201
35	310.68	2.871	0.971	85.085	1.320
40	308.92	3.026	0.945	81.909	1.431
45	306.67	3.326	0.916	78.333	1.537

4.3.3 Optimizations for Various Current Densities

To obtain predictions of the maximum peak currents that might be obtained with the SSRL system, I have done a series of optimizations for $(\Delta p/p)_i = \pm 10\%, \pm 5\%$ and $\pm 2.5\%$ using MASK-generated initial longitudinal distributions for $E_{p2} = 75 MV/m$ and $0 \leq J \leq 80 A/cm^2$. Since the initial longitudinal distribution is affected by space-charge in the gun, it is necessary to do the optimization for each current level. I chose $\phi_o = 15°$ as a compromise between minimum bunch length, maximum momentum gain, and minimum momentum spread. As before, I assumed 45 MeV as the linac energy gain. The results are summarized in Table 4.4 and in Figures 4.10 through 4.12. (Note that the data points in the figures are connected as an aid to the eye.

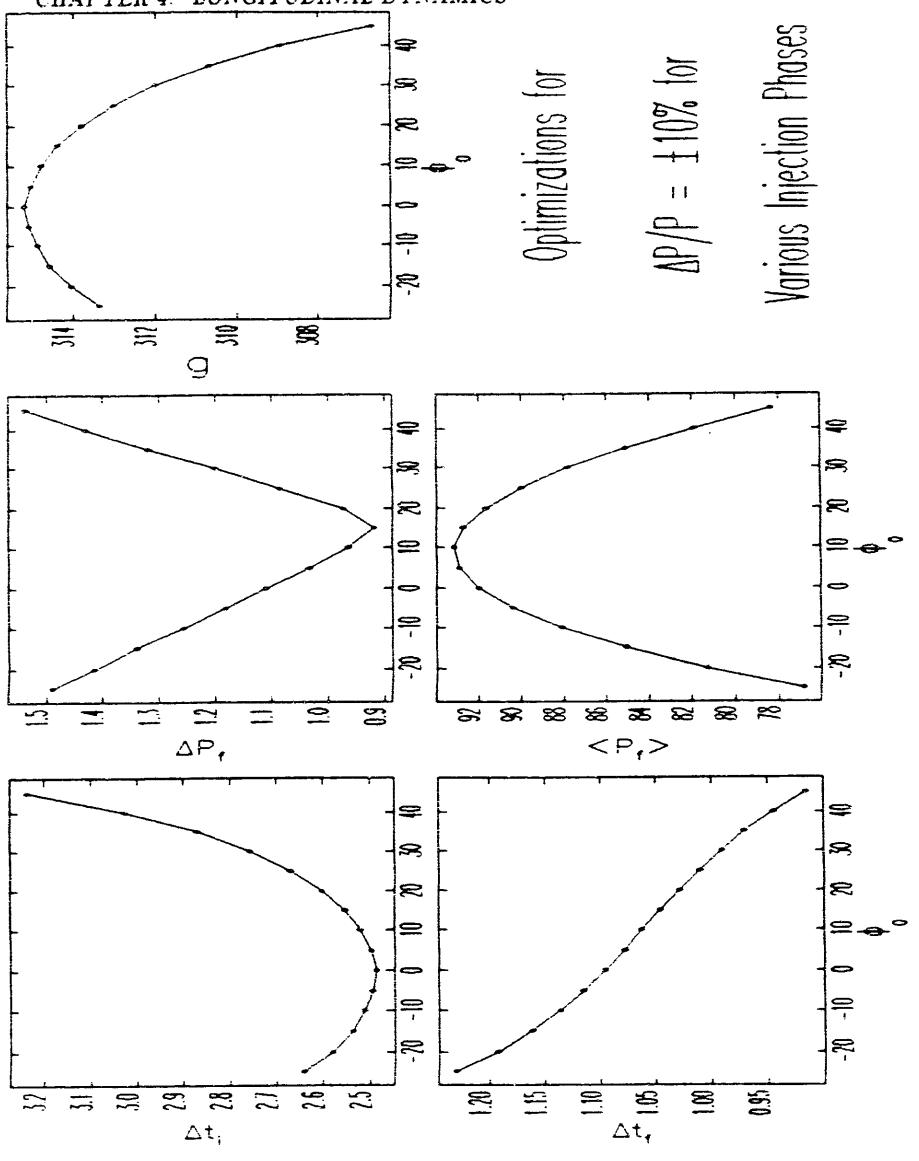

Figure 4.9: Optimization for a Short Bunch at the Linac Exit for Various ϕ_o

and not to indicate any assumed variation in the quantities between data points.)

Table 4.4: **Optimizations for Various Cathode Current Densities**

J (A/cm^2)	g_α (G/cm)	Δt_f (ps)	$\Delta t_{f,mad}$ (ps)	$\langle p_f \rangle$ (m_ec)	Δp_f (m_ec)	$\Delta p_{f,mad}$ (m_ec)	Q (pC)	$\langle I \rangle$ A
$(\Delta p/p)_i = \pm 10\%$								
0.2	313.20	1.070	0.292	92.737	0.906	0.263	2.3	2.1
10	314.36	1.097	0.278	92.700	0.922	0.262	110.9	101.1
20	315.20	1.121	0.287	92.661	0.951	0.267	217.8	194.3
40	318.61	1.193	0.277	92.630	0.979	0.265	416.6	349.2
80	321.28	1.272	0.315	92.573	1.027	0.275	762.8	600.7
$(\Delta p/p)_i = \pm 5\%$								
0.2	313.40	0.879	0.183	92.896	0.556	0.130	1.7	1.9
10	318.51	0.685	0.142	92.879	0.537	0.128	80.3	117.2
20	318.30	0.802	0.157	92.842	0.577	0.138	155.6	194.0
40	329.60	0.567	0.100	92.850	0.540	0.128	289.3	510.2
80	339.68	0.494	0.090	92.850	0.531	0.126	501.6	1015.4
$(\Delta p/p)_i = \pm 2.5\%$								
0.2	293.80	0.531	0.100	92.949	0.280	0.066	1.1	2.1
10	303.48	0.475	0.084	92.943	0.276	0.066	50.4	106.1
20	309.94	0.432	0.086	92.934	0.278	0.069	94.6	219.0
40	325.18	0.457	0.086	92.953	0.262	0.065	162.8	356.2
80	335.18	0.409	0.087	92.960	0.281	0.062	260.7	637.4

4.3.4 Effects of Transport Aberrations

These predictions of high peak currents neglect space-charge forces in the gun-to-linac transport line and in the linac itself. They also neglect the effects of non-chromatic t_{5jk} and u_{5jkl} terms ("aberrations") in the alpha-magnet, and of field errors in the alpha-magnet (see Chapter 3). Other effects that are not included in the analysis are wake-fields in the accelerator section. In Chapter 3, I discuss the effect of field errors, and show that the effect of field errors on r_{5j} matrix elements is small, from which I conclude that the ability of the alpha-magnet to compress the bunch is unaffected by field errors.

To evaluate the effects of space-charge, both in the GTL and in the accelerator

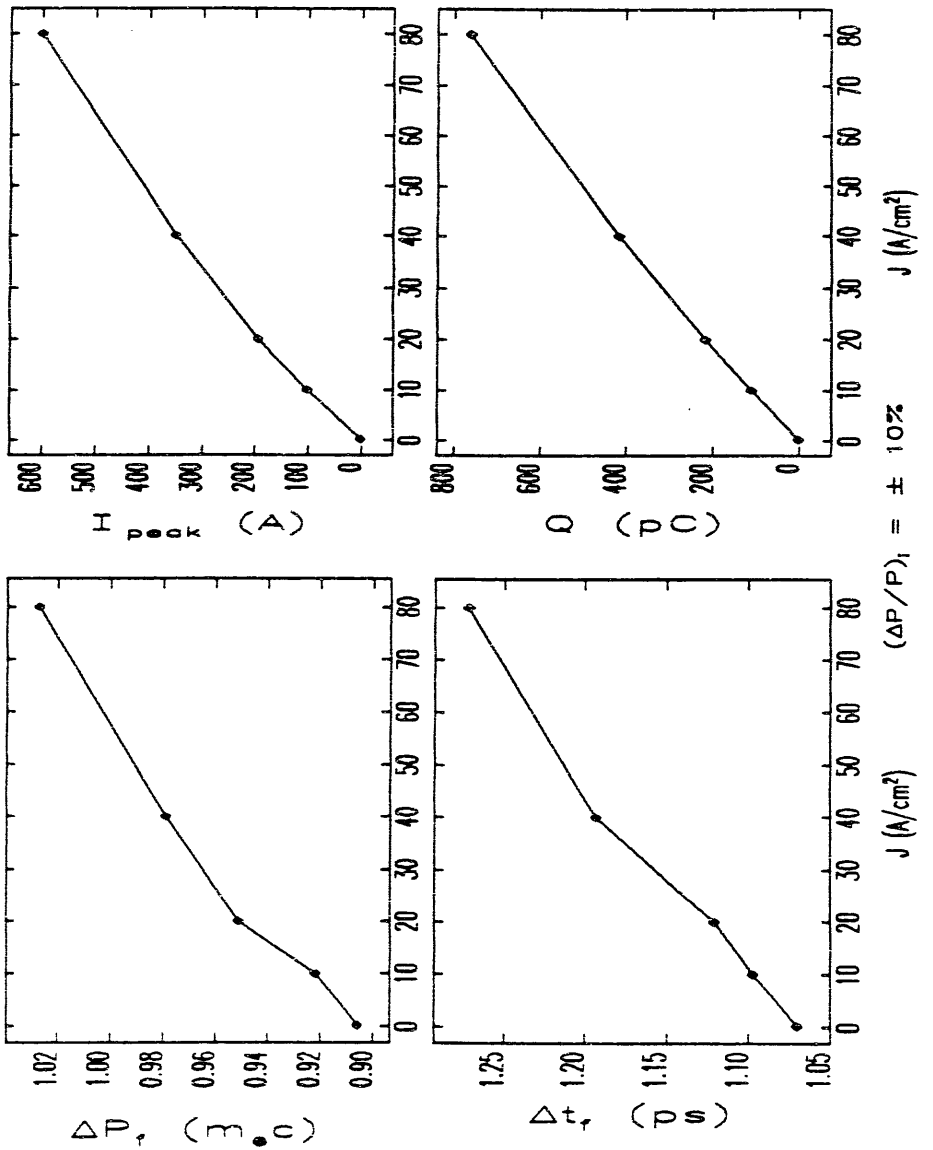

Figure 4.10: Results of Optimized Compression for Various Cathode Current Densities. for $(\Delta P/P)_i = \pm 10\ \%$

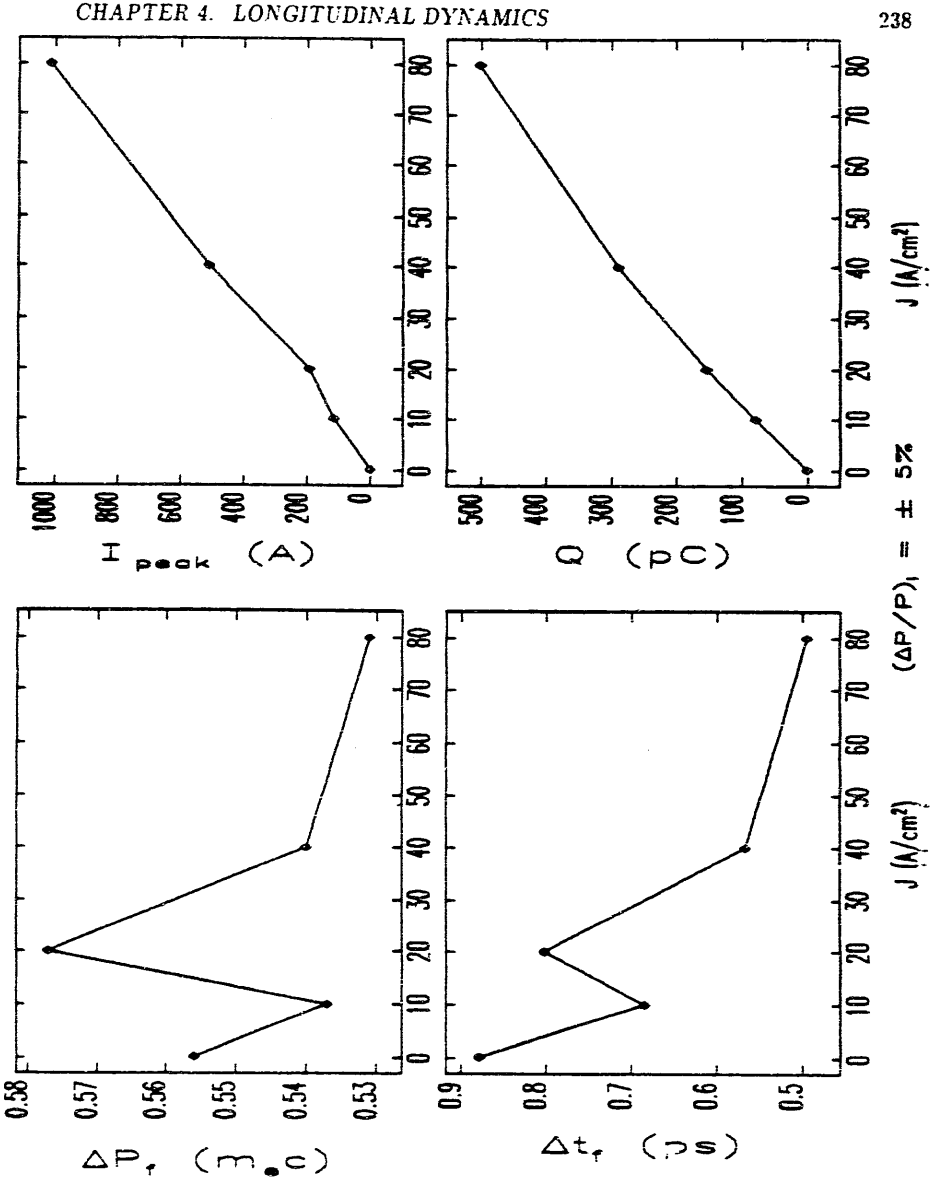

Figure 4.11: Results of Optimized Compression for Various Cathode Current Densities. for $(\Delta P/P)_i = \pm 5\ \%$

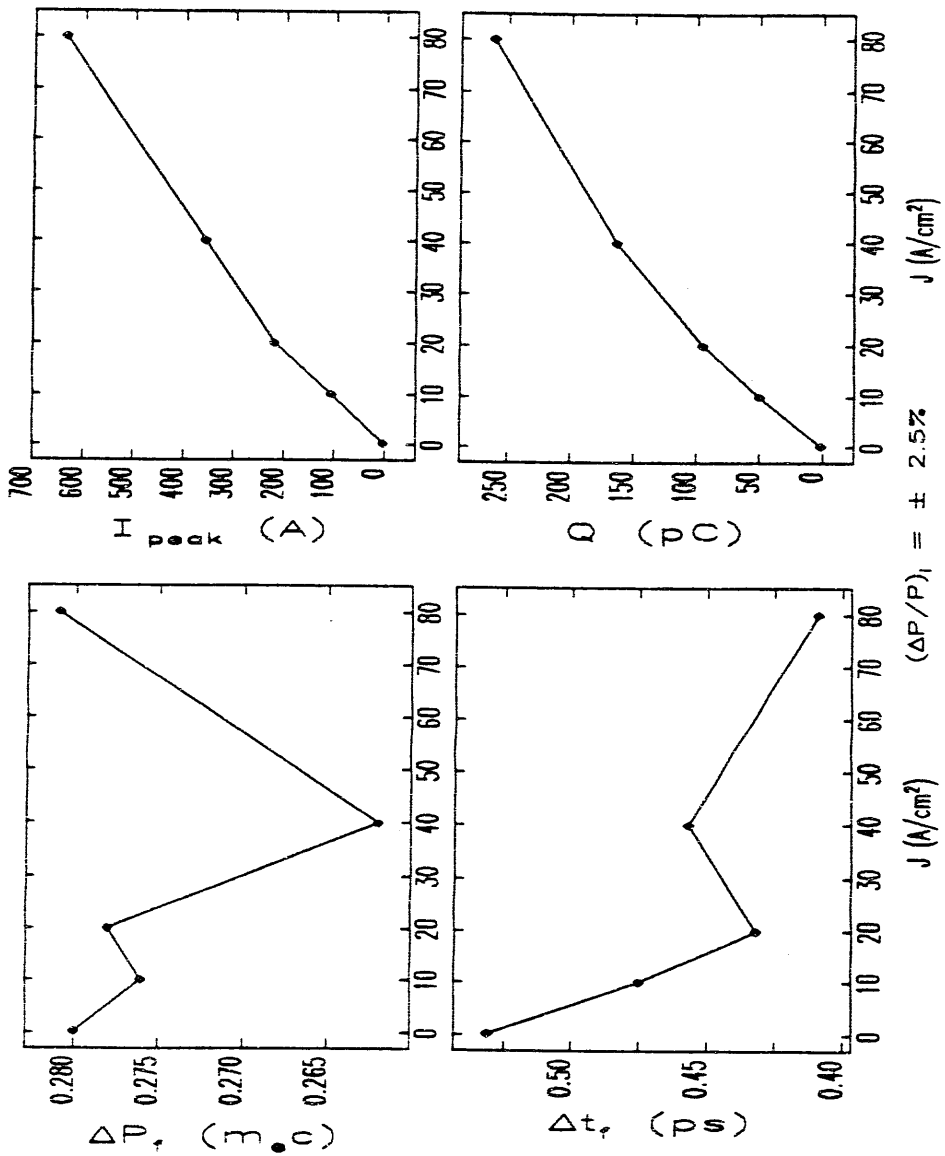

Figure 4.12: Results of Optimized Compression for Various Cathode Current Densities, for $(\Delta P/P)_i = \pm 2.5$ %

section, it would be necessary to employ a program such as PARMELA that is capable of simulating beam-transport with space-charge. Unfortunately, PARMELA does not include alpha-magnets, nor does time permit me to modify the program to remedy this deficiency (more would be required than simply inserting the transport matrix). In addition, the space-charge algorithm used by PARMELA is not well-suited to use for thermionic RF guns, where there is a large velocity spread in the beam. Finally, I have not found that PARMELA performs accurately in calculating simple test cases, such as the spread of a uniform cylindrical beam. Hence, evaluation of the effects of space-charge in the GTL must await the development of a suitable program, and will not be pursued here.

However, the program **elegant**[49] is capable of accurately simulating the GTL, ignoring space-charge. **elegant** includes chromatic aberrations (see Chapter 5) in the quadrupoles and alpha-magnet as well as other aberrations in the alpha-magnet (see Chapter 3). I will discuss the GTL optics and such issues as chromatic aberrations in Chapter 5. For the present, I simply present the results of **elegant** simulations of the GTL and the first linac section, which use the same initial phase-space data as was used in the previous calculations. That is, the **elegant** simulations took initial phase-space data generated by MASK for $E_{p2} = 75$MV/m, for a range of current densities, and for initial momentum spreads of $\pm 5\%$ and $\pm 10\%$. The results are shown in Figures 4.13 and 4.14, which are to be compared to Figures 4.10 and 4.11, respectively.

In addition to showing the peak current at the end of the linac section, I have shown the peak current at the gun exit, and the cathode current (i.e., $\pi R_c^2 J$), to illustrate the increase in peak current due to the bunching processes in the gun and GTL/linac. I have not shown the momentum spread, in order to use the space for other quantities, and because it is essentially the same as the previous results.

One sees that the peak currents predicted by **elegant** are considerably less than those obtained previously. The reason is that path-length aberrations in the GTL increase the broadness of the momentum versus time curves, making compression to very short bunches more difficult. In addition, the transmission through the first section is only 70% (particles are lost on the approximately 18 mm diameter apertures between linac cells), which reduces the amount of charge reaching the end of the

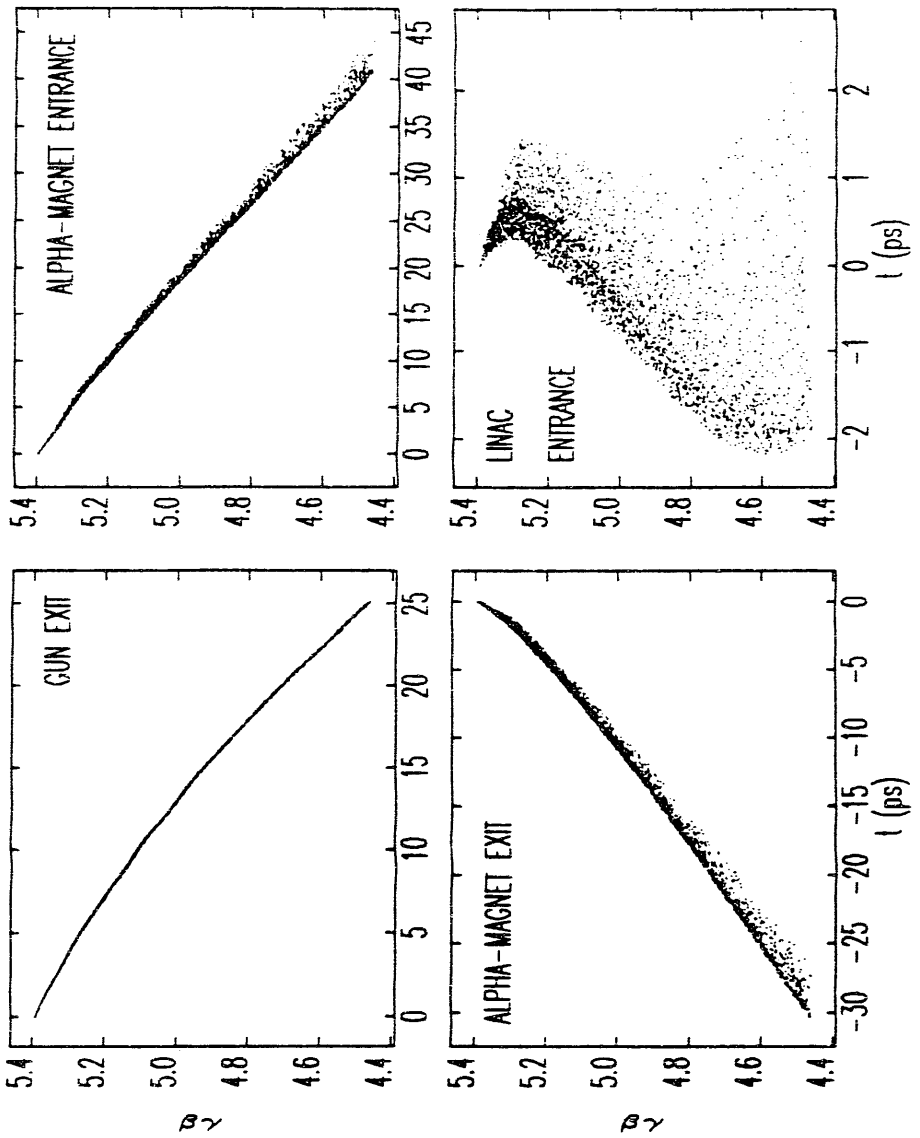

Figure 4.13: **elegant**/MASK Results after First Accelerator Section, for Various Cathode Current Densities and $E_{p2} = 75$MV/m, for $(\Delta P/P)_i = \pm 10$ %

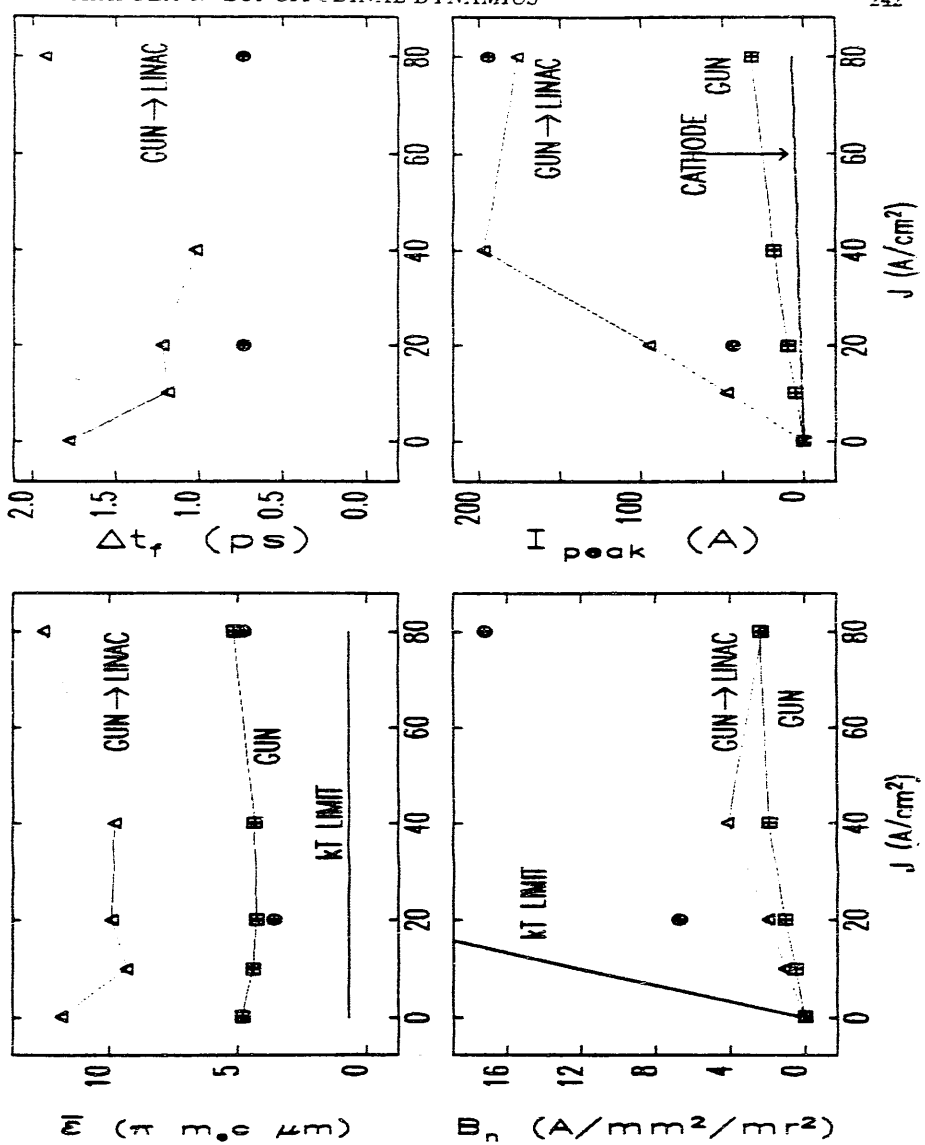

Figure 4.14: **elegant**/MASK Results after First Accelerator Section, for Various Cathode Current Densities and $E_{p2} = 75$MV/m, for $(\Delta P/P)_i = \pm 5$ %

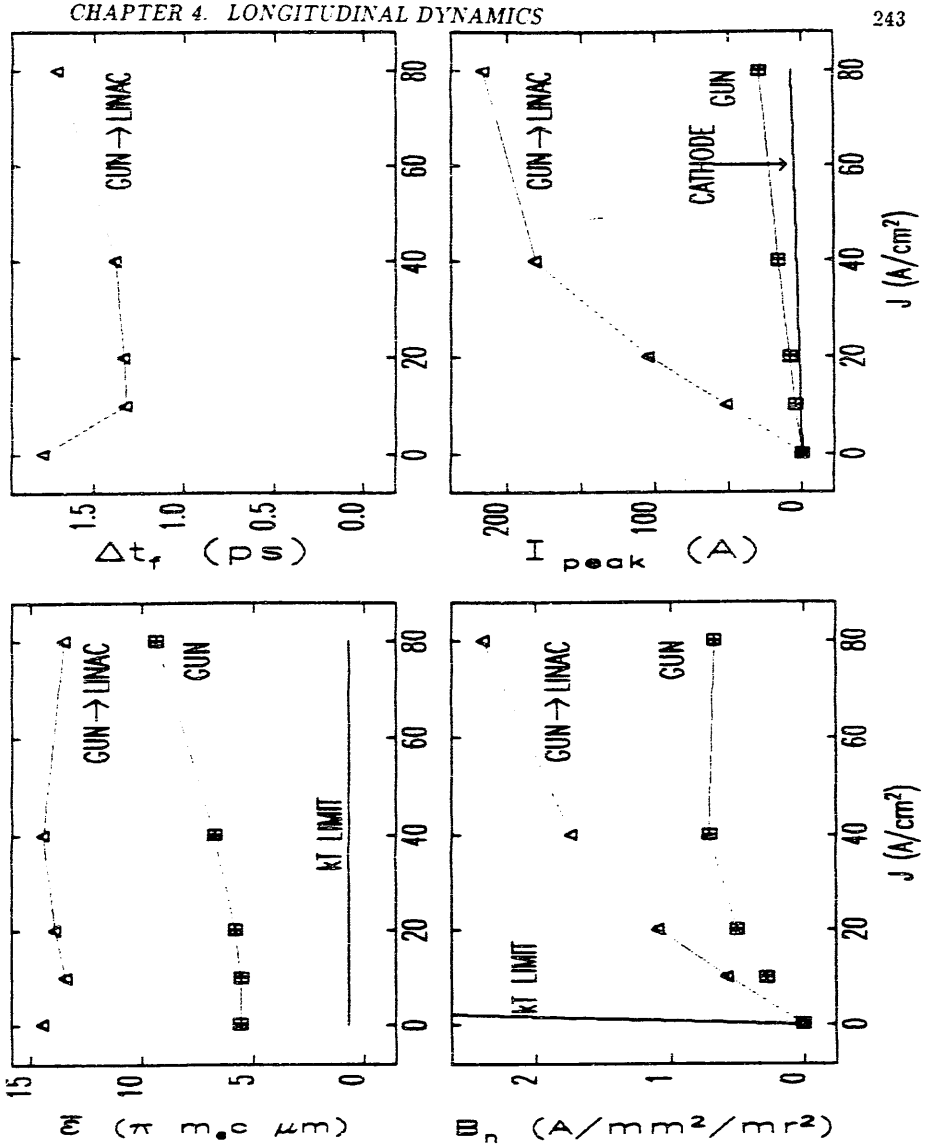

Figure 4.15: Longitudinal Phase-Space at Various Points in the GTL (elegant/MASK results for $E_{p2} = 75$MV/m, J = 10A/cm^2, and $(\Delta P/P)_i = \pm 10$ %).

linac. Figure 4.15 shows the evolution of the longitudinal phase-space in the GTL, for J = 10A/cm^2 and $\Delta P/P = \pm 10\%$. As one would expect, the bunch length at the alpha-magnet entrance is greater than that at the gun exit, due to the higher velocity of the lead particles in the bunch.

The broadening of the longitudinal phase space is due to path-length-affecting aberrations in the quadrupoles and drift spaces between the gun and alpha-magnet. To see that this is reasonable, note that for a drift space of length L_o, the path length traveled by a particle with non-zero slope is

$$L = L_o\sqrt{1 + x'^2 - y'^2} \approx L_o - \frac{L_o}{2}(x'^2 + y'^2). \tag{4.40}$$

At the gun exit, $x'_{rms} = y'_{rms} \approx 10$ mrad, and the straight-line distance from the gun exit to the alpha-magnet entrance is 60 cm. The path-length increase for $x' = x'_{rms}$ and $y' = y'_{rms}$ is 60μm, which corresponds to a time delay of 0.2 ps. Since there are particles in the beam with x' and y' the several times the RMS value, the broadening seen is larger than this estimate. As a result of such aberrations, the phase-space at the linac entrance differs considerably from the results show in Figure 4.8, because the latter results did not include any consideration of transverse motion.

Figures 4.13 and 4.13 also show the normalized RMS emittance and brightness at the end of the first linac section, as well as results at the gun exit, for comparison with those at the end of the linac. Recall that the emittance is defined as

$$\varepsilon_{n,x} = \pi m_e c\sqrt{\langle x^2\rangle\langle p_x^2\rangle - \langle p_x x\rangle^2} \tag{4.41}$$

and the brightness as

$$B_n = \frac{2I_{peak}}{\varepsilon_{n,x}\varepsilon_{n,y}}(\pi m_e c)^2. \tag{4.42}$$

The emittance shown in the Figures is the geometric mean of the emittances for the x and y planes, $\bar{\varepsilon}_n = \sqrt{\varepsilon_{n,x}\varepsilon_{n,y}}$.

The emittance at the end of the linac section is larger than that at the gun exit, but not as large as the emittance at the entrance to the linac. The emittance is "filtered" in the linac because particles with large transverse amplitudes are lost on the linac disc apertures. Put another way, the emittance numbers do not refer to the same particles, since 30% are lost. One sees that the emittance depends only

weakly on the current density; this is due to the overwhelming effect of chromatic aberrations. The emittance at the entrance to the linac is about twice that at the end of the section. These points are discussed further in Chapter 5.

In addition to showing the simulation results for the emittance, I have shown the thermal limit on the emittance for a cathode of R_c = 3mm, using[16]

$$\varepsilon_{n,x} = \varepsilon_{n,y} = \frac{1}{2} R_c \sqrt{\frac{kT}{m_e c^2}}, \tag{4.43}$$

where T is the cathode temperature, which is 1200°K for the SSRL gun. One sees quite clearly that the thermal limit is far from being approached: the emittance is dominated by RF focusing, non-linear fields in the gun, and chromatic aberrations in the GTL.

Figure 4.14 also shows two data points obtained by simulating the gun with a smaller emitting area on the cathode. (These appear as crossed circles in the graphs.) In particular, an emitter radius of 1.5 mm was used, with the physical cathode size kept at 3mm radius. In effect, the region from r = 1.5mm to r = 3mm was taken to be a "dead region" on the cathode. In this situation, particles are emitted much closer to the axis in the gun, so that non-linear fields in the gun have less of an effect, resulting in a smaller emittance. In addition, the smaller emittance leads to smaller effects from path-length-affecting aberrations in the GTL, so that shorter bunch-lengths are achieved. While the amount of charge drops due to the decrease in emitting area, this is balanced to some extent by the shorter bunch-length, so that the peak current at J = 80A/cm^2 is increased. Because of the strong effect on the emittance, the brightness is dramatically increased. These results make a strong case for operating the gun with such a cathode, especially since the cathode is currently operated well below its maximum current density, meaning that a reduced emitting area could be used with no loss of total charge.

4.3.5 Comparison with Other Injectors

The data of Figures 4.13 and 4.14 permit comparison of the predicted performance of the SSRL preinjector (i.e., the RF gun, GTL, and linac) with other RF-linac-based preinjectors. In order to do this, I have reviewed recent literature giving parameters

of existing and planned injectors. There is always the chance of confusion in any such compilation, especially since many authors do not state their definition of the emittance or bunch length. Where doubt exists, I have assumed that the quoted emittance is the 4-σ or "edge-emittance" and that the bunch length refers to 90-95% of the beam, since these appear to be the most commonly used definitions.

One extremely useful resource in this regard was C. Travier's review article on RF guns[14], which gives extensive performance data for RF guns and state-of-the-art DC gun systems (i.e., those with high-performance guns and multiple subharmonic bunchers). I have also taken data from T. I. Smith's review[46], which lists several systems planned for or already in use as FEL drivers; these are not necessarily state-of-the-art systems. (Where Travier and Smith differ on the same system, I have used Travier's data, which is more recent.) I also show data points for several other systems that are intended for FEL use[73, 74, 75] as well as SLAC's SLC[76] (including damping rings) and the original SLAC injector[48].

Note that I will compare *injectors*, rather than *guns*. From an applications-oriented viewpoint, this is the most appropriate comparison to make among systems using various types of guns, since it includes all of the effects that come into play when one actually makes use of the beam from a gun. It also avoids issues such as whether a multi-cell thermionic RF gun should be compared to a DC gun with prebunchers, given that the multi-cell RF gun is in some sense a combined gun and prebuncher.

The data for DC-gun-based and microtron-based systems are in Table 4.5, while those for RF guns are in Table 4.6. Two data points are listed for the SSRL system. Both are for E_{p2} = 75MV/m and f=0.05, but one assumes J = 40A/cm^2 with R_c = 3mm, while the other assumes J = 80A/cm^2 with R_c = 1.5mm. (These are both consistent with less than 4 MW incident RF power, which is the anticipated upper limit that will be supplied to the gun after some recent, but untested, hardware upgrades.) Figure 4.16 shows some of this data in graphical form, with addition points supplied for the SSRL system, as explained on the graph.

One sees that the SSRL system is predicted to perform quite well in terms of peak current and brightness, achieving levels comparable to those achieve by much more sophisticated and complicated systems. One also sees, however, that the high brightness and high peak current are achieved by generating very short pulses, which

are not appropriate to FEL work at wavelengths that are not very long compared to the electron bunch length (see Chapter 1 for a discussion).

Table 4.5: **Performance of DC-Gun and Microtron- Based Injectors**

Project	ε_n $\pi \cdot m_e c \cdot \mu m$	I_{peak} A	Q nC	δt ps	B_n A/mm^2/mr^2
SLC (1986) [76]	30	2400	8	3.3	5.3
SLC 2 [14]	75	580	10.4	18	0.21
SLC 1 [14]	43	430	7.7	19	0.47
Boeing [14]	13	350	4.9	14	4.2
LANL [14]	60	300	9.0	30	0.17
ALS [14]	40	200	4.0	20	0.25
CLIO [14]	7.5	100	1.5	15	3.46
Trieste FEL [74]	50	15	0.15	10	0.012
UK FEL [46]	13	10	-	-	0.12
Frascatti [75]	1.4	6	-	-	6.1
SCA/TRW [46]	1.3	4	-	-	4.73
Orig.SLAC [48]	5.7	0.3	-	-	0.018
NIST-NRL [73]	5	0.3	14	15	0.024

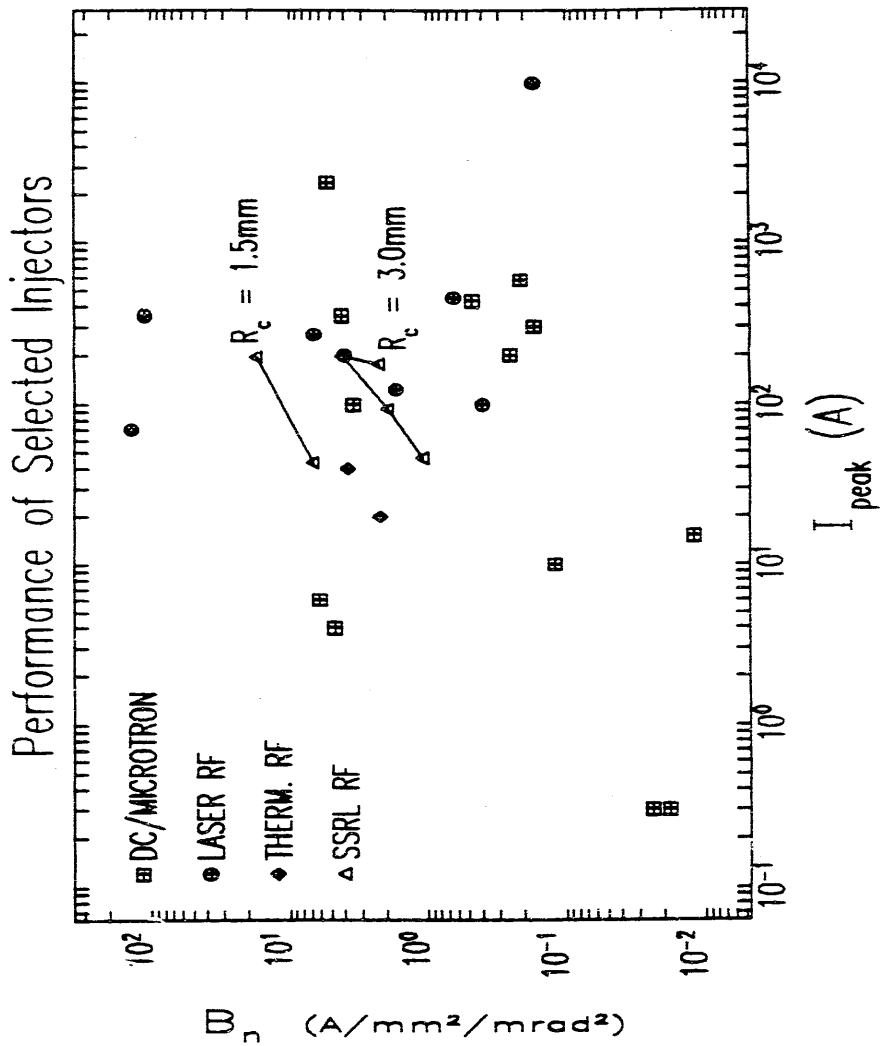

Figure 4.16: Brightness and Peak Current for Various Injectors

Table 4.6: **Performance of RF-Gun-Based Injectors**

Project	ε_n $\pi \cdot m_e c \cdot \mu m$	I_{peak} A	Q nC	δt ps	B_n $A/mm^2/mr^2$	type
ANL 14	340	$10 \cdot 10^3$	100	8	0.17	laser
CERN 14	37.5	450	9	30	0.64	laser
LANL:						
AFEL 14	2.5	350	5	16	112	laser
HIBAF 14	9.0	270	4	15	6.7	laser
PHASE I 14	10.0	200	11	70	4.0	laser
SSRL:						
3mm, 40A/cm^2	9.5	196	0.2	1.0	4.4	therm.
1.5mm, 80A/cm^2	4.8	144	0.2	0.7	17.2	therm.
BNL 14	12.0	125	1	8	1.7	laser
CEA 14	22.5	100	10	100	0.4	laser
DFELL 14	1.0	70	0.17	2.5	140	laser
DFELL 14	4.6	20-40	.05-.1	2-3	1.9-3.8	therm.
IHEP 14	4.3	10-20	.080-.1	4-5	1.1-2.2	therm.

Chapter 5

Gun Experimental Characterization

In the previous chapters, I have made detailed predictions of how the gun is expected to perform. In this chapter, I report on the experimental tests I have made of those predictions. Characterization of the gun can be broken down into several categories. There are steady-state properties and there is the evolution of the system into the steady-state.

System evolution deals with the response of the beam–cavity system to the RF pulse, and includes such topics as the evolution of beam current, reflected power, and beam momentum. This topic is beyond the scope of the present work, and will be addressed in future publications. Theoretical and numerical treatment of steady-state properties appeared in Chapter 2.

For the steady-state, there is a natural subdivision into transverse and longitudinal beam properties. Of course, by transverse beam properties I mean primarily the emittance, but also the transverse phase-space ***distribution***, of which the emittance is only one parameter. The emittance measurements I have done relied on the variation of a quadrupole upstream of a phosphorescent screen. From the variation of beam-size with quadrupole strength one can, as I will show, deduce the emittance (under certain assumptions). While this cannot, even ideally, provide complete knowlege of an arbitrary transverse phase-space distribution, it does provide all of the second-order moments. Also, if the measured beam size as a function of quadrupole strength

is compared to a simulation of the same experiment, using an initial beam distribution calculated by MASK, and if the two are found to be similar, then one may reasonably conclude that not only are the measured and calculated emittances in agreement, but also that the measured and calculated phase-space distributions may well be in agreement.

Longitudinal beam properties include the momentum spread and the dependence of momentum on cavity field levels, in addition to the bunch length. Momentum spectra proved one of the best diagnostics available on the SSRL system, with measurements being possible over the full range of gun operating conditions. In spite of this, there is ambiguity in the measurements in that direct knowledge of the cavity fields and cathode current density was not available.

Bunch length measurements were performed using the third preinjector linac section, phased at the null so as to impart a time-dependent momentum spread to the beam. By analysing the momentum spread with a bending magnet as a function of the fields in the linac section, the bunch length can be determined.

The term "steady-state" refers to a hypothetical condition reached by the gun after a sufficient time has elapsed since the beginning of the driving RF pulse. As I discussed in chapter 3, this condition is expected to be reached only during the later part of the (rather short) 2μs RF pulse delivered to the gun at SSRL, particularly when beam-loading is small (as happens for low-current running). This injects some ambiguity into any experiment that purports to be a measurement of a steady-state property. Additional complications arise from the fact that the RF pulse delivered to the gun had a "flat-top" with an upward slope amounting to (typically) 10% of the average power level of the flat-top.

Being a combination of a research project and a construction project serving a "higher goal", the RF gun project did not include the most sophisticated diagnostics possible. Certainly, in order to characterize the gun sufficiently to verify many of the detailed predictions I have made in previous chapters, more complete experiments are needed than I have been able to do. However, the experiments that have been done demonstrate that the gun performs largely as expected. As my discussion of the experiments proceeds, I will indicate what the short-comings of the measurements are thought to be, and how these might be overcome in future work.

While the experiments I will report on represent my own efforts, they could never have been done without the efforts and support of many people, whom I am pleased to recognize in the acknowledgements.

5.1 Experimental Configuration

In this section, I will give an overview of the primary experimental configuration that I used at SSRL[54]. While some experiments were done at Varian or with different configurations at SSRL, the data I shall cover were taken with the configuration I shall describe here.

5.1.1 Gun-to-Linac Components

Figure 5.1 shows a schematic layout of the "Gun-to-Linac" region (or "GTL") of the SSRL Preinjector. This system is similar to that used for the HEPL RF gun[32], in that both systems employ an RF gun with an alpha-magnet for bunch compression (see Chapters 3 and 4), along with quadrupoles to control the transverse beam size.

In addition to providing bunch compression, necessary in order to obtain low momentum spread at the end of the linac, the alpha-magnet also allows momentum filtration. As discussed in Chapter 3, the beam inside the alpha-magnet is dispersed according to momentum. A scraper (referred to as "the scraper" or "the alpha-scraper") inside the alpha-magnet can be moved into the beam from the low-momentum side, thus allowing the operator to let through only those particles with momenta greater than a certain value (this is discussed fully in a following section). I will discuss how the alpha-scraper is used for measurement of the momentum spectrum in Section 5.3.

The scraper is only one of the variables under the control of the operator. The alpha-magnet gradient and the gradients in the quadrupoles are all independently controllable, allowing great flexibility in the optics. The RF power delivered to the gun, the phase and frequency of that power, and the cathode filament power are the gun-specific "knobs" at the operator's disposal.

Steering magnets are also included in the transport line. These are useful for magnetic optics tests and to compensate for alignment errors, magnetic field errors, and stray magnetic fields. Each steering magnet provides both horizontal and vertical steering, using a picture-frame design with four coils. Steering magnets are placed just after the gun, before and after the alpha-magnet, and between the last quadrupole and the chopper.

Figure 5.1 also shows the traveling-wave beam-chopper[77], which is used to select

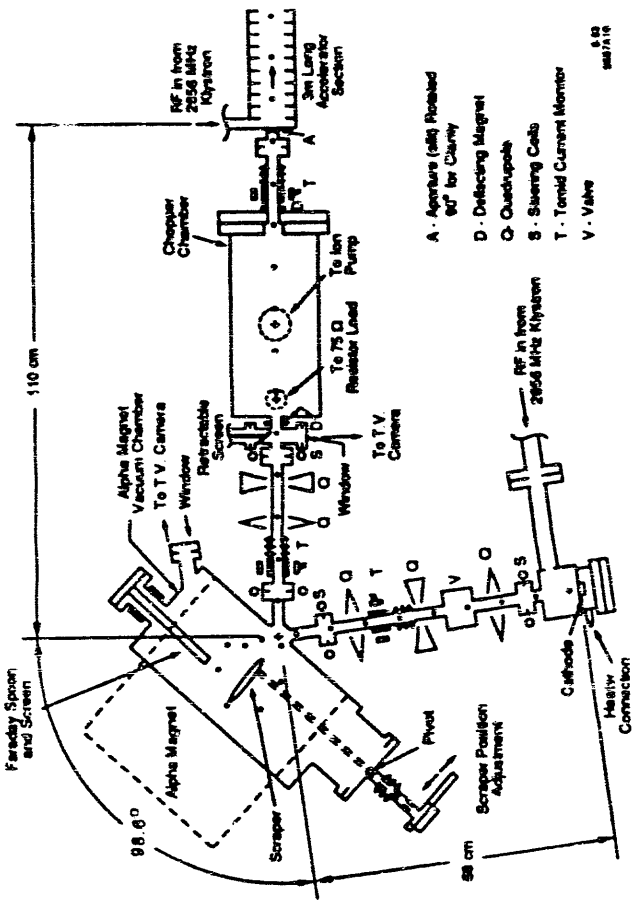

Figure 5.1: Gun-to-Linac Region of the SSRL Preinjector

three to five bunches for injection into the linac out of the train of several thousand bunches that is emitted by the gun during each RF pulse. I will not discuss the chopper further here.

5.1.2 Gun-to-Linac Instrumentation

Several types of diagnostic instrumentation are available in the GTL[78, 79]. Forward and reflected RF power signals are available from a directional coupler in the rectangular waveguide leading to the gun. The coupler is not as close to the gun as one might like: originally, the coupler was quite near the gun, but the ceramic window covering the coupling hole broke, forcing us to find a new position for the coupler. The coupler used for these experiments is about 1 m from the gun body, with a wave-guide "H-bend" and a high power RF window between the gun and the coupler. These signals are converted into voltages by calibrated diodes, so that one obtains a signal that is related to the power envelope of the RF signal in question. Because of the H-bend between the coupler and the waveguide, reflected power signals will include the effect of any mismatch between the H-bend and the rectangular waveguide leading up to it, instead of simply being a product of the match between the gun cavity and the waveguide. Cold-test measurements (performed by others) indicate, however, that the match of the H-bend to the waveguide is very good.

Beam-current measurements can be obtained from either of two toroids, one before the alpha magnet ("GT1") and one following it ("GT2"), as well as from a "Faraday-spoon" mounted inside the alpha magnet. In addition to being the current monitors that are used in day-to-day operation of the GTL, the toroids find application in the measurement of momentum spectra. As I will discuss in more detail below, by measuring beam transmission from GT1 to GT2 as a function of scraper position, one can obtain the momentum spectrum and the beam power.

The toroids are essentially transformers, with the beam current acting as one winding of the transformer. The name "toroid" derives from the use of a ferrite torus as the core of the transformer. The torus is mounted symmetrically around the path of the beam, so that the beam goes down the axis of the torus. At the position of the torus, there is a ceramic break installed, since otherwise the image currents induced

in the metallic vacuum pipe would cancel the beam current. The other winding of the transformer consists of wire coiled around the ferrite. A shield is also installed around the torus, providing a path for image currents that goes around the outside of the ferrite. Because this image current is equal in magnitude and opposite in sign to the beam current, the effect is equivalent to having the beam current make a loop around the ferrite. Hence, the "turns ratio" of the transformer is 1 :: n, where n is the number of turns of wire wrapped on the ferrite.

The Faraday-spoon is essentially a copper target that is held inside the alpha-magnet in such a position that is is struck by the beam whenever the alpha-magnet is turned off. The target is supported by copper tubes (which double as conduits for cooling water), which are themselves eventually supported by an electrically isolated vacuum feed-through. Hence, the electrons that are absorbed by the Faraday-spoon do not return immediately to ground, but can rather be made to return to ground through an external resistor. The voltage across this resistor is proportional to the current absorbed by the Faraday-spoon, from Ohm's law. The surface of the Faraday spoon where the electrons impact the copper was not shaped into a cup, since for electrons of 2—3 MeV, back-scattering should be negligible.

Mounted on the front surface of the Faraday-spoon is a phosphorescent screen, which emits visible light when struck by electrons. This screen, referred to as the "alpha-magnet screen", is viewed by closed-circuit TV through an optical window in the alpha-magnet vacuum chamber. This screen is heat-sunk to the Faraday-spoon, with the front surface of the Faraday-spoon being very slightly convex in order to ensure good thermal contact between the screen and the copper.

Another phosphor screen is installed down-stream of the alpha-magnet, in front of the chopper. This screen, referred to as the "chopper screen", is retractable and, unlike the alpha-magnet screen, is not water-cooled. Hence, while the alpha-magnet screen can take full beam current (more than 900 mA) without apparent harm, the chopper screen is easily damaged by too much beam current, and in addition will cause an unacceptable increase in GTL and gun vacuum pressures if inserted with more than about 50 mA current exiting the alpha-magnet. Because it follows the alpha-magnet, the chopper screen provides the only way to measure the emittance of the beam without the inaccuracy that would occur if the full momentum spread were

5.1 Experimental Configuration

In this section, I will give an overview of the primary experimental configuration that I used at SSRL[54]. While some experiments were done at Varian or with different configurations at SSRL, the data I shall cover were taken with the configuration I shall describe here.

5.1.1 Gun-to-Linac Components

Figure 5.1 shows a schematic layout of the "Gun-to-Linac" region (or "GTL") of the SSRL Preinjector. This system is similar to that used for the HEPL RF gun[32], in that both systems employ an RF gun with an alpha-magnet for bunch compression (see Chapters 3 and 4), along with quadrupoles to control the transverse beam size.

In addition to providing bunch compression, necessary in order to obtain low momentum spread at the end of the linac, the alpha-magnet also allows momentum filtration. As discussed in Chapter 3, the beam inside the alpha-magnet is dispersed according to momentum. A scraper (referred to as "the scraper" or "the alpha-scraper") inside the alpha-magnet can be moved into the beam from the low-momentum side, thus allowing the operator to let through only those particles with momenta greater than a certain value (this is discussed fully in a following section). I will discuss how the alpha-scraper is used for measurement of the momentum spectrum in Section 5.3.

The scraper is only one of the variables under the control of the operator. The alpha-magnet gradient and the gradients in the quadrupoles are all independently controllable, allowing great flexibility in the optics. The RF power delivered to the gun, the phase and frequency of that power, and the cathode filament power are the gun-specific "knobs" at the operator's disposal.

Steering magnets are also included in the transport line. These are useful for magnetic optics tests and to compensate for alignment errors, magnetic field errors, and stray magnetic fields. Each steering magnet provides both horizontal and vertical steering, using a picture-frame design with four coils. Steering magnets are placed just after the gun, before and after the alpha-magnet, and between the last quadrupole and the chopper.

Figure 5.1 also shows the traveling-wave beam-chopper[77], which is used to select

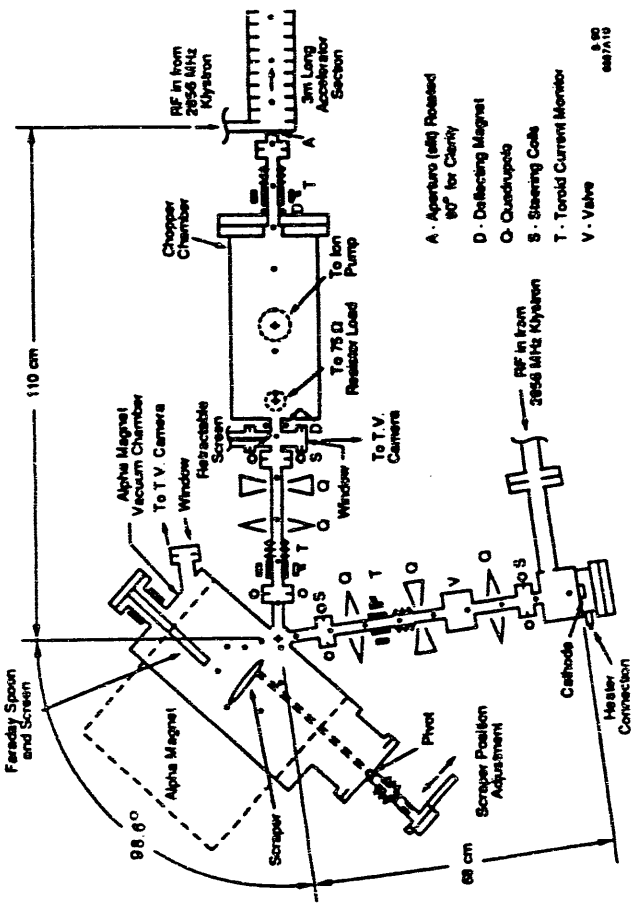

Figure 5.1: Gun-to-Linac Region of the SSRL Preinjector

three to five bunches for injection into the linac out of the train of several thousand bunches that is emitted by the gun during each RF pulse. I will not discuss the chopper further here.

5.1.2 Gun-to-Linac Instrumentation

Several types of diagnostic instrumentation are available in the GTL[78, 79]. Forward and reflected RF power signals are available from a directional coupler in the rectangular waveguide leading to the gun. The coupler is not as close to the gun as one might like: originally, the coupler was quite near the gun, but the ceramic window covering the coupling hole broke, forcing us to find a new position for the coupler. The coupler used for these experiments is about 1 m from the gun body, with a wave-guide "H-bend" and a high power RF window between the gun and the coupler. These signals are converted into voltages by calibrated diodes, so that one obtains a signal that is related to the power envelope of the RF signal in question. Because of the H-bend between the coupler and the waveguide, reflected power signals will include the effect of any mismatch between the H-bend and the rectangular waveguide leading up to it, instead of simply being a product of the match between the gun cavity and the waveguide. Cold-test measurements (performed by others) indicate, however, that the match of the H-bend to the waveguide is very good.

Beam-current measurements can be obtained from either of two toroids, one before the alpha magnet ("GT1") and one following it ("GT2"), as well as from a "Faraday-spoon" mounted inside the alpha magnet. In addition to being the current monitors that are used in day-to-day operation of the GTL, the toroids find application in the measurement of momentum spectra. As I will discuss in more detail below, by measuring beam transmission from GT1 to GT2 as a function of scraper position, one can obtain the momentum spectrum and the beam power.

The toroids are essentially transformers, with the beam current acting as one winding of the transformer. The name "toroid" derives from the use of a ferrite torus as the core of the transformer. The torus is mounted symmetrically around the path of the beam, so that the beam goes down the axis of the torus. At the position of the torus, there is a ceramic break installed, since otherwise the image currents induced

in the metallic vacuum pipe would cancel the beam current. The other winding of the transformer consists of wire coiled around the ferrite. A shield is also installed around the torus, providing a path for image currents that goes around the outside of the ferrite. Because this image current is equal in magnitude and opposite in sign to the beam current, the effect is equivalent to having the beam current make a loop around the ferrite. Hence, the "turns ratio" of the transformer is 1 :: n, where n is the number of turns of wire wrapped on the ferrite.

The Faraday-spoon is essentially a copper target that is held inside the alpha-magnet in such a position that is is struck by the beam whenever the alpha-magnet is turned off. The target is supported by copper tubes (which double as conduits for cooling water), which are themselves eventually supported by an electrically isolated vacuum feed-through. Hence, the electrons that are absorbed by the Faraday-spoon do not return immediately to ground, but can rather be made to return to ground through an external resistor. The voltage across this resistor is proportional to the current absorbed by the Faraday-spoon, from Ohm's law. The surface of the Faraday spoon where the electrons impact the copper was not shaped into a cup, since for electrons of 2—3 MeV, back-scattering should be negligible.

Mounted on the front surface of the Faraday-spoon is a phosphorescent screen, which emits visible light when struck by electrons. This screen, referred to as the "alpha-magnet screen", is viewed by closed-circuit TV through an optical window in the alpha-magnet vacuum chamber. This screen is heat-sunk to the Faraday-spoon, with the front surface of the Faraday-spoon being very slightly convex in order to ensure good thermal contact between the screen and the copper.

Another phosphor screen is installed down-stream of the alpha-magnet, in front of the chopper. This screen, referred to as the "chopper screen", is retractable and, unlike the alpha-magnet screen, is not water-cooled. Hence, while the alpha-magnet screen can take full beam current (more than 900 mA) without apparent harm, the chopper screen is easily damaged by too much beam current, and in addition will cause an unacceptable increase in GTL and gun vacuum pressures if inserted with more than about 50 mA current exiting the alpha-magnet. Because it follows the alpha-magnet, the chopper screen provides the only way to measure the emittance of the beam without the inaccuracy that would occur if the full momentum spread were

used. These considerations (among others, such as phosphor saturation) combine to militate against measurements of beam-emittance at high currents.

Rather than rely on bench-top calibrations of the toroids, I chose to calibrate the toroid signals (i.e., in amperes of beam current per volt of output signal) relative to the Faraday spoon. In particular, this was done by first optimizing beam transmission to the Faraday spoon. By viewing the beam spot on the alpha-magnet screen, it could be verified that no beam-scraping was occurring, which implies that the current measured by the Faraday cup is the same as that passing through GT1. Simultaneous measurements of the peak Faraday cup signal versus the peak GT1 signal were then taken as the current from the gun was varied, thus providing a calibration of GT1. (Note that since the beam current increases during the RF pulse, the peak for each signal occurs near the end of the pulse.) For a fixed-length and fixed-shape beam pulse, this procedure includes the effect of the L/R time-constant of the toroid response (estimated to be 10 μs[80]), since the toroid is a linear device. For a shorter beam pulse than that used for the calibration, use of the same calibration would result in an over-estimate of the beam current. I will return briefly to this issue in the Section 5.3, where the measurement of spectral distributions is discussed.

5.1.3 Beamline Control and Data Acquisition

External to the concrete vault that houses the GTL is a computer-aided measurement and control system, based partly on an IBM PC compatible and partly on a DEC MicroVax II. The PC, in concert with a LeCroy 9450 digital oscilloscope (accessed via GPIB) and a MetraByte DAS-20 Data Acquisition Board, collects and stores data and controls the alpha-scraper through a stepper-motor driver. On-line, real-time analysis of data is available through software that I wrote specifically for the experiments, and through a subset of the `mpl` Scientific Toolkit (described in Appendix 1). In particular, any waveform acquired from the oscilloscope may be plotted or subjected to various types of analyses and transformations (such as fitting, Fourier analysis, integration, differentiation, and noise subtraction) on the PC. Scope parameters may be stored on the PC and later restored to the scope under user or program control. Data taken for momentum spectra may be immediately analyzed and plotted on the PC.

The LeCroy 9450 oscilloscope can be synchronized with the sweep of the close-circuit TV cameras that are used to view the phosphorescent screens. Hence, it is possible to use the combination of the scope and PC as a "frame grabber", allowing storage of screen images for later analysis. As I will describe in more detail below, these frames were transferred to SSRL's VAX 8700 for image analysis to determine beam sizes.

The MetraByte Data Acquisition Board is used for sampling relatively slow signals and for controlling the position of the alpha-scraper. The board has eight differential A/D converter inputs, with a maximum sampling rate of 100 kHz (distributed over however many channels are sampled), as well as digital I/O and 2 D/A convertors. I have developed software to provide versatile measurement capabilities using this board. For example, an eight-channel "chart recorder" (`chart`) is available, as is a 100-kHz digital oscilloscope program (`scope`), and several programs (`dchart`, `ramp`) for automated and semi-automated magnetic measurements (or analogous measurements) using a stepper-motor or computer-controlled power supply. I used the later programs for all the magnetic measurements on the GTL. The chart-recorder program is used to monitor the gun filament current and voltage, as well as vacuum pump currents; this is particularly useful during cathode processing. For RF processing, a program (`bursts`) is available that counts vacuum bursts, to help the experimenter assess progress; as bursts are seen to decrease (or increase) in frequency, the experimenter can increase (or decrease) the RF power.

While the PC was used for all data acquisition and some control functions, primary control of GTL components was through the Injector Control System (ICS). Only the alpha-scraper was controlled from the PC. No link was available between the PC and the ICS, and hence some experiments were only semi-automated. For example, emittance all measurements required grabbing screen images as a quadrupole was varied. The experimenter was required to act as the "link" between the PC and the ICS, telling the PC when to grab a frame, and the ICS when to change the quadrupole strength. Also, data transfer from the PC to SSRL's main computer, a VAX 8700, was accomplished by copying the data from the PC to diskettes, from which they could be transferred from another PC to the VAX 8700 over a terminal line. This is not a trivial matter, since the TV frame files are quite large, so that the transfer

might take several hours for the data from a single shift of experiments. Since the floating-point representations on the PC and the VAX 8700 are different, the frame files were converted from the binary format used by the LeCroy 9450 to a packed ASCII format before being transferred.

5.1.4 The Preinjector Linear Accelerator

The SSRL Preinjector utilizes three SLAC-built 30 MW, XK5 S-Band klystrons, each driven by a separate modulator, with all the modulators sharing a single high-voltage power supply. Each klystron supplies RF power to one of the linear accelerator sections, with the second klystron (i.e., the klystron for the second accelerator section) also supplying RF to the gun via a 7-dB coupler. (Since the experiments, the coupler has been replaced with a 5-dB coupler, to supply more power to the gun.) Downstream of the 7-dB coupler is a high-power attenuator, followed by a high-power phase-shifter and then the gun. The RF drive for the second klystron is derived from an oscillator which drives a 1 kW RF amplifier. The RF drive for the other two klystrons is derived from couplers inserted in the waveguide from the second klystron. This configuration has the advantage that only one RF amplifier is needed. Also, since the power for the gun is taken from the power going to the second linac section, the first section has the full power of the first klystron, giving rapid initial acceleration in the first section.

The disadvantage of this configuration is that all RF phases are referenced to the second section, when a more congenial configuration would have all RF phases referenced to the gun. In order to phase up the linac, on must first phase the first section relative to the gun, in order to get beam into the second section; this is done by adjusting the low-power phase-shifter for the RF drive to the first klystron in order to maximize the signal on a toroid between sections 1 and 2. One then phases the gun to the second section using the high-power phase-shifter, while adjusting the phase of the first section to keep it in phase with the gun. Finally, one phases the third section to accelerate the beam exiting the second section. Any change in the phase of the bunches coming from the gun, such as will occur whenever the beam momentum changes due to beam-loading or whenever the alpha-magnet gradient is

changed, requires rephasing all three sections. This creates difficulties for certain experiments, as I will discuss later in this chapter.

5.2 Gun-to-Linac Optics

In this section, I discuss the GTL optics. While this topic is not really an experimental one, it is relevant to the experiments, and so I have chosen to treat it in this chapter. I will present some experimental data, namely magnetic and optics measurements on the GTL quadrupole.

5.2.1 Modeling of the Quadrupoles

The optical properties of the alpha-magnet are discussed in Chapter 3. The remaining optical elements are the five quadrupoles. While I will not go into the details of magnetic measurements on these quadrupoles, Figure 5.2 shows the gradient as a function of z, the longitudinal position. Also shown in the Figure is a trapezoidal approximation to this distribution. This trapezoid was obtained by first setting the endpoints of the trapezoid to lie at the same z as the 10% points of the actual gradient. The flat-top of the trapezoid was required to be of the same height as the actual flat-top, with the length of the flat-top adjusted to make the area under the trapezoid the same as the area under the actual distribution. That is,

$$\int g(z)dz = g_o l_{top} - g_o l_{fringe}, \tag{5.1}$$

where $2l_{fringe} - l_{top}$ is the distance between the two z locations where $g = g_o/10$. For the GTL quadrupoles, $l_{fringe} = 3.28$cm and $l_{top} = 4.43$cm. There is no compelling reason for choosing l_{fringe} as I did (i.e., based on the points at which $g = g_o/10$): it simply seemed a "reasonable" choice. Later in this section, l_{fringe} will be adjusted based on beam-optics experiments.

The equation of motion (see [6, 10] for background material) in the entrance fringe-field of the trapezoidal model is

$$x'' = -\frac{k_o}{l_{fringe}}\frac{xz}{1-\delta}, \tag{5.2}$$

where primes represent derivatives with respect to z,

$$k_o(m^{-2}) = 0.029979\frac{g_o(kG/m)}{pc(GeV)}. \tag{5.3}$$

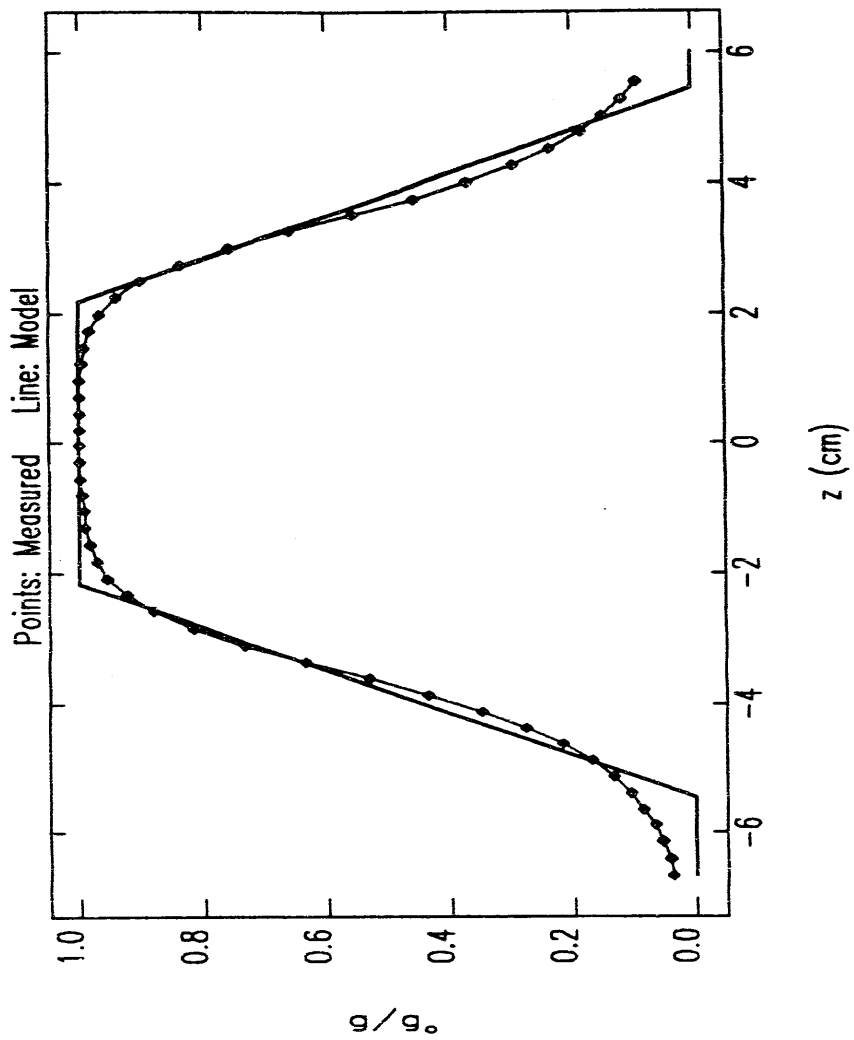

Figure 5.2: GTL Quadrupole Gradient vs Longitudinal Position

and where $\delta = (p - p_o)/p_o$ is the fractional momentum deviation. Equation 5.2 is valid for either positive or negative k_o, i.e., for either a focusing or defocusing quadrupole, respectively. It can be solved by a series of the form

$$x = \sum_{n=0}^{\infty} a_n z^n. \tag{5.4}$$

Inserting this into the differential equation yields

$$a_{3n} = a_0 \left(-\frac{k_o}{1-\delta}\right)^n \prod_{m=1}^{n} \frac{1}{3m(3m-1)} \tag{5.5}$$

$$a_{3n-1} = a_1 \left(-\frac{k_o}{1-\delta}\right)^n \prod_{m=1}^{n} \frac{1}{3m(3m-1)} \tag{5.6}$$

$$a_{3n-2} = 0, \tag{5.7}$$

where $n = 1, 2, \ldots \infty$. Evaluation of the series and its derivative at $z=0$ reveals that $a_0 = x(0)$ and $a_1 = x'(0)$, which completes the solution. The solution for $x'(z)$ is obtained by simply taking the derivative of $x(z)$.

The first-order matrix-elements ($r_{11}, r_{12}, r_{21}, r_{22}$) are found by taking derivatives of $x(z)$ and $x'(z)$ with respect to $x(0)$ or $x'(0)$, with $\delta = 0$. The non-zero second-order matrix elements (t_{126}, t_{216}) are found by taking derivatives with respect to δ and then with respect to either $x(0)$ or $x'(0)$. For the exit fringe fields, one obtains the matrix by finding the reverse matrix [6] of the matrix for the entrance fringe fields. The chromatic terms clearly swap in the same way as the first-order matrix terms from which they derive. (I use the unconventional lower-case letter "t" for the transport matrix for consistency with Chapter 3.)

I have incorporated these results into the tracking/integrating code elegant [49], which I wrote specifically for the SSRL RF gun project. For matching purposes, I have used the program MAD [71], modified to include alpha-magnets. The quadrupole fringe-fields in MAD were simulated by breaking up each of the fringe regions into four constant-gradient sections. This was found to give good agreement with trials done with up to 50 constant-gradient sections (which were also used to confirm the exact solution given above). Later in this section, I will show results of experimental tests of this quadrupole model.

5.2.2 Optical Matching

Matching of the optics in the GTL proceeded from several assumptions and under several constraints:

1. The output phase-space of the gun varies with beam current and fields in the gun. Hence, for matching, I used an initial betatron ellipse that enclosed all of the useful beam (as defined in Chapter 2) for the full range of operating conditions. This ellipse has parameters $\beta = 0.074$m, $\alpha = 0$, and $\epsilon = 27\pi$mm $-$ mrad.

2. Various beam-pipe and other chamber apertures must be accommodated:

   (a) The beam-pipe radius is 18mm throughout most of the GTL, except where otherwise noted in this list.

   (b) For 11 cm before and after the alpha-magnet, the beam-pipe radius is 11mm.

   (c) Inside the alpha-magnet, the aperture is highly irregular (due to the shape of the magnet poles) but large enough to safely ignore.

   (d) The horizontal aperture in the chopper is $\pm$ 10mm and the vertical aperture is $\pm$ 4mm.

   (e) There is a 30mm-long, 4mm-diameter differential pumping aperture immediately after the gun, and a slightly-constricting aperture near toroid GT1, which acts as shield for the ceramic. (The later was added well after the matching studies, and as it does not alter them, I have ignored it.)

3. A gently convergent beam in both planes is desirable at the entrance to the linac, in order to match to the linac transverse acceptance (i.e., to avoid scraping beam on the apertures inside the linac). In addition, the vertical beam size should be less than 2mm at the final chopper slits (just before the linac) in order to obtain proper chopping [77], and the horizontal beam size at the linac entrance should be less than 9mm

4. The alpha-magnet gradient is dictated solely by the needs of bunch compression (see Chapter 4), and hence the rest of the optical elements must be adjusted to

accommodate the alpha-magnet. The alpha-magnet matrix can be expressed as a function of the size of the alpha-shaped loop made by the central particle, $\hat{q}_1$ (which is inversely proportional to the parameter α used in Chapter 3, and which is also called x_{max} in some of the literature[45, 32]). Similarly, the compression is related to this parameter, so that in order to accommodate a wide range of alpha-magnet settings, I did the matching for a number of values of $\hat{q}_1$ between 9cm and 15cm. Some results of this matching are shown below.

5. It was desired that the momentum resolution of the alpha-magnet momentum filter be a few percent. The momentum resolution is roughly[6]

$$\delta_{res} = \frac{\sqrt{\epsilon\beta}}{D}. \tag{5.8}$$

where the dispersion at the midpoint in the alpha-magnet is $D = \hat{q}_1/2$, and is typically 4-8 cm, depending on how the alpha-magnet is set for pulse compression and what the central momentum is. Hence, β at the midpoint is required to be less than 0.006 m for 1% resolution. Typically, I achieved better than 2% nominal resolution. Baring excessive second-order effects, the actual resolution should be better than the nominal resolution, since the initial phase-space is smaller than what I used for the matching. Second-order simulations with elegant confirm this expectation, as I shall show in the next section.

Table 5.1: **Quadrupole Strengths for GTL Optics Solutions for Various Values of $\hat{q}_1$**

$\hat{q}_1$ (cm)	k_o for Q1 ($1/m^2$)	k_o for Q2 ($1/m^2$)	k_o for Q3 ($1/m^2$)	k_o for Q4 ($1/m^2$)	k_o for Q5 ($1/m^2$)	$\delta_{res,nom}$ (%)
9.00	244.04	-142.95	150.50	107.71	-84.24	2.0
10.00	234.39	-142.05	150.08	105.61	-83.54	1.9
11.00	229.30	-141.17	148.96	104.01	-83.04	1.9
12.00	224.88	-140.36	147.76	102.54	-82.55	1.8
13.00	222.18	-139.60	146.37	101.25	-82.12	1.8
14.00	210.71	-138.74	145.92	99.43	-81.31	1.7
15.00	225.68	-138.36	142.71	99.44	-81.58	1.7

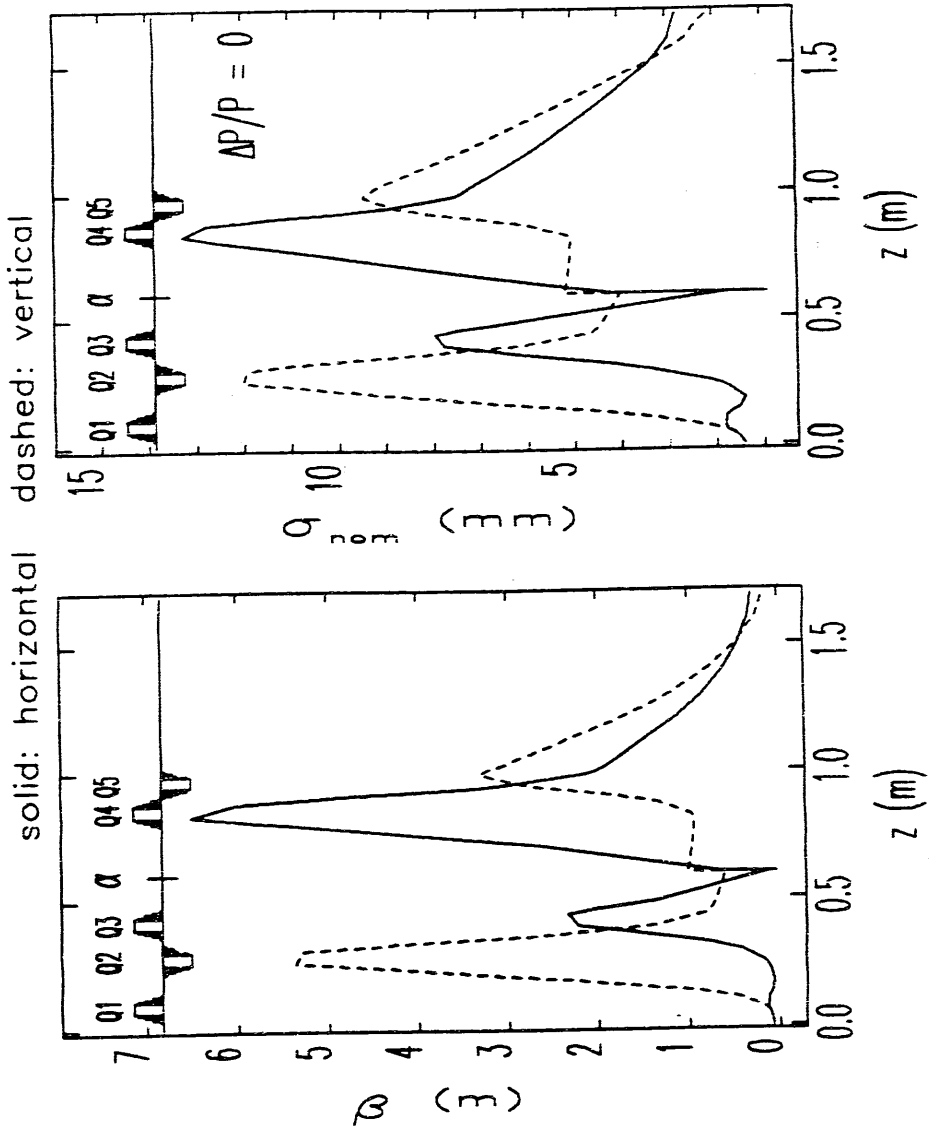

Figure 5.3: GTL Optics Solution for $\hat{q}_1 = 11$cm

Using the results of Chapter 3, I have added alpha-magnets to MAD. Figure 5.3 shows a "typical" optics solution for the GTL, in this case for $\hat{q}_1$ = 11cm. The alpha-magnet appears as a zero-length element at z = 0.6m. The quadrupole strengths for this solution and several similar ones for different values of $\hat{q}_1$ are listed in Table 5.1. These optics solutions are designed for the case when one wants to accelerate the beam in the linac. When one is interested only in doing experiments in the GTL, some of the constraints are changed or eliminated, resulting in qualitatively different solutions. I will discuss such optics solutions as needs dictate.

5.2.3 Higher-Order Effects

These optics solutions assume that a first-order treatment of the beam-optics problem is adequate. This is known not to be the case for the GTL, because of the large momentum spread of the gun beam. One of the design goals, as discussed in Chapter 2, was to be able to make use of ± 10% momentum spread out of the gun, in order to increase the useful beam current. Because the focal length of a quadrupole of a given gradient is inversely proportional to the momentum, a ± 10% variation in momentum implies a ±10% variation in quadrupole focal length. Hence, a lattice that is matched for the "on-momentum" particles in the beam may well be seriously mis-matched for the off-momentum particles in the beam. The best way to investigate the seriousness of this problem is to do second-order tracking, which I have done using **elegant**.

In particular, Figure 5.4 shows the results of first- and second- order tracking with **elegant**. The initial particle distribution was obtained from a **MASK** simulation (with E_{p2} = 75MV/m and J = 10A/cm^2, with the particle momenta being pre-filtered to ensure that the beam-sizes refer to the desired ±10% momentum spread all along the beamline. (Had this not been done, the beam-sizes would have dropped dramatically after the alpha-magnet, due to the filtering action of the scraper.) The optics is the same as that used to make Figure 5.3; note, however, that the beam-sizes will not agree, since the initial conditions assumed for Figure 5.3 are general acceptance parameters which do not match the initial conditions for any particular beam from the gun.

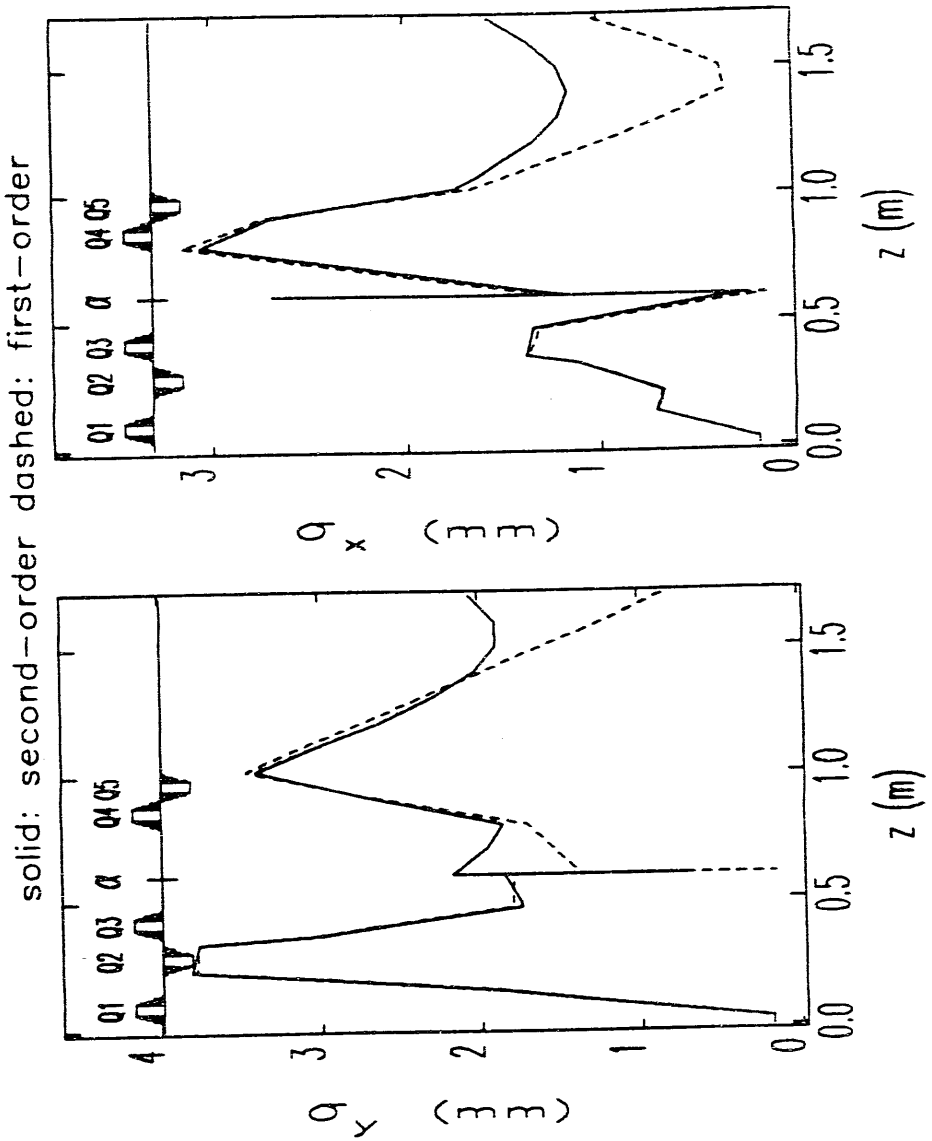

Figure 5.4: Beam-sizes for GTL Optics Solution for $\hat{q}_1$ = 11cm, from First- and Second-Order Tracking with **elegant**, for **MASK**-generated Initial Particle Distribution. $\Delta P/P$ = 10%.

Because the focusing is different for different momenta, the transverse beam emittance must change as the beam goes through the GTL. To see this, consider a beam with initial phase-space described by $x' = \alpha x$ and some spread of momentum deviations, $\delta = (p - p_o)/p_o$, where I assume that $\langle x\delta \rangle = 0$. The RMS emittance for such a distribution is zero, as can be verified using the definition of the emittance:

$$\epsilon = \pi\sqrt{\langle x^2\rangle\langle x'^2\rangle - \langle x'x\rangle^2} \tag{5.9}$$

If this beam goes through a thin-lens quadrupole with focal length f_o for the central momentum, then the coordinates after the quadrupole are

$$\bar{x} = x \tag{5.10}$$

$$\bar{x}' = x\left(\alpha - \frac{1-\delta}{f_o}\right), \tag{5.11}$$

where I have expanded the focal length to first order in δ. The emittance for this new distribution is

$$\bar{\epsilon} = \pi\frac{x_{rms}^2\delta_{rms}}{f_o}. \tag{5.12}$$

For a beam with δ uniformly distributed on $[-\delta_o, \delta_o]$, $\delta_{rms} = \delta_o/\sqrt{3}$. Thus, for example, if $x_{rms} = 10$mm and $\delta_o = 0.1$, then for $f_o = 0.5$ m, one sees that $\bar{\epsilon} \approx 12\pi$-mm · mrad. This indicates that emittance blow-up due to chromatic aberrations should be significant in the GTL, as the simulations confirm. In particular, Figure 5.5 shows the emittance for the x and y planes as a function of position in the GTL, for the same lattice and beam conditions used in the previous Figure. Note that the changes in the emittance are much more apparent than the changes in the beam sizes. There are several reasons for this. First, the chromatic aberrations in the quadrupoles change the emittance by changing the angular coordinates, so that a large change in the emittance at the exit of a quadrupole does not entail a large change in the beam-size. Only after the beam has traveled some distance down the beam-line will the beam-size change significantly. Second, the beam-size scales like the square-root of the emittance, so that fractional changes in the beam-size are expected to be generally less than fractional changes in the emittance.

It is difficult to characterize the emittance degradation in general, since it depends on the initial phase-space, the lattice, and the momentum spread under consideration.

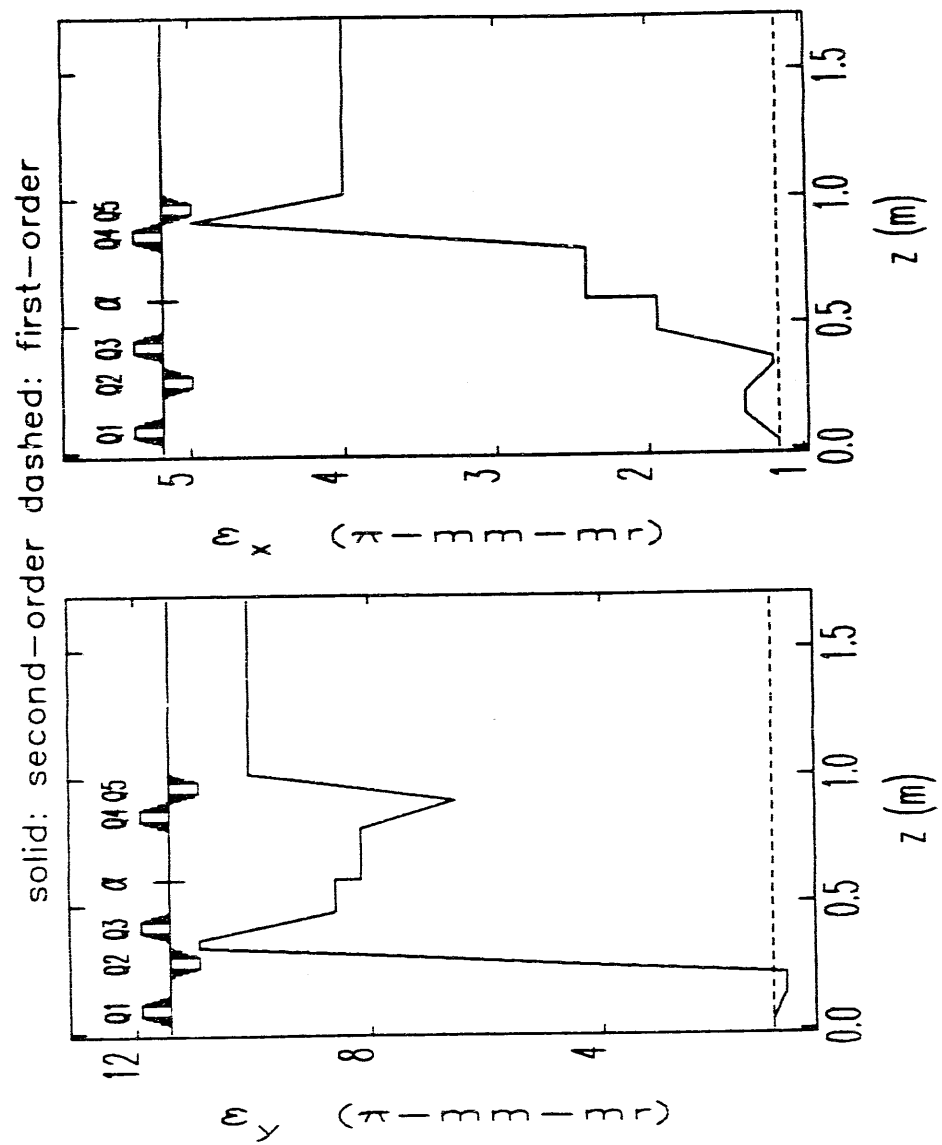

Figure 5.5: Emittance Degradation for GTL Optics Solution for $\hat{q}_1 = 11$cm, from First- and Second-Order Tracking with **elegant**, for MASK-generated Initial Particle Distribution. $\Delta P/P = 10\%$.

However, the results shown in Figure 5.5 are typical of the effects on MASK-generated initial particle distributions for the lattices listed in Table 5.1, when ± 10% momentum spread is taken from the initial distribution. In general, these lattices blow up the x-plane emittance to about twice what the y-plane emittance is blown up to. Later in this chapter, I will show comparable data to that in Figure 5.5 for the lattice used in the emittance measurements, where in contrast the relative degradation of the two planes is reversed. Because of such difficulties, I will simply give a rough characterization of the emittance at the end of the GTL by stating that the geometric mean of the x- and y-plane emittances, $\epsilon_m = \sqrt{\epsilon_x \epsilon_y}$, is 5–6 $\pi \cdot$ mm $\cdot$ mr. This result is indepedent of the current density used in the MASK simulation, because chromatic effects after the gun overwhelm space-charged induced emittance growth in the gun.

I have not investigated the possibility of finding a lattice that has a minimal effect on the emittance while still satisfying the constraints listed above. The optics is to a large extent dictated by the need for small horizontal beam-sizes at the vertical midplane in the alpha-magnet (for momentum filtration), and by the long drift-space between Q5 and the linac (required for the chopper). Creating a small horizontal beam-waist inside the alpha-magnet requires a large beam-size at Q3, which requires strong Q1 and Q2. Similarly, to create beam-waists at the linac entrance requires large beam-sizes at Q4 and Q5, which in turn requires strong focusing prior to Q4 and Q5. As equation (5.12) illustrates, the emittance-degrading effect of a quadrupole increases with increasing beam size and quadruple strength.

5.2.4 Experimental Tests of the Quadrupole Model

Finally for this section, I discuss the results of experiments designed to test the modeling of the GTL quadrupoles. This is important, since the quadrupoles are used for emittance measurements, as reported in a later section of this chapter. The present experiment involved GTL quadrupoles Q4 and Q5, as well as the steering magnet (GTL_CORR3, which provides independent horizontal and vertical steering) between the alpha-magnet and Q4. The current in the steering magnet was varied in order to vary the angular "kick" it imparted to the beam. This kick in turn causes the beam position at the chopper screen to vary. By measuring the movement of the

beam spot as a function of the kick, one can probe the optics of the elements between the steering magnet and the screen. In particular, by making vertical deflections one can determine the r_{34} matrix element for beam transport from the steering magnet's principle plane to the screen[6, 10]. I confine myself to the vertical plane, since the horizontal and vertical planes are equivalent in this case (the quadrupoles have bipolar power supplies), and since there were problems with the steering magnet power supply for the horizontal plane when I took the data.

The measurements made use of the LeCroy 9450 oscilloscope's ability to trigger off of and digitize TV signals, so that the position of the center of the beam spot could be determined accurately using the cursors on the scope. This yields positions in terms of sweep time, which can be converted to Δy relative to the position for zero kick-angle by making use of a calibration of the sweep obtained by noting the positions in the sweep of two identifiable features on the screen, the distances between which are known. For the present work (both for the r_{34} measurements and the emittance measurements reported later), the sweep was calibrated by directing the beam to the edge of the screen, thereby producing a sharp feature that was readily identified on the oscilloscope. Since the calibration changed from experiment to experiment (usually because the camera was "adjusted" or bumped), I will not record the calibrations here.

With quadrupoles Q4 and Q5 turned off (and degaussed), the transport matrix from GTL_CORR3 is simply that for a drift space, and hence $r_{34} = z_s - z_k$, where z is the longitudinal coordinate and the subscripts stand for Screen and Kick, respectively. Ideally, one would like to know z_k and z_s accurately, and use the measurement with Q4 and Q5 off to calibrate the kick vs current. This is what was done for the experiments reported on in the last section of Chapter 3, where I discuss how to calculate the offset of z_k from the geometric center of the steering magnet (the offset occurs because of the long fringe fields). Unfortunately, mechanical difficulties prevented GTL_CORR3 from being placed at the expected location, and it instead had to be placed much closer to Q4 and to toroid GT2 (which has a ferrite core) than was expected. Hence the magnetic measurements made for this steering magnet prior to installation are not valid, and I simply assumed that the principle plane is 1.8 ± 1cm ahead of the mechanical center the steering magnet (the offset for GTL_CORR2 was found to

be 1.8cm). With this assumption, I calibrated the steering kick vs current from data taken with Q4 and Q5 off by requiring that $r_{34} = z_s - z_k = 0.274 \pm 0.010$m and assuming the kick to be linear in the current (a reasonable assumption for these low-field steering magnets).

Several factors need to be taken into account in assessing the results with the quadrupoles on. The beam from the gun has a very large momentum spread (see Chapter 2, and the next section of this chapter), and hence I set the alpha-scraper to filter out a fairly small momentum spread (about $\pm 3\%$), consistent with getting a sufficiently intense beam spot. The magnet settings (i.e., the nominal k_o values of the quadrupoles) were referenced to some nominal "lattice" momentum that was not the same as the actual beam momentum. Correction for this small effect is made by multiplying the nominal quadrupole strengths by 0.97 ± 0.01 to get the actual k_o values. This also introduces the uncertainty in the beam momentum into the quadrupole strength.

Figures 5.6 and 5.7 show the results two sets of measurements. The first set (Figure 5.6) was made with quadrupole Q5 turned off and degaussed; the strength of quadrupole Q4 was varied, and data was taken to allow r_{34} to be deduced by fitting a line to $(\theta, \Delta y)$. Similarly, the data in Figure 5.7 was taken with quadrupole Q4 turned off and degaussed, with the strength of Q5 being varied. Also shown in the Figures are the results of matrix calculations (done with **elegant**) made using the trapezoidal quadrupole model discussed above. Each figure shows the simulation result for the quadrupole model with the nominal parameters given above (i.e., $l_{fringe} = 3.28$cm and $l_{top} = 4.43$cm), as well as a simulation result obtained with l_{fringe} changed to 3.00 cm.

This latter result is an approximate best fit (in the least-squares sense) to the data, taking l_{fringe} as the parameter to be fit. This best-fit model has approximately 4% less integrated strength than the model with nominal parameters. It is a compromise between best fits for the individual data sets, and hence isn't a best fit for either data set by itself. While it is possible to fit the two data sets separately, I decided to find a single fit that matched both reasonably, to make subsequent work less complicated.

The point at $k_o = 81$ m^{-2} for Q4 was considered spurious and was ignored in the fitting; it would appear that the quadrupole was set incorrectly when this data point was taken. It may appear that this point could be the result of saturation. However,

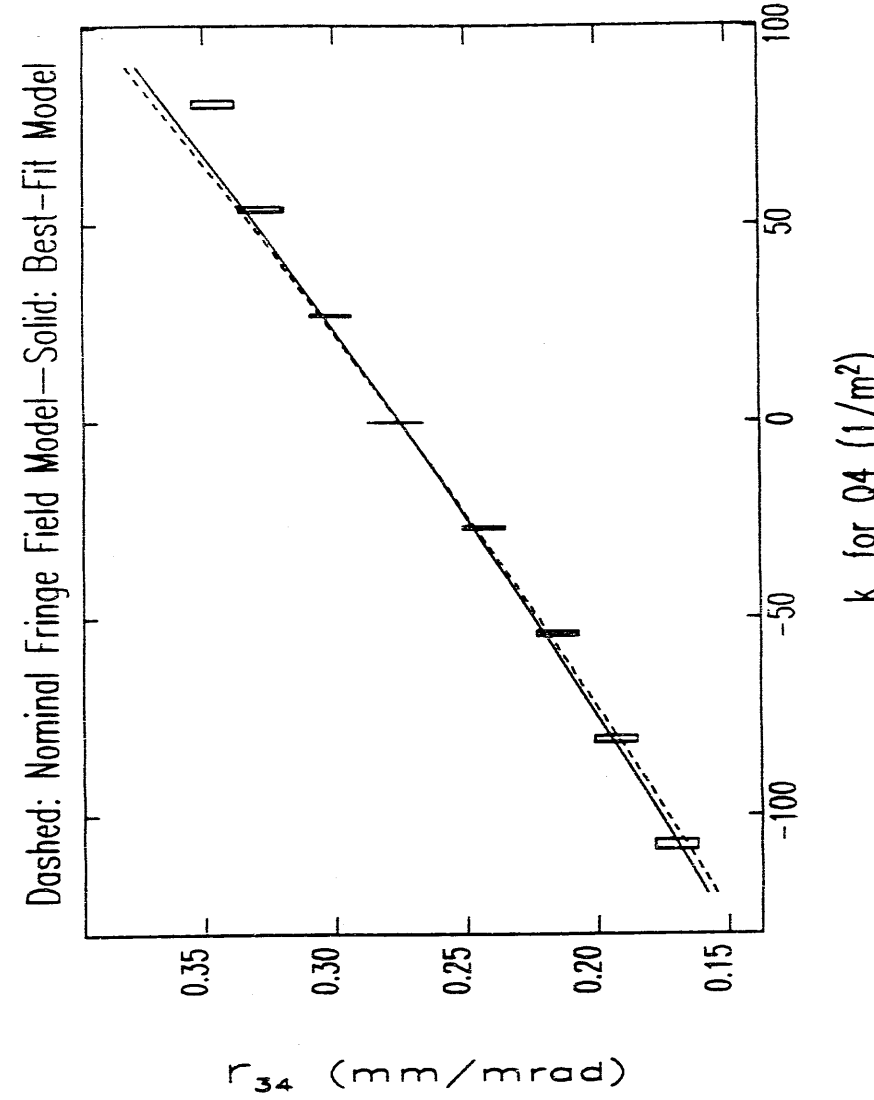

Figure 5.6: r_{34} vs Strength of Quadrupole Q4, with Q5 off.

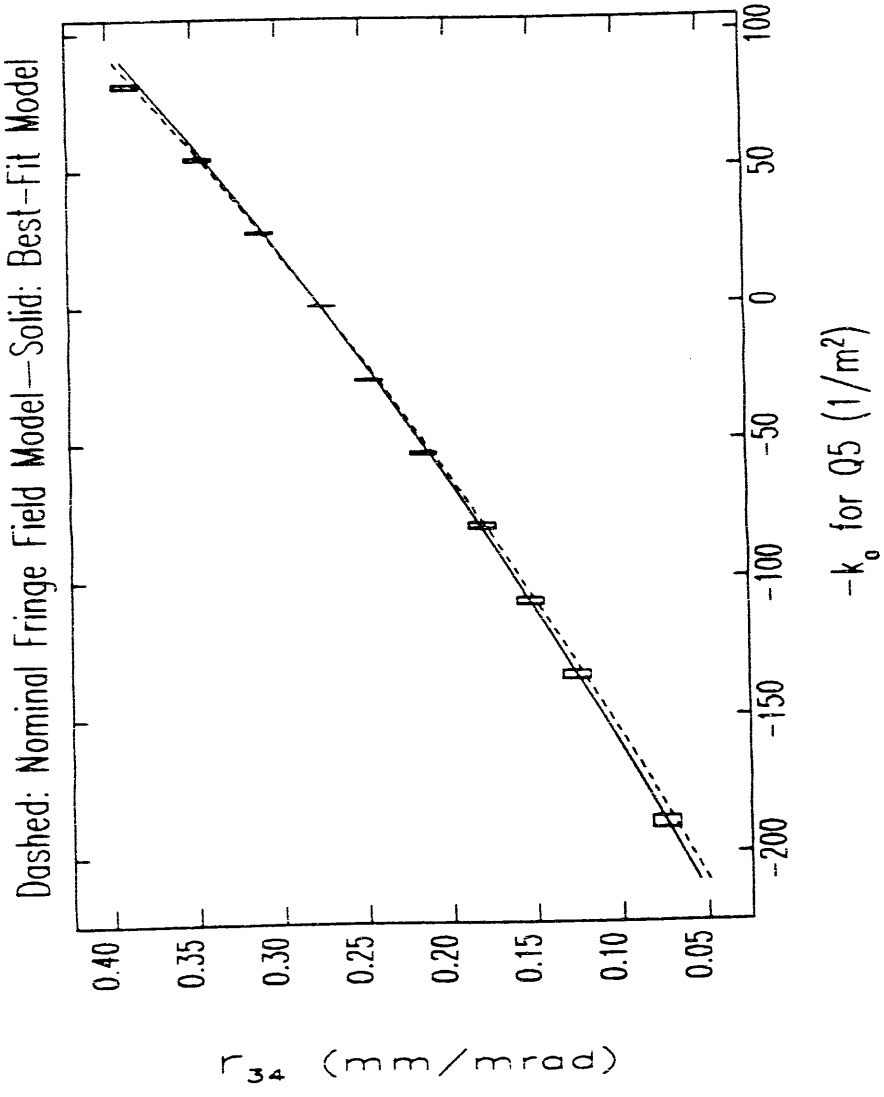

Figure 5.7: r_{34} vs Strength of Quadrupole Q5, with Q4 off.

saturation was taken into account in the calibration of k_o versus driving current, and is in any case a negligible consideration for these quadrupoles at the currents used.

There are a number of possible explanations for the discrepencies between the model and the experimental results. The measurements were made on a magnetically isolated quadrupole, whereas the quadrupoles are installed with other magnetic materials in close proximity. Both Q4 and Q5 have a magnet-steel picture-frame steering magnet less than 30 mm away. The magnets are supported by rails that are made of magnetic steel. Finally, Q4 and Q5 are close enough that the fringe field of one quadrupole extends into the other quadrupole. Also, the nominal fringe-field model was itself guess, and should not be expected to agree exactly with the experimental data. The fact that I have fit the data using l_{fringe} as a parameter does not, of course, mean that the length of the fringe fields is actually responsible for the discrepancy, though it is the most likely choice. Whatever the source of the quadrupole strength error, the data are consistent with $l_{fringe} = 3.00$cm, and I shall use this value in the all of the analysis in this chapter.

5.3 Momentum Spectra

5.3.1 Principle of Spectrum Measurements

As I mentioned in the previous section, measurement of momentum spectra made use of current-measuring toroids GT1 and GT2 and the alpha-scraper. If the alpha-magnet gradient is g (in G/cm), then the size of the alpha-shaped trajectory executed by an on-axis particle with momentum $p = \beta\gamma$ is (see Chapter 3)

$$q_1(\text{cm}) = 75.05\sqrt{\frac{p}{g}}. \tag{5.13}$$

where q_1 is measured from the inner surface of the alpha-magnet front plate. (I will ignore the effects of less-than-perfect momentum resolution introduced by the non-zero beam emittance for now.) Thus, if the outer edge of the alpha-scraper is at $q_1 = q_{1,outer}$, then momenta above

$$p_{outer} = g\left(\frac{q_{1,outer}}{75.05}\right)^2 \tag{5.14}$$

will be allowed through the alpha-magnet. The inner edge of the alpha-scraper is at $q_1 = q_{1,inner} \approx q_{1,outer} - l_s$, where $l_s = 7.3$cm is the length of the absorbing copper block that does that actual scraping of the beam. (These relations are only approximately true, since the scraper swings in an arc rather than moving linearly in q_1, as I shall discuss presently.) As the scraper is moved toward larger q_1, low momentum particles can in principle begin to pass through the alpha-magnet by virtue of having $q_1 > q_{1,inner}$. Particles with momenta below

$$p_{inner} = g\left(\frac{q_{1,inner}}{75.05}\right)^2 \tag{5.15}$$

will be able to pass through. Hence, the scraper acts as a notch momentum filter, eliminating momenta between p_{inner} and p_{outer}. The current though GT2 (which is ideally proportional to signal measured on GT2) as a function of the "scraper position", $q_s = q_{1,outer}$, is thus given by

$$S_2(q_s) = \int_0^{p_{inner}(q_s - l_s)} \rho_p(p)dp - \int_{p_{outer}(q_s)}^{p_{max}} \rho_p(p)dp \tag{5.16}$$

where p_{max} is the maximum momentum that can be transmitted through the alpha-magnet and $\rho_p(p)dp$ is proportional to the number of particles in the range $[p, p - dp]$. p_{max} is determined by the position of the outer wall of the vacuum chamber, and is typically well beyond the upper momentum in the beam. $\rho_p(p)$ is, of course, "the spectrum" and is what is to be determined. For convenience, I choose to normalize $\rho_p(p)$ in such a way that

$$\int_{p_{min}}^{p_{max}} \rho_p(p) = 1, \tag{5.17}$$

where the integral is over the range of momenta that exit the alpha-magnet. Hence, $S_2(q_s)$ is the transmitted current with the scraper at q_s, normalized to the total transmitted current (i.e., with the scraper moved out of the beam).

5.3.2 Practical Considerations and Simplifications

Because of the limited momentum-acceptance of the beamline leading to the alpha-magnet, and because of the length of the alpha-scraper, it is a very good approximation to ignore the first term in equation (5.16). A typical gradient for a momentum spectrum measurement is 100-200 G/cm, with the maximum momentum in the beam being 2-3 MeV/c, or 4-6 m_ec. The worst-case is when the gradient is small and the peak momentum is large, since this produces the largest values of q_1 relative to the fixed size of the scraper. Hence, when $p_o = 6$ and $g = 100$ G/cm, one sees that $p_i \approx 1.9$. That is, when the scraper is in such a position so as to eliminate the highest momentum particles in the beam, then low-momentum particles of momenta up to 1.9 are also let through. These particles, however, never reach the alpha-magnet, being eliminated by the combined effect of the strong quadrupoles and beam-pipe apertures between the gun and alpha-magnet. This is confirmed by simulations and can be tested experimentally by noting whether transmission increases as the scraper is moved past the point at which the maximum beam momentum is being intercepted. In any given measurement, one can verify that there are no low-momentum particles getting past the scraper on the low-momentum side by verifying that the transmitted current does not increase until on moves well beyond the point where the maximum momentum in the beam is intercepted.

Similarly, by choosing the gradient properly, one can place p_{max} well beyond the

maximum momentum in the beam. Hence, through judicious choice of the experimental parameters, one can employ a simplified form of equation 5.16, namely

$$S_2(q_s) = \int_{p_{outer}(q_s)}^{\infty} \rho_p(p)dp \qquad (5.18)$$

In order to obtain $\rho_p(p)$, one simply takes the derivative of $S_2(q_s)$ with respect to p_{outer}:

$$\rho_p(p) = -\frac{\partial S_2(q_s)}{\partial q_s}\frac{\partial q_s}{\partial p_{outer}}. \qquad (5.19)$$

Thus, one can find the momentum spectrum, $\rho_p(p)$, from readings of the current on GT2 as a function of scraper position, i.e., by taking $S_2(q_s)$ as a function q_s.

There are several additional points that require discussion here:

1. The current into the alpha-magnet, as measured by the signal on the first toroid, S_1, is supposed to be held constant during such a measurement. Often there are slight drifts in S_1, due to slight drifts in RF power or filament power. Hence, it is best to normalize $S_2(q_s)$ to S_1, with both signals being read simultaneously after each motion of the scraper.

2. Because the gun current varies during the RF pulse, it is necessary to read both toroid signals at the same point in the pulse. Just what point that is depends on what one is interested in. Typically, I was interested in the momentum spectrum for the steady-state, and hence I read the toroids late in the pulse, but before the RF pulse had started to fall off. (This assumes that it is reasonable to neglect the inductive nature of the toroid response—see below.) Figure 5.8 shows two typical pulses for the current toroids, along with the corresponding RF pulses.

3. There is some pulse-to-pulse jitter in the toroid signals, as well as some baseline drift if the linac is not synchronized to the 60-Hz line. Hence, I averaged the toroid signals over (typically) ten pulses, and triggered the oscilloscope in such a way that the baseline could be read from the first part of the waveform (i.e., I pretriggered the scope ahead of the modulator pulse).

4. The toroid pulses in Figure 5.8 look quite clean, because I have subtracted stored RFI (radio-frequency interference) pulses. Had I not done this, the pulses would

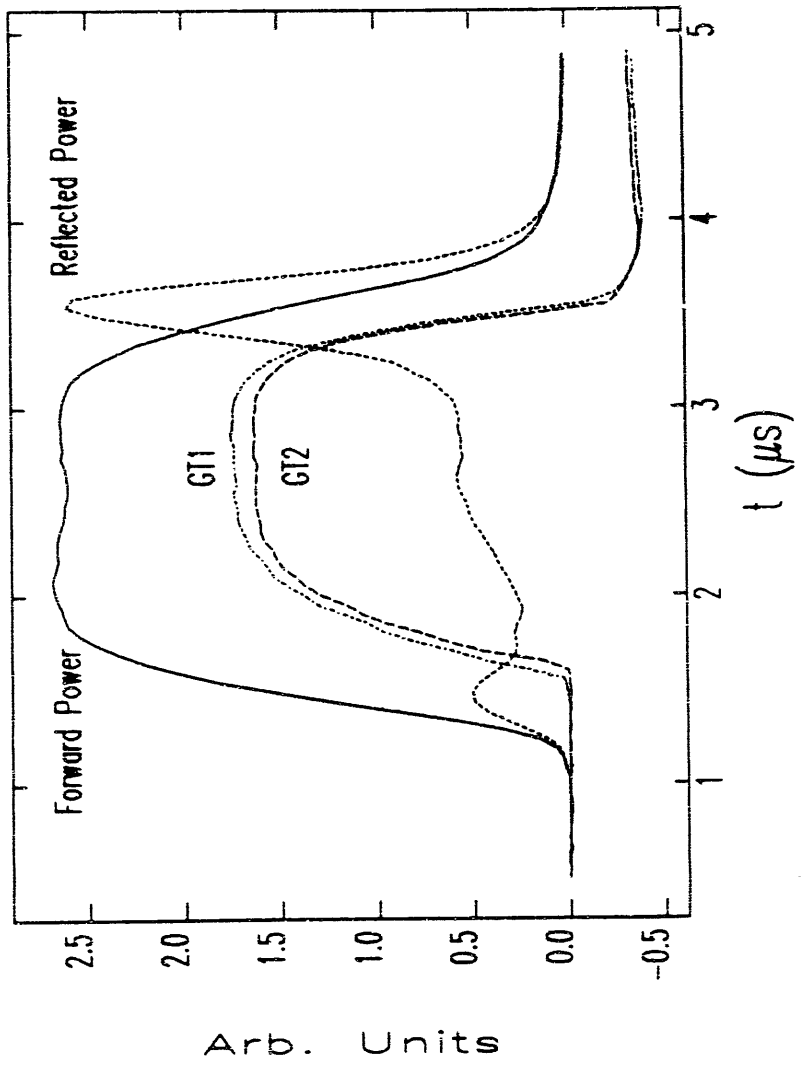

Figure 5.8: Typical Toroid and RF Pulses

show oscillations due to stray signals present in the experimental environment (mostly due to the firing of the thyratrons in the modulators). While the LeCroy 9450 permits subtraction of signals, it cannot simultaneously average a signal and do subtraction. Hence, the subtraction was performed on the PC, using previously sampled toroid "noise" signals taken with the RF drive turned off.

I have incorporated all of these considerations into the program spect3, which I have used for all of the spectrum measurements presented in this work. The program first reads an input file that specifies how much to move the scraper between readings, as well as giving other parameters of the measurement. It requests the user to turn off the RF power, after which it reads S_1 and S_2 to get RFI samples, which are subtracted from the toroid signals as the program takes the data. After turning the RF back on and waiting an appropriate interval for the gun emission to restabilize (typically less than 30 seconds, depending on the current), the user lets the program continue. It moves the scraper and reads S_1 and S_2 after each movement, averaging over as many pulses as the user requested in his input file, then subtracting noise. The signals read from the scope are equi-spaced samplings of whatever appears on the oscilloscope screen. The program averages over several (how many is specified by the user) samples at the very beginning of the waveforms to obtain the baseline, and over several samples at the very end of the waveform to obtain the desired current. Hence, the user must set up the scope so that the portion of the pulse he is interested in obtaining the spectrum for comes at the end of the waveform.

Some complexity is introduced into the measurements by the fact that the width of the beam pulse varies with scraper position, since the momentum spectrum changes as the fields build up in the gun. As mentioned in Section 5.1, the toroids have an L/R time-constant of 10 μs. The response to a square-wave current pulse is a decaying exponential with this time-constant. For a pulse of length $\tau \ll$ L/R, the response falls off to $1 - \tau R/L$ by the end of the pulse. The FWHM of the beam pulse on the Faraday spoon is at most 1.5 μs. Hence, the calibration of the toroids against the Faraday spoon, made for a full-length beam pulse, would be at most 15% in error for a very short pulse (i.e., $\tau \to 0$, provided τ is large compared to 2 ns, which is the response time of the ferrite[80]).

This issue was not appreciated at the time the experiments were done, and I must

content myself with estimating the likely effect on the measurements that have been taken. The measurements could have been taken using a Faraday cup at the alpha-magnet exit, to avoid the toroid response issue altogether. However, this would have interfered with commissioning efforts, and was not possible. In the future, it would be advantageous install an insertable Faraday cup after the alpha-magnet.

The toroid pulses were read near the end of the flat top (see Figure 5.8). Assuming a square current pulse, the signal on GT2 at the end of the pulse is (compare with equation (5.18)

$$S_2(p_s) = \left(1 - \frac{\tau(p_s)}{T}\right) \int_{p_s}^{\infty} \rho_p(p)dp, \tag{5.20}$$

where T=L/R and where $\tau(p_s)$ is the pulse width when particles of momentum greater than p_s are let through the alpha-magnet. (I have expressed S_2 as a function of p_s rather than q_s in order to simplify the analysis.) Taking the derivative with respect to p_s, one obtains

$$\frac{\partial S_2}{\partial p_s} = -\rho_p(p_s)\left(1 - \frac{\tau(p_s)}{T}\right) - \frac{S_2(p_s)}{1 - \frac{\tau p_s}{T}} \frac{1}{T} \frac{\partial \tau(p_s)}{\partial p_s}. \tag{5.21}$$

In my analysis, I have equated $\frac{\partial S_2}{\partial p_s}$ with $-\rho_p(p_s)$. To first order in τ, the fractional error made in doing this is

$$\frac{\Delta\rho_p(p_s)}{\rho_p(p_s)} \approx -\frac{\tau(p_s)}{T} + \frac{S_2(p_s)}{\rho_p(p_s)} \frac{1}{T} \frac{\partial \tau(p_s)}{\partial p_s}, \tag{5.22}$$

where $\Delta\rho_p > 0$ corresponds to over-estimation of ρ_p. Since $\tau(p_s)$ and $S_2(p_s)/\rho_p(p_s)$ are decreasing functions of p_s, the error in the spectral density *decreases* as p_s increases toward the maximum momentum in the beam. In fact, both $\tau(p_s)$ and $S_2(p_s)/rho_p(p_s)$ go to zero as p_s approaches the maximum beam momentum, p_{max}. Since the peak of the spectral distribution (see subsection 5.3.5) is near p_{max}, the error in the measurement of ρ_p will become increasingly small as p_s approaches the peak. This indicates that the relative error in the spectral densities measured near the momentum peak is significantly less than the 15% variation in the calibration, so that it is likely that measurements of the widths of momentum peaks are reasonably accurate.

As for the beam power measurments, note that the width of the current pulse varies from 1.5 μs to essentially zero as the scraper is moved into the beam. Hence,

the error in beam power measurements is less than 15 %, and probably more like half of that.

As I mentioned above, the scraper does not move only in the q_1 direction, but rather swings in an arc from a pivot point on a lever arm $L_a = 46$cm long. The length of the chord between the two extremal positions of the scraper is 8.3 ± 0.14cm. Hence, the angular range is 10°, so that the motion in q_1 is linear with the angle to within less than 0.5%. The angular motion is accomplished through a stepper-motor mechanically connected to the scraper-arm in such a way that the motor pivots through the same angle as the scraper. Hence, as the motor pulls itself toward (or pushes itself away from) the pivot block into which its threaded shaft runs, it changes the scraper-arm angle linearly (to within 0.5%) with the distance it pulls or pushes itself. Hence, the motion of the scraper is linearly related to the motion of the motor, which is in turn linearly related to the number of steps taken.

5.3.3 Scraper Calibration and Sources of Error

Two limit switches serve to limit the range of the motion, and additionally provide calibration of the range of motion, which is such that the outer edge of the scraper goes from $q_1 = 8.55$cm to 16.9cm (these values are believed good to within $\pm$ 0.1 cm). Since the stepper motor takes 11180 ± 20 steps in going continuously (i.e., without pause) from one limit to the other, the position of the outer edge is seen to be

$$q_1(N_s) = q_1(0) - N_s C_s \tag{5.23}$$

where N_s is the number of steps taken from the inner limit switch (where the scraper is closest to the alpha-magnet front plate), $q_1(0) = 8.55 \pm 0.1$cm, and $C_s = (7.42 \pm 0.13) \times 10^{-4}$ cm/step. Since the stepper is accurate to within 100 steps even for a long sequence of short bursts of steps such as are needed for a spectrum measurement, the uncertainty in $q_1(N_s)$ is primarily due to uncertainty in the knowledge of $q_1(0)$ and C_s. In general, the uncertainty in q_1 is given by

$$\sigma_{q_1}(\text{cm}) = \frac{1}{10}\sqrt{1 + \left(\frac{N_s}{1000}\right)^2}\, 0.13^2 \tag{5.24}$$

If the gradient is g, then the momentum being filtered at N_s steps is

$$p(N_s) = g\left(\frac{q_1(N_s)}{75.05}\right) \tag{5.25}$$

Propagating uncertainties through this equation gives

$$\frac{\sigma_p}{p} = \sqrt{\left(\frac{\sigma_g}{g}\right)^2 + \frac{N_s^2\sigma_{C_s}^2 + \sigma_{q_{10}}^2}{(q_{10} - C_s N_s)^2}}. \tag{5.26}$$

The uncertainty in the gradient, σ_g, is a result of uncertainty in the alpha-magnet current, and is at most $\pm 1\%$. This is as large as it is because the power supply does not have a precision shunt, and I had to rely on the front-panel meter, which has a resolution of 1 A. Since the excitation curve for the magnet was taken using the same meter, there is no concern about the absolute accuracy of the meter. The alpha-magnet is seldom used at gradients below 100 G/cm, which means the current is always greater than 50 A. Gradients of 150-250 G/cm are more typical.

The maximum fractional uncertainty in the momentum will occur either at $N_s = 0$ or $N_s = (N_s)_{max}$, depending on whether the fractional uncertainty in q_{10} or C_s is larger, respectively. Since it is the fractional uncertainty in q_{10} that is larger, one sees that the maximum value of σ_p/p is 1.6%. Rather than propagate errors through the analysis of every spectrum, I will simply take this value as the uncertainty in any momentum measurement. This uncertainty is not the same as the momentum resolution of the momentum filter, which is determined solely by the mono-energetic beam size and the dispersion, as discussed in the last section.

A check of the accuracy of the momentum measurements can be made by taking spectra at several alpha-magnet gradients for the same beam conditions. One expects that if the calibration of the alpha-magnet gradient (vs driving current) and the scraper are correct, then these spectra should be very nearly the same. In fact, the calibrations were checked and then corrected by doing such an experiment. The weak link in the spectrum measurements is the calibration of the alpha-scraper position. As I discussed above, the scraper is on the end of a long arm, which is moved by means of a stepper motor attached to a much shorter arm. Hence, small motions of the short arm produce much larger motions of the scraper, with an amplification

factor of about 4.6. Any accidental movement of the limit switches or bending of the copper cooling tubes that act as the scraper arm would affect q_{10}, leaving C_s unchanged (to a good approximation).

Data taken at several alpha-magnet gradients show very clearly that there is a change in the scraper calibration compared to that done in before installation of the alpha-magnet vacuum chamber. Spectra were taken at fixed beam conditions (i.e., at fixed beam current and RF power) for a series of gradients between 170 and 340 G/cm. The peaks of these spectra were found by analysing the data as discussed above. Figure 5.9 shows the inferred momenta of the spectral peak for the $q_{10} = 8.1$cm (the "nominal" value from the pre-installation calibration) and for $q_{10} = 8.55$cm (the corrected value). The data reduced with the corrected value of q_{10} are clearly much more consistent than those reduced with the nominal value. I found that I could not bring the inferred momenta of the spectral peaks into agreement by varying C_s, nor could I do so by using a hypothetical gradient miscalibration factor (i.e., $g \rightarrow f \cdot g$, with f a constant).

The data points for the highest gradient are apparently spurious, perhaps because the small trajectories produced by such a strong gradient never go outside the region that is strongly influenced by the hole-induced gradient errors (see Chapter 3). Since the spatial region over which the hole-induced errors strongly effect the field is constant as the gradient varies, one would expect that such errors would have the most effect for the strongest gradients.

Note that, in addition to correcting the spread in the inferred momentum peaks, changing q_{10} from its nominal value increases the inferred momentum of the peaks by about 8%. This shift is consistent with other observations, namely, that if the nominal calibration was used to determine the momentum, then the quadrupoles appeared to be about 12% weaker than would have been expected based on simulations; using the corrected calibration of the scraper reduces this discrepancy to about 4%.

As for the source of the calibration error, there is a strong suspicion that the copper cooling tubes that support the scraper were bent during connection to the water supply system: this may be the source of the offset. The offset might also be a result of rough handling during transport and installation: the scraper arm oscillates quite readily, and hence may have been bent as a result of a resonant oscillation

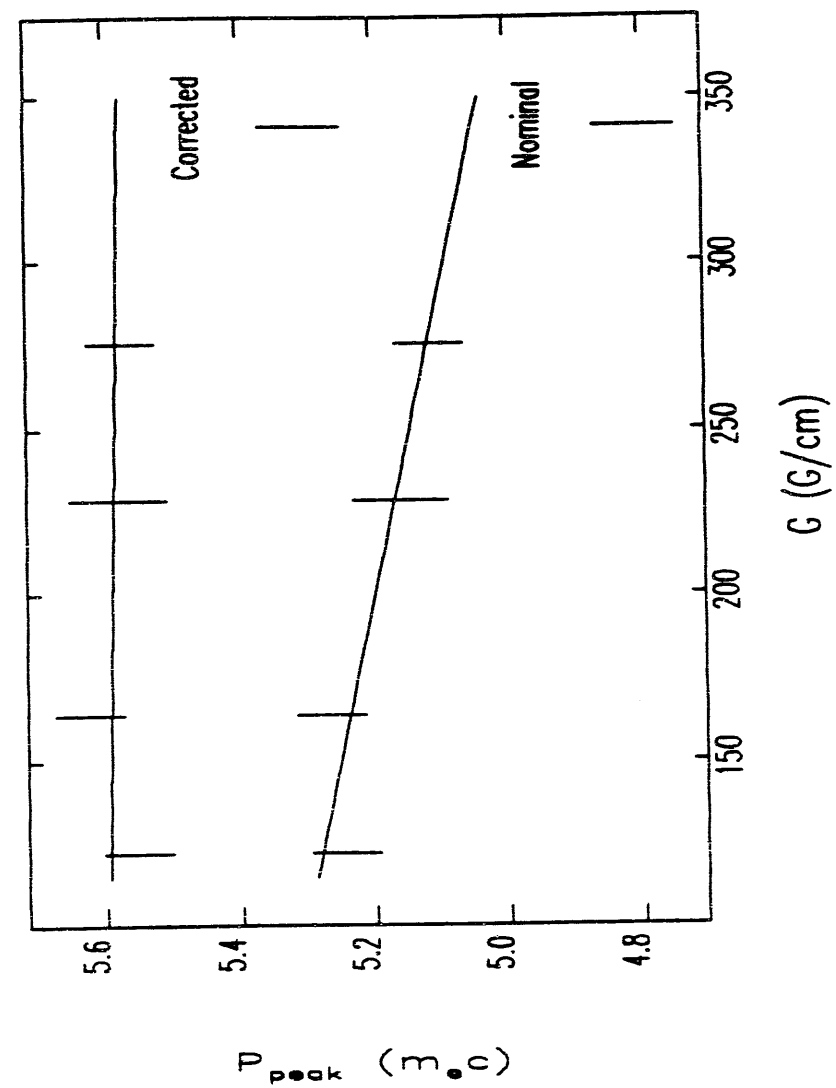

Figure 5.9: Inferred Positions of Spectral Peak for Nominal and Corrected Scraper Calibration

excited during transport.

5.3.4 Measurement of Beam Power

In addition to the momentum distribution of the beam, the data collected by **spect3** can be used to find the beam power transmitted through the alpha-magnet. This is useful in estimating the cavity electric field level, as I will discuss later in this section.

Let I_{tot} be the total current exiting the alpha-magnet, so that the number of electrons in the interval $(p, p + dp)$ passing out of the alpha-magnet during time Δt (where Δt is large compared to the RF period) is

$$dN(p, \Delta t) = \frac{I_{tot}\Delta t}{e}\rho_p(p)dp. \tag{5.27}$$

The kinetic energy of a single electron is given by $m_e c^2(\gamma - 1)$, and so the total kinetic energy of the particles exiting the alpha-magnet during Δt is

$$\Delta U = \int_0^{p_{max}} m_e c^2(\sqrt{p^2 - 1} - 1)\frac{dN(p, \Delta t)}{dp}dp. \tag{5.28}$$

Hence, the transmitted beam power is

$$P_{beam} = \frac{\Delta U}{\Delta t} = I_{tot}\frac{m_e c^2}{e}\int_0^{p_{max}}(\sqrt{p^2 - 1} - 1)\rho_p(p)dp. \tag{5.29}$$

5.3.5 Experimental Results

I performed two series of spectrum measurements designed to test the validity of simulations of the gun and GTL. For the first series, I took data while varying the RF power for constant cathode filament power, at low gun current ($I_{GT1} < 100$mA). For each RF power level, I recorded forward and reflected power waveforms and GT1 and GT2 waveforms. In addition, the program **spect3** was run in order to collect data necessary for determining the spectral distribution and transmitted beam power. For the second series, the RF power was held constant while varying the cathode filament power, and hence the gun current.

Simulation Tests of the Lattice

All relevant magnetic elements (i.e., quadrupoles Q1 through Q3 and the alpha-magnet) were kept at the same settings throughout both sets of measurements. While better spectra would have been obtained for the lower-momentum beams had the lattice momentum (i.e., the central momentum the lattice is designed to transport) been varied to match the beam momentum, this would have greatly increased the time needed for the experiment. It would have been necessary to measure the beam momentum for each RF power level or cathode filament power level, then turn off the RF to allow standardization of the magnets, ramp them to the proper central momentum, turn the RF back on and wait for the gun to equilibriate, and only then take the final scans. This would roughly have tripled the time required for each measurement, and since experimental time was short, I elected instead to use a lattice with a fairly weak alpha-magnet (170 G/cm) and relatively gentle quadrupole settings, and therefore relatively large momentum acceptance.

Simulations with **elegant** verify that this lattice provides reasonably accurate momentum spectra, and, in particular, accurate determination of the position of the spectral peak, for a wide range of beam momenta, for a fixed lattice momentum. **elegant** includes alpha-magnets with moveable inner and outer scrapers in the vertical midplane, which idealizes the experimental situation. For the simulations, I placed the outer scraper at $q_1 = 20$ cm, which is the position of the outer vacuum chamber wall, and directed the program to vary the inner scraper from $q_1 = 8.55$ cm to $q_1 = 17.05$ cm, in 1 mm steps (the experiments used 1mm steps also). The simulations ended at the position of toroid GT2, and included all beam apertures. The output from **elegant** was then post-processed by another program (**alpha_spect**) that is nearly identical to the program (**spect_proc**) used to process experimental data collected with **spect3** (the only significant difference is the data format the programs accept). In this way, I am able to make valid comparisons of simulation and experiment.

Figure 5.10 shows the results of two sets of **elegant** simulations, performed using initial particle distributions generated by **rfgun** for E_{p2} between 50 MV/m and 90 MV/m. (Note that, as always, the areas under the spectral distributions are normalized to unity.) The first set included transverse dynamics to second order, while for

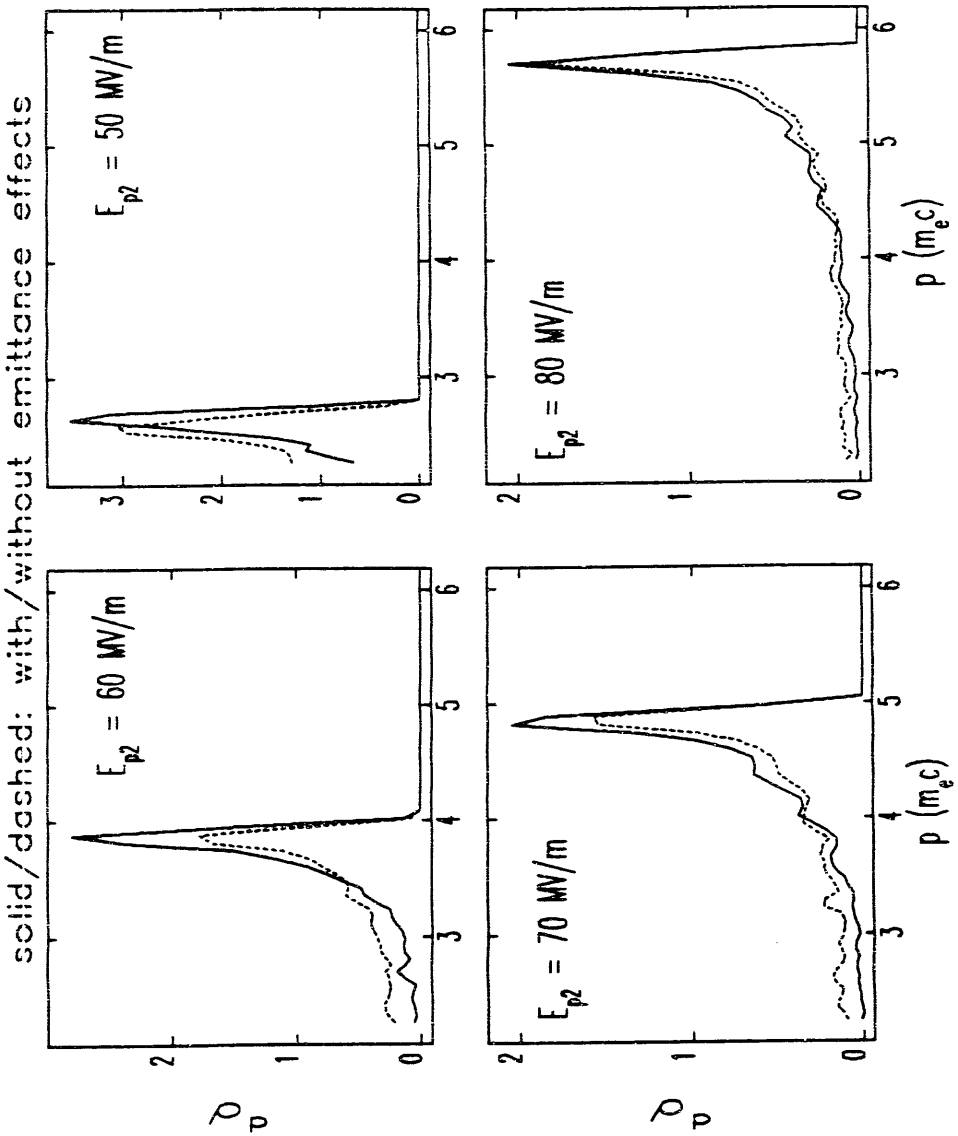

Figure 5.10: elegant-Simulated Spectrum Measurements for Various E_{p2}, Using `rfgun`-Generated Initial Particle Distributions

the second set I directed elegant to set all transverse coordinates to zero (i.e., $\epsilon \rightarrow 0$) before tracking. In both cases, the alpha-magnet was simulated to third order, with the expansion being done about the median momentum of the initial particle distribution. It is not a foregone conclusion that the results for the two cases should be the same. First, for $\epsilon \neq 0$, the momentum resolution of the system must be expected to degrade as the momentum of the highest-momentum particles being intercepted by the scraper becomes significantly less than the central momentum of the beamline (for which the matching for good momentum resolution was done). Second, the simulations with $\epsilon \rightarrow 0$ will not include the effects of apertures in the beamline, which will cause particle losses before and after the alpha-magnet.

One sees from the Figure that for the higher values of E_{p2}, the spectral distributions for the two sets of simulations are quite similar, while for the lower values of E_{p2} there are significant differences. For the $\epsilon \rightarrow 0$ simulations, the spectra include more low-momentum particles, indicating that in the simulations including transverse dynamics, low-momentum particles are being lost between the gun and alpha-magnet (or even after the alpha-magnet but before GT2). This is because the low-momentum particles are over-focused by the quadrupoles (which are optimized for higher momenta), and hence get scraped off on beamline apertures. Similar results are obtained when this comparison is done with MASK-generated initial particle distributions.

The conclusion is that for a valid comparison of experimentally-measured spectral distributions and predictions, the predictions must take the form of a simulation of the actual measurement. That is, one cannot simply compare the measured spectral distribution to the predicted spectral distribution at the gun exit. However, from Figure 5.10 also shows that the *positions* of the spectral peaks are quite close for the two cases, with a definite difference appearing only for $E_{p2} = 50$MV/m. Hence, an experimentally-determined momentum peak at $p \geq 3m_ec$ can be assumed to be at the very close to the position of the spectral distribution of the beam from the gun; this conclusion will help to simplify some of the analysis that follows.

Comparison of Measured and Simulated Spectra

Figure 5.11 shows representative spectral distributions from the first series of measurements. The forward RF power was varied between 2.6 MW and 1.3 MW. Though

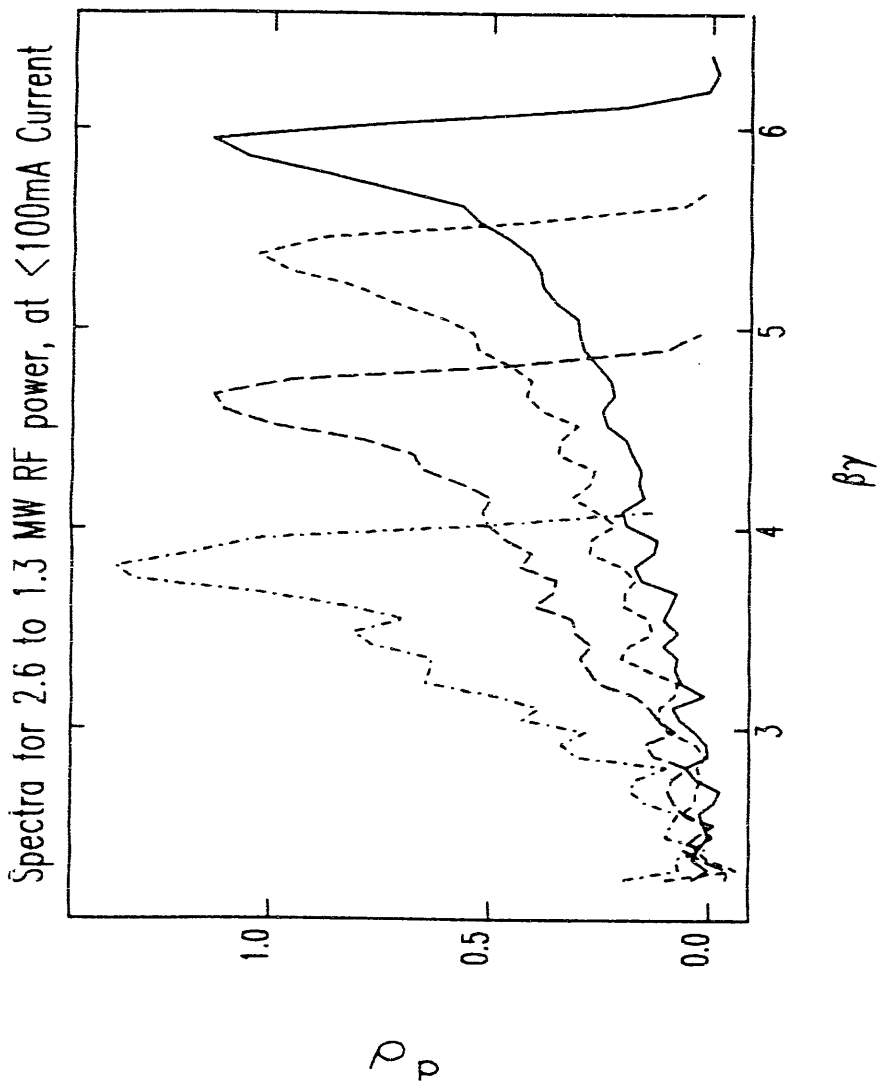

Figure 5.11: Experimentally-Measured Spectral Distributions at Low Current for 1.3 to 2.6 MW Forward Power

the cathode filament power was kept fixed, the beam current varied between 100 mA and 50 mA (respectively) as a natural result of the different cavity fields that prevail at different RF power levels. One sees that these are similar to the simulated spectra shown in Figure 5.10, but noticeably broader.

Better agreement between experiment and simulation is obtained when MASK-generated initial particle distributions are used, as seen from Figure 5.12. The MASK runs used for these simulations were done with very small current densities, so that no space-charge effects are present. Note that the horizontal scale is the same as that used in Figure 5.11. (I did not attempt to match the positions of the peaks between the simulations and the experiments, since this would have required additional MASK runs with little gained.)

In order to further investigate the issue of the broadness of the momentum peaks, I used the results of the `elegant/rfgun` simulations just described (for $\epsilon \neq 0$), along with additional simulations for MASK-generated initial particle distributions. I computed the full-width-at-half-maximum (FWHM) for the spectral peak from every simulation. The results are displayed in Figure 5.13, where the FWHM is plotted against the position of the spectral peak. The results of the same analysis for the experimental data is shown as well.

It is clear from this Figure that `rfgun` seriously underestimates the broadness of the spectrum. This was discussed in Chapter 2, as well, were I pointed out the differences in the spectral distributions predicted by `rfgun` and MASK. The fact that the MASK results correspond more closely to experiment lends support to the conclusion I drew in Chapter 2, namely, that the off-axis expansion used by `rfgun` is of too low an order to accurately model the important non-linear fields near the cathode. Note that the differences persist even for MASK runs done with J = 0 (i.e., without space-charge), implying that the difference is not a result of the fact that `rfgun` does not include space-charge effects.

Figure 5.13 also indicates that, according to the simulations done with MASK-generated distributions, it should be possible to detect current-related (i.e., space-charge induced) broadening of the spectral peak. Indeed, there seems to be some difference between the experimental points for variable RF power (done at < 100mA) and variable filament power (where in excess of 300 mA was used in some cases),

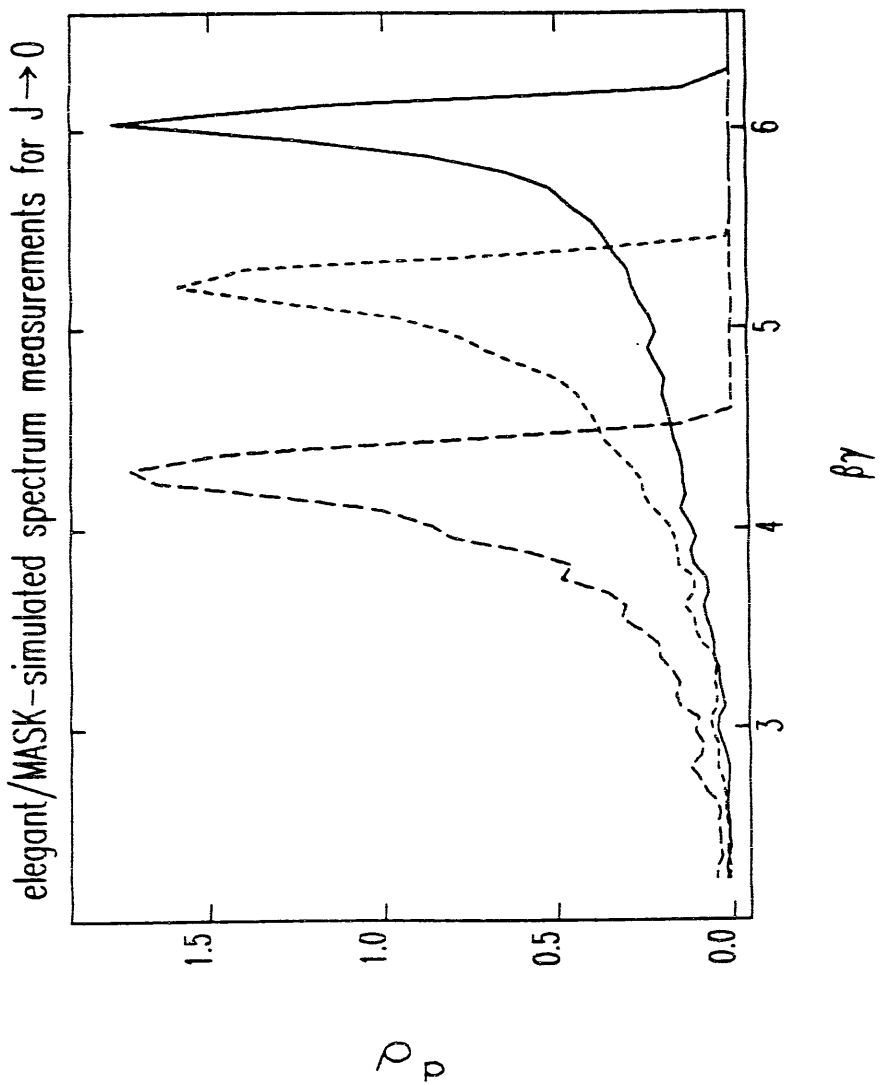

Figure 5.12: **elegant**-Simulated Spectral Distribution Measurements for MASK-Generated J — 0 Initial Particle Distributions

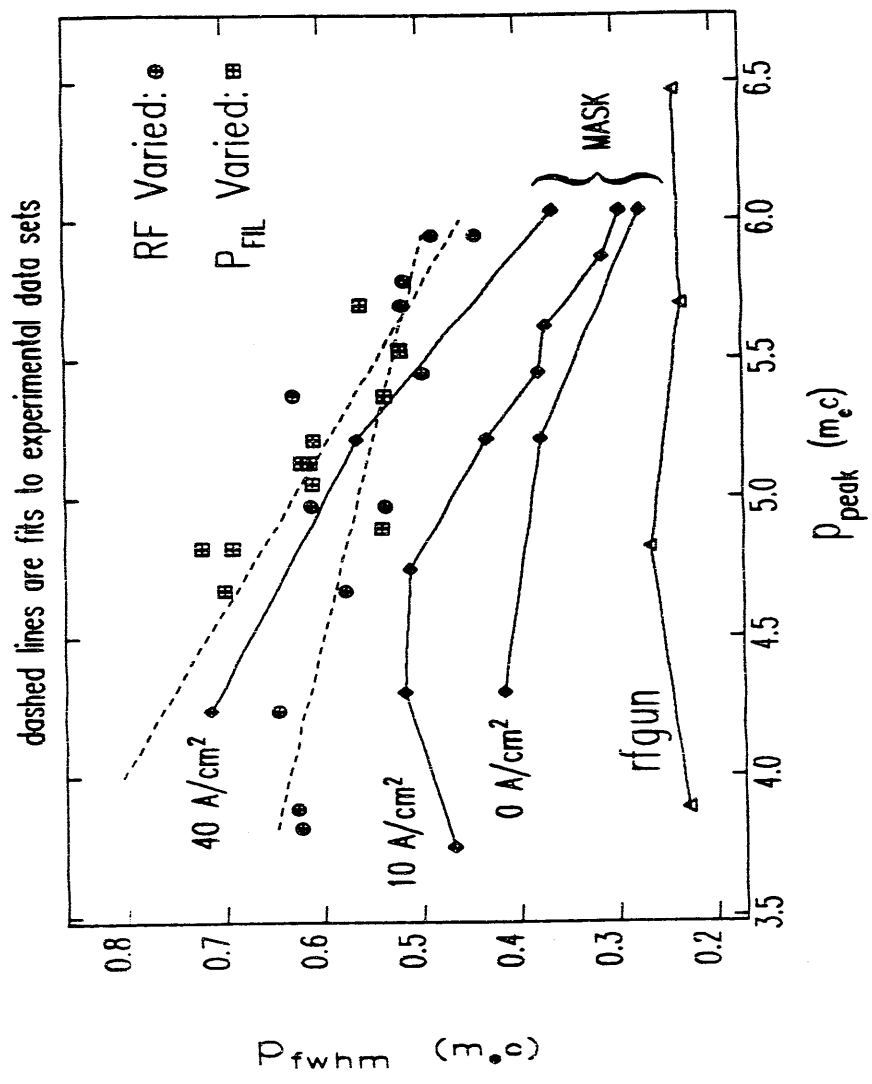

Figure 5.13: FWHM of the Momentum Peak from Experiments and from Simulated Experiments with **elegant**, **rfgun**, and MASK

in that the slope of p_{fwhm} versus p_{peak} is greater in magnitude for the latter set of data than for the former. The comparison of the two series of experimental points is unclear however, in that (unlike the MASK runs) the current density (i.e., the filament power) was varied in one set.

A clearer measurement of the effect of current density on the broadness of the spectrum could have been done by taking two data sets versus RF power at two different filament power levels. If one runs at sufficiently low currents, then the only difference between two such data sets would have been the cathode current density (running at lower currents is required in order to prevent back-bombardment from heating the cathode and thus making the current density vary with RF power). This experiment was, however, not carried out.

Of course, the current density could be inferred for each of the experimental measurements based on the momentum, the current seen at GT1, and simulations of beam-losses in between the gun and GT1. However, this is likely to be very uncertain, for reasons to be seen presently. Still, it is interesting to see how the current density can be estimated.

Inferring the Current Density

Figures 5.14 and 5.15 show additional data from the two series of measurements. While this data is not directly useful for comparison with simulations, it is useful in that it gives an indication of the quality of the measurements, in terms of the reproducibility and smoothness of the dependence of the various quantities on the varied quantity. (I chose I_1 as the independent variable for the second series because it is more physically meaningful than the cathode filament power.)

Returning now to the issue of inferring the current density, recall that in Chapter 2, I used the effective cathode area.

$$A_{eff} = \frac{I_{gun-exit}}{J}. \tag{5.30}$$

to indicate the efficiency with which charge is extracted from the gun. Similarly, I now define an effective cathode area for transmission to GT1:

$$A_{GT1} = \frac{I_{GT1}}{J}. \tag{5.31}$$

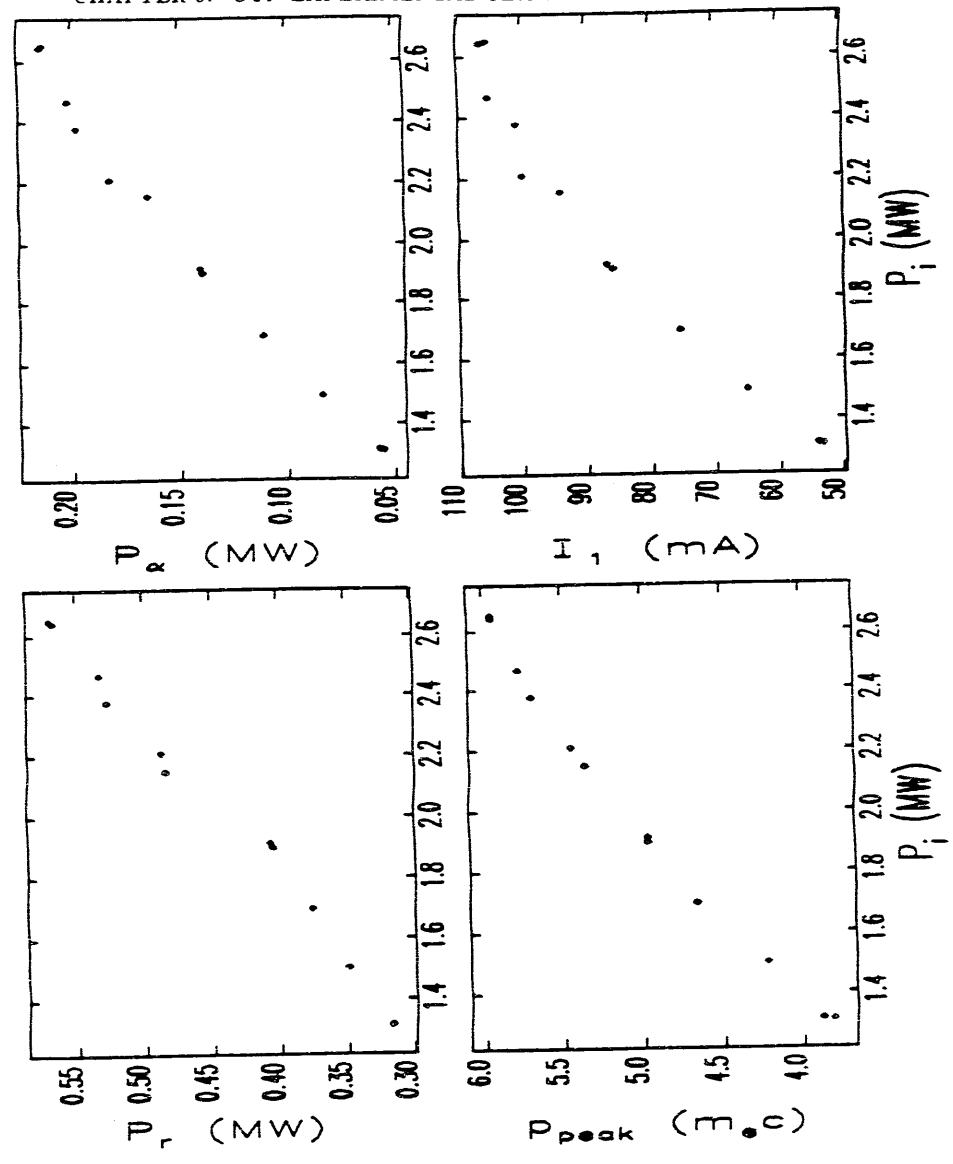

Figure 5.14: Selected Experimentally-Measured Parameters as a Function of Incident RF Power, for Constant Cathode Filament Power

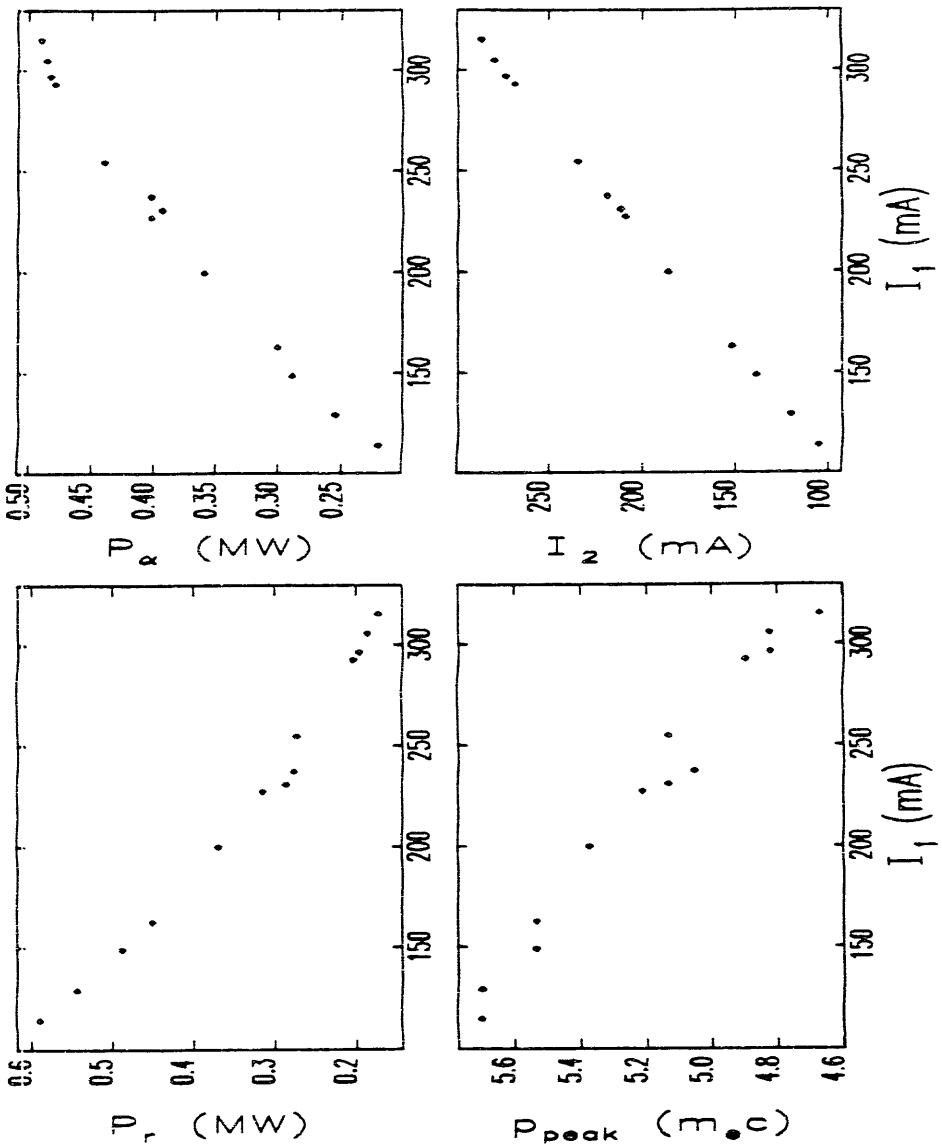

Figure 5.15: Selected Experimentally-Measured Parameters as a Function of GT1 Current, for Constant Incident RF Power

I_{GT1} is a function of the fields in the gun (which determine the beam momentum) and the GTL optics (including the effects of apertures). For a fixed optical configuration, one can parametrize I_{GT1} as a function of p_{peak}, the position of the spectral peak. (I_{GT1} is also a function of J through the effects of space-charge on the emittance and the momentum spectrum, and hence on the transport, but I shall ignore this for simplicity.) Equivalently, one may invert the relationship and calculate J from

$$J = \frac{I_{GT1}(p_{peak})}{A_{GT1}(p_{peak})}. \tag{5.32}$$

The experiments provide measurements of I_{GT1} versus p_{peak}, as shown in the Figures 5.14 and 5.15. $A_{GT1}(p_{peak})$ may be obtained from `elegant` simulations, and I have done this using MASK-generated initial particle distributions calculated for $J = 10 A/cm^2$ (again, I'll ignore any dependence of the transport on J, which is acceptable since, as will be seen, J varies over a small range). The result is shown in Figure 5.16, along with the analogous result for transport to GT2. (I note in passing that these results *do not* indicate the best achievable transmission for the GTL, since the lattice was not matched to each momentum, as discussed above. It is not uncommon to obtain 90% transmission between GT1 and GT2 with a properly tuned lattice.)

Figure 5.17 shows the inferred current density for the experimental data shown in Figures 5.14 and 5.15. One sees that the current densities are apparently quite modest compared to 140 A/cm^2, which is the upper limit for the cathode. As I will discuss presently, there is reason to believe that these results may underestimate the current density by perhaps a factor of two. Some indication that there may be a problem here is the fact that the current density clearly varies for the case of constant filament power, even though the current is low enough to make back-bombardment negligible. (I estimate that there is about 0.5 W of average back-bombardment power for 800 mA to GT1 and 2.6 MW RF, while the heater power is in the range of 9–11 W. Hence, for less than 100 mA, I expect the back-bombardment power to be less than 1% of the filament power.)

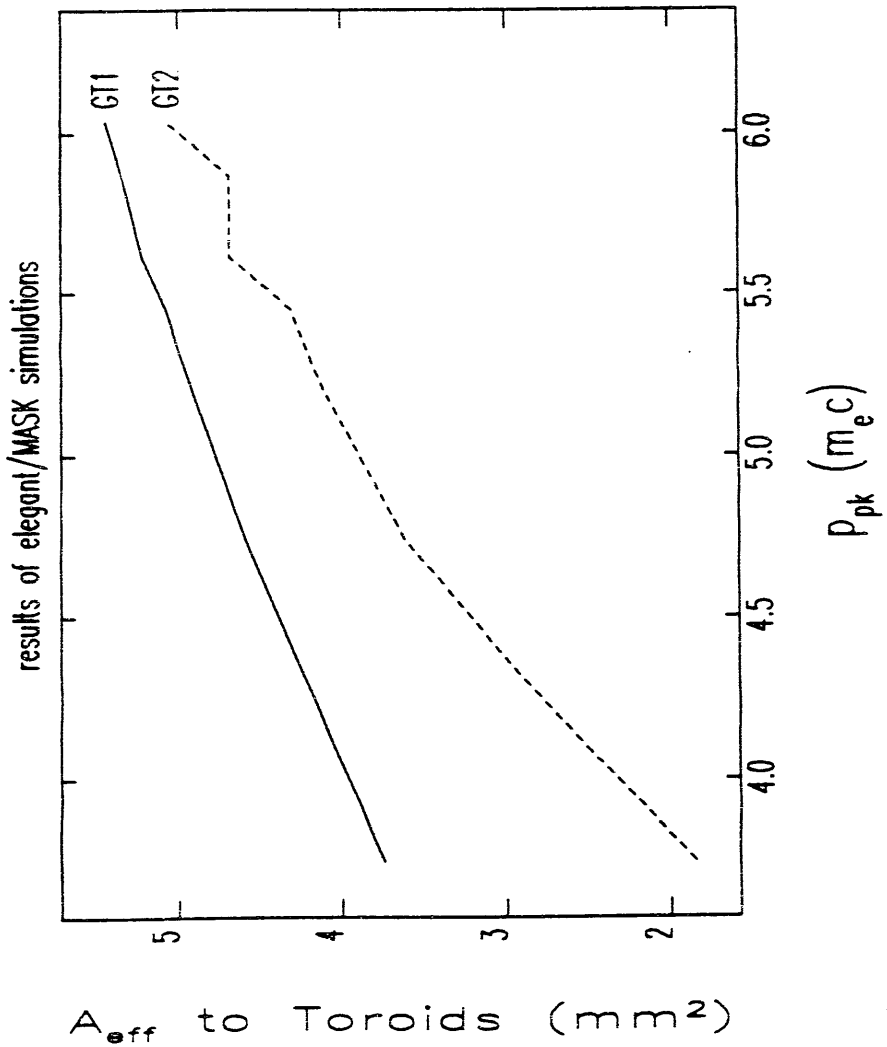

Figure 5.16: Effective Cathode Areas for Transmission of Current to GT1 and GT2, as a Function of the Momentum Peak, from `elegant`/MASK Simulations.

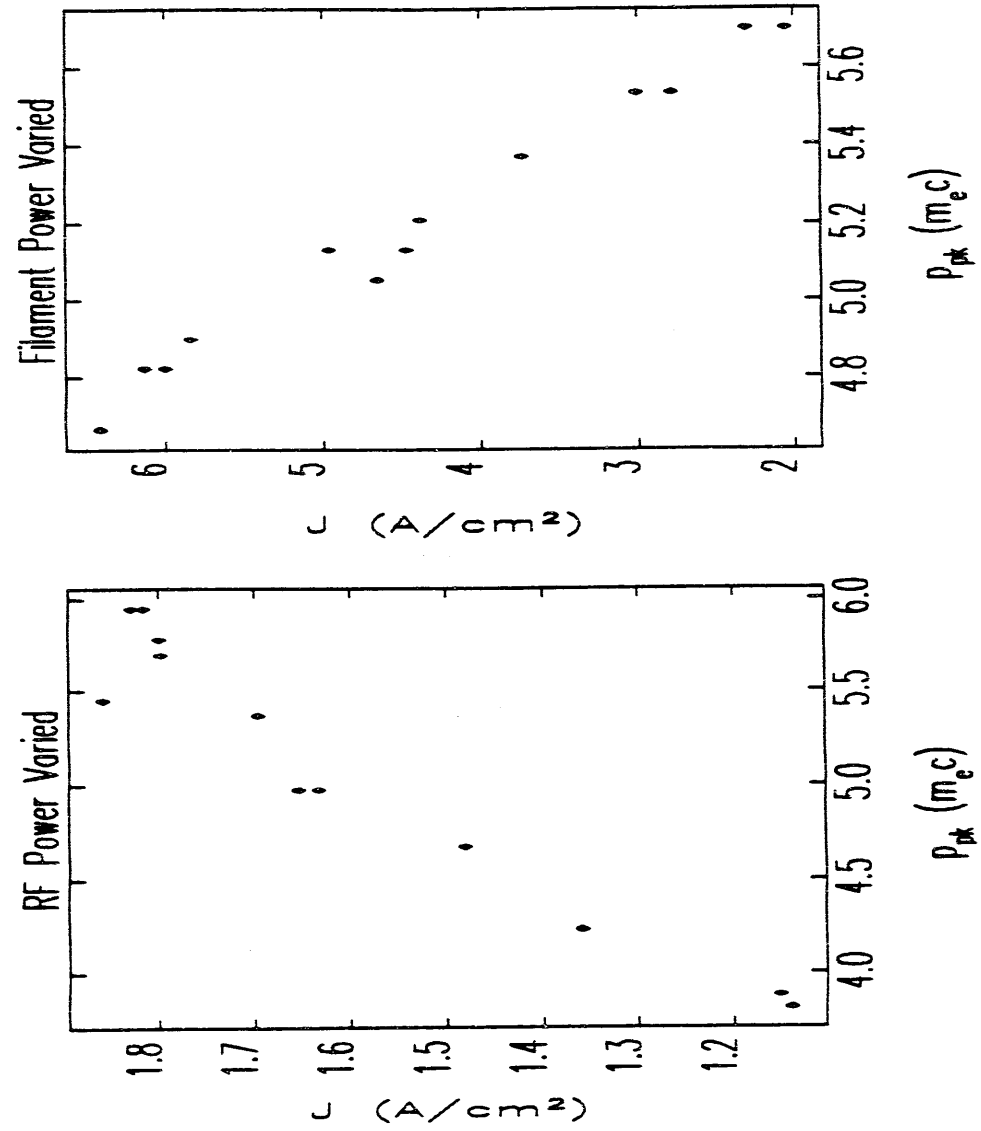

Figure 5.17: Current Density for Experiments, as Inferred from Experimental Data and Simulations

Inferring the Gun Field Level

In Chapter 2, I gave predictions of how the beam momentum is expected to depend on the fields in the cavity, which I parametrized in terms of the peak on-axis field amplitude, E_{p2}. In order to check these predictions against experiment, it is necessary to measure or deduce the E_{p2} levels during the experiments. Since there are no field probes in the gun cavity, I have had to rely on the relationship between E_{p2} and the power dissipated in the cavity walls

$$E_{p2} = K_f \sqrt{P_w}, \tag{5.33}$$

where the nominal value of K_f (obtained from a combination of cold-test measurements and simulation results, as discussed in Chapter 2), is 70.5 MV/m/$\sqrt{\text{MW}}$.

To see how to make use of this, note that conservation of energy dictates that

$$\frac{dU_s}{dt} = P_f - P_r - P_w - P_p, \tag{5.34}$$

where U_s is the energy stored in the cavity fields, P_w is the RF power dissipated in the cavity walls, P_f is the incident RF power, P_r is the reflected RF power, and P_p is the power into particles. P_p includes power into any particles, whether they exit the gun or impact the cathode or the cavity walls. At equilibrium, the cavity fields are constant, and hence so is U_s. Therefore

$$P_w = P_f - P_r - P_p. \tag{5.35}$$

Experimentally, P_f and P_r are obtained from calibrated RF detector diodes, which produce a voltage signal that can be related to the applied RF power.

P_p must be obtained by a more circuitous route, since there is no way to measure the power that goes into particles hitting the cathode, the cavity walls, or apertures between the gun and the alpha-magnet. What can be measured, as discussed above, is the beam power transmitted through the alpha-magnet. To relate this to P_p, I have used simulations to model the particle losses in the gun and between the gun and alpha-magnet. For the latter simulations, I used `rfgun`-generated particle distributions as input to `elegant`, which tracked the particles through the GTL lattice using the beam-optics used in the actual experiment (with the quadrupole model

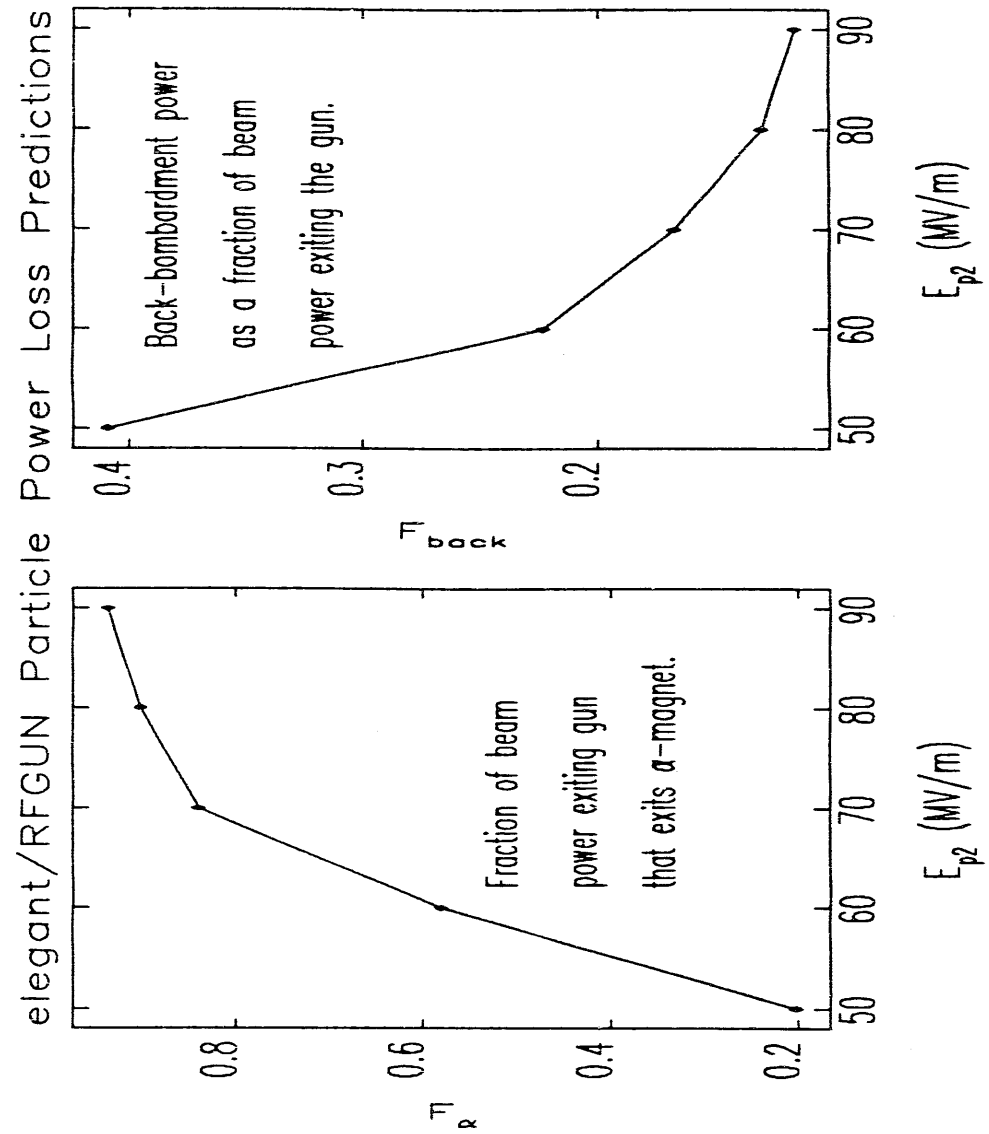

Figure 5.18: Particle Power Loss Parameters as a Function of the Spectral Peak

adjusted as per the previous section). I did such simulations for E_{p2} between 50 and 90 MV/m. (Essentially the same results are obtained using MASK-generated initial particle distributions for J=10 A/cm^2.) These simulations allowed me to deduce the ratio of initial beam power (i.e., the power in the particles that exit the gun) to transmitted beam power as a function of E_{p2}, or, equivalently and more usefully, as a function of p_{peak}. Figure 5.18 shows the results of these calculations, which I will represent formally as

$$P_{ib} = \frac{P_o}{F_o(p_{peak})}. \tag{5.36}$$

where P_{ib} is the initial beam power, P_o is the beam power transmitted through the alpha-magnet, and F_o is the fraction of the initial beam power that exits the alpha magnet.

Next, I used `rfgun` to simulate back-bombardment as a function of E_{p2}, allowing me to deduce the ratio of back-bombardment power to initial beam power as a function of the spectral peak:

$$P_{back} = P_{ib}F_{back}(p_{peak}). \tag{5.37}$$

These results are shown in Figure 5.18. Simulations with MASK and `rfgun` indicate that it is a good approximation to equate the beam power lost in the gun with the beam power lost by particles hitting the boundary at $z = 0$ (i.e., the plane of the cathode), and I will employ this approximation, so that

$$P_p = P_{ib} - P_{back}, \tag{5.38}$$

or

$$P_p = \frac{P_o}{F_o(p_{peak})}(1 - F_{back}(p_{peak})). \tag{5.39}$$

The quantities P_f, P_r, P_o, and p_{peak} can all be measured, and thus it is now possible to calculate the cavity wall power for the experiments, using

$$P_w = P_f - P_r - \frac{P_o}{F_{meas}(p_{peak})}, \tag{5.40}$$

where

$$F_{meas}(p_{peak}) = \frac{F_o(p_{peak})}{1 - F_{back}(p_{peak})} \tag{5.41}$$

is the fraction of the power going into particles that is actually measured (i.e., at the exit of the alpha-magnet). Equation (5.40) can then be used in equation (5.33) to find the peak, on-axis electric field for each RF power level used in the experiment.

Figure 5.19 shows the results of this analysis for the two sets of data. From the data at constant filament power, one might conclude that there was a calibration error in the RF power measurements or that K_f was in error. However, the data for variable filament power, where beam-loading is significant (and is indeed the cause of the variation of p_{peak}), show that the error is most likely to be in the calculation of F_{meas}. Figure 5.20 shows the results of repeating the analysis with $F_{meas} = F_{meas}/1.47$ and $K_f = 66$MV/m, a combination that was found to bring both sets of data into reasonable agreement with simulation results. I found that I could not obtain a similar level of agreement for both data sets by modifying only K_f or by modifying only F_{meas} by a multiplicative constant. Hence, it would appear that the beam losses are considerably greater than predicted, and that the Q of the cavity (which is related to K_f) is lower than predicted.

Possible Explanations of Discrepencies

Several explanations of the higher-than-expected beam power losses are possible:

1. If the emittance is larger than predicted, then more particles would be lost on the apertures between the gun and alpha-magnet. In particular, there is a 2mm radius aperture constriction at the exit of the gun, designed to isolate the gun vacuum from the GTL vacuum. This is observed to be a significant source of radiation[81], suggesting that, contrary to expectations, significant beam power is lost here. Simulations predict that only about 10% of the beam power should be intercepted by this aperture.

2. Space-charge effects in the gun or GTL might alter the momentum spectrum sufficiently to falsify the beam-power-loss predictions. I do not believe this is the explanation, as the effect of space-charge on the position of the momentum peak is slight.

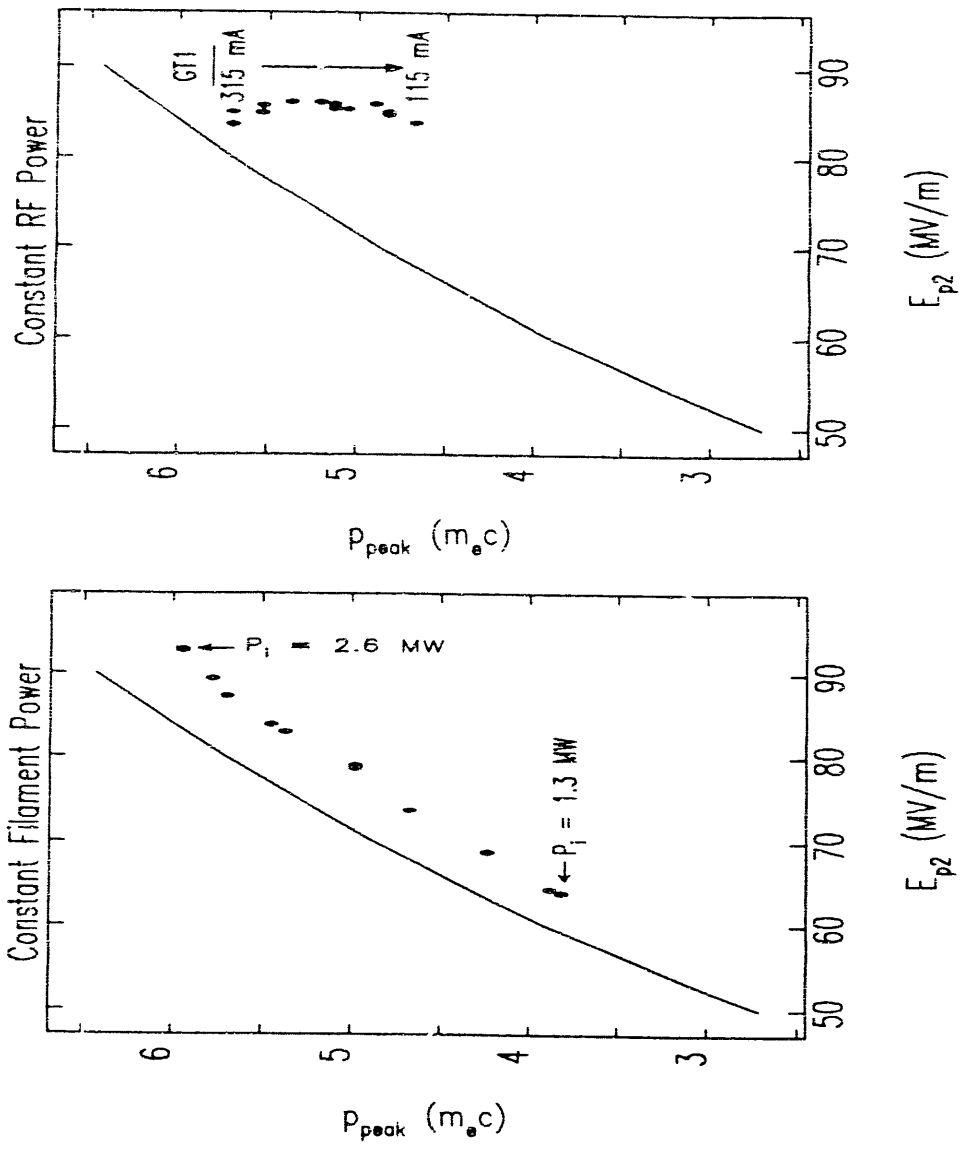

Figure 5.19: Spectral Peak as a Function of Inferred E_{p2}, for Constant Filament Power and Constant RF Power Series

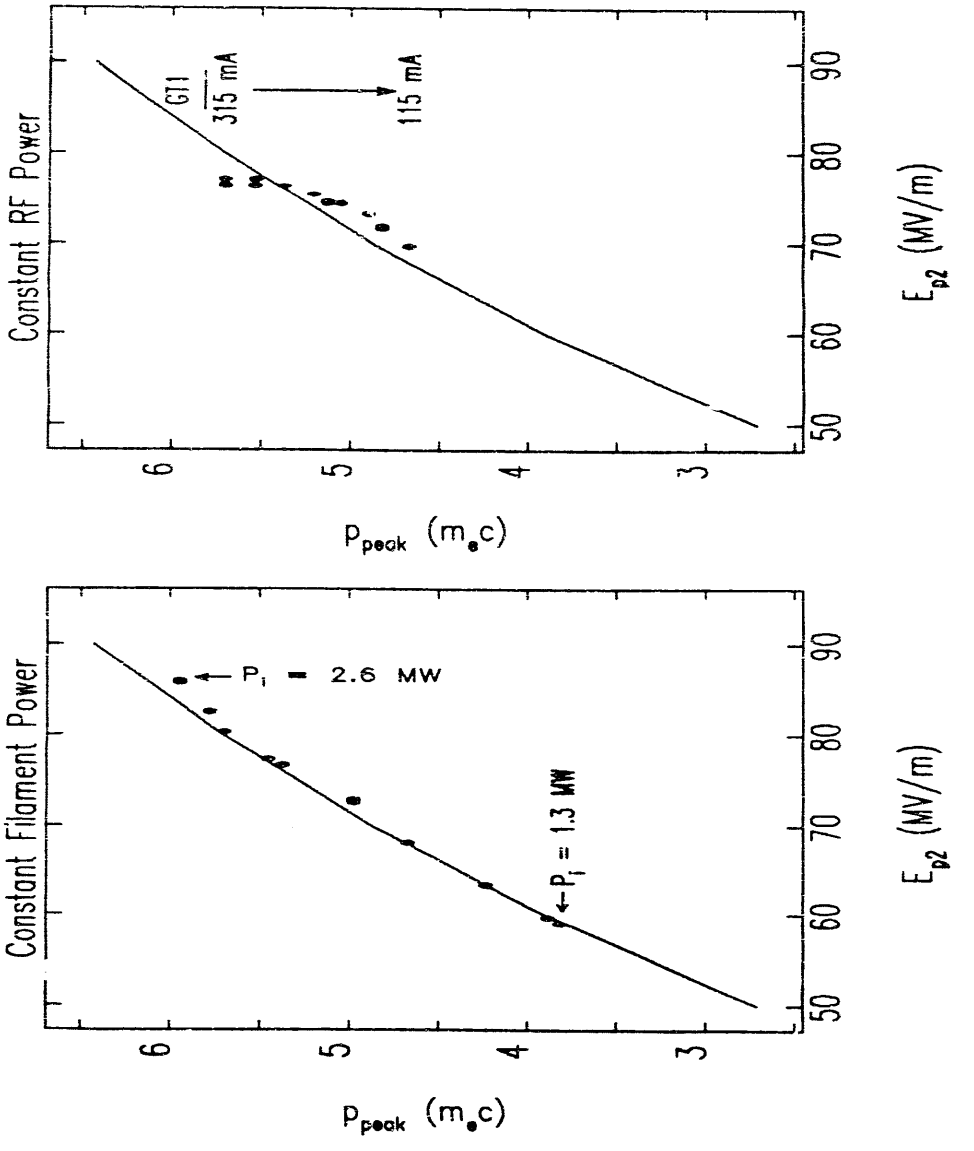

Figure 5.20: Spectral Peak as a Function of Corrected E_{p2}, for Constant Filament Power and Constant RF Power Series

3. Dark-current electrons or secondary electrons created by electrons hitting the cavity walls could be absorbing RF power, without contributing to the beam [82]. There is evidence that secondaries (or dark-current electrons) are present in significant quantities, in that shape of the initial part of the reflected power pulse is seen to be very unstable when there is no beam. This suggests that electrons are being emitted sporadically in sufficient quantities to load the cavity, though no observable beam reaches GT1. This sporadic emission appears to stabilize (i.e., become continuous) when the cathode is heated by the filament, suggesting that the addition of a primary beam brings about continuous emission of these useless electrons.

4. The estimation of K_f does not include losses in the side-coupling cavity, though clearly there must be some in actual operation if net power is to flow from the second on-axis cell to the first. This issue remains to be investigated.

5. Miscalibration of the reflected power diode could cause over-estimation of the amount of "left-over" power, and hence over-estimation of the amount of power that must have gone into the beam. I have made several tests of the calibrations of the forward and reflect power diodes:

   (a) The combined diodes, loads, attenuators, and cabling for the two RF signals are identical to within 5%.

   (b) Simultaneously sampled forward and reflected power waveforms as a function of the RF drive show that the dependence of the reflected power on the forward power is fairly linear, except at low power, where the secondary problem just mentioned has a serious effect on the reflected power and also where the forward RF power waveform began to change shape (a property of the particular klystron being used). Figure 5.21 shows the pulse-average forward power and various parameters of the reflected power pulse (pulse average, power level at the second peak, power level at equilibrium), as a function of the average forward power. Also shown is the peak forward power as a function of the pulse-average forward power, which provides a test of the accuracy of the experiment and the effect of distortion of the

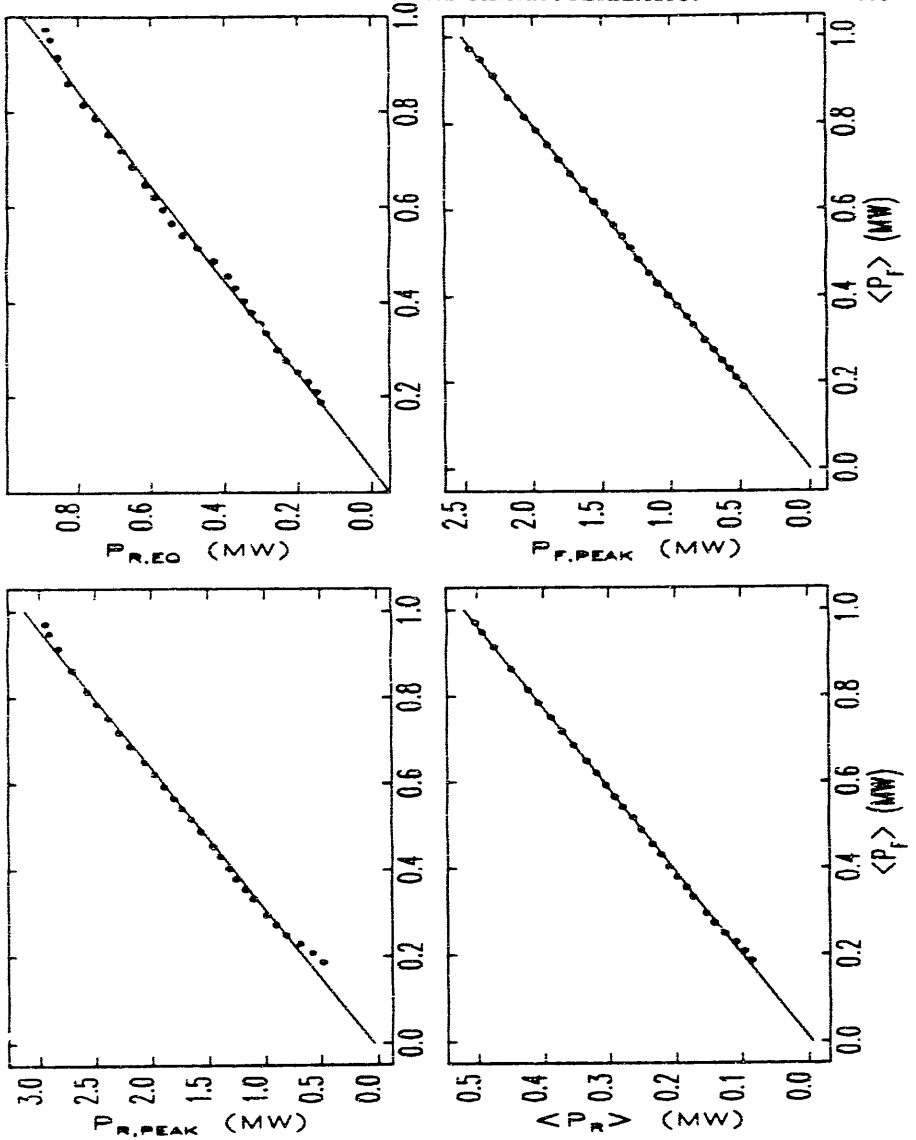

Figure 5.21: RF Calibration Test Results (See Text for Explanations)

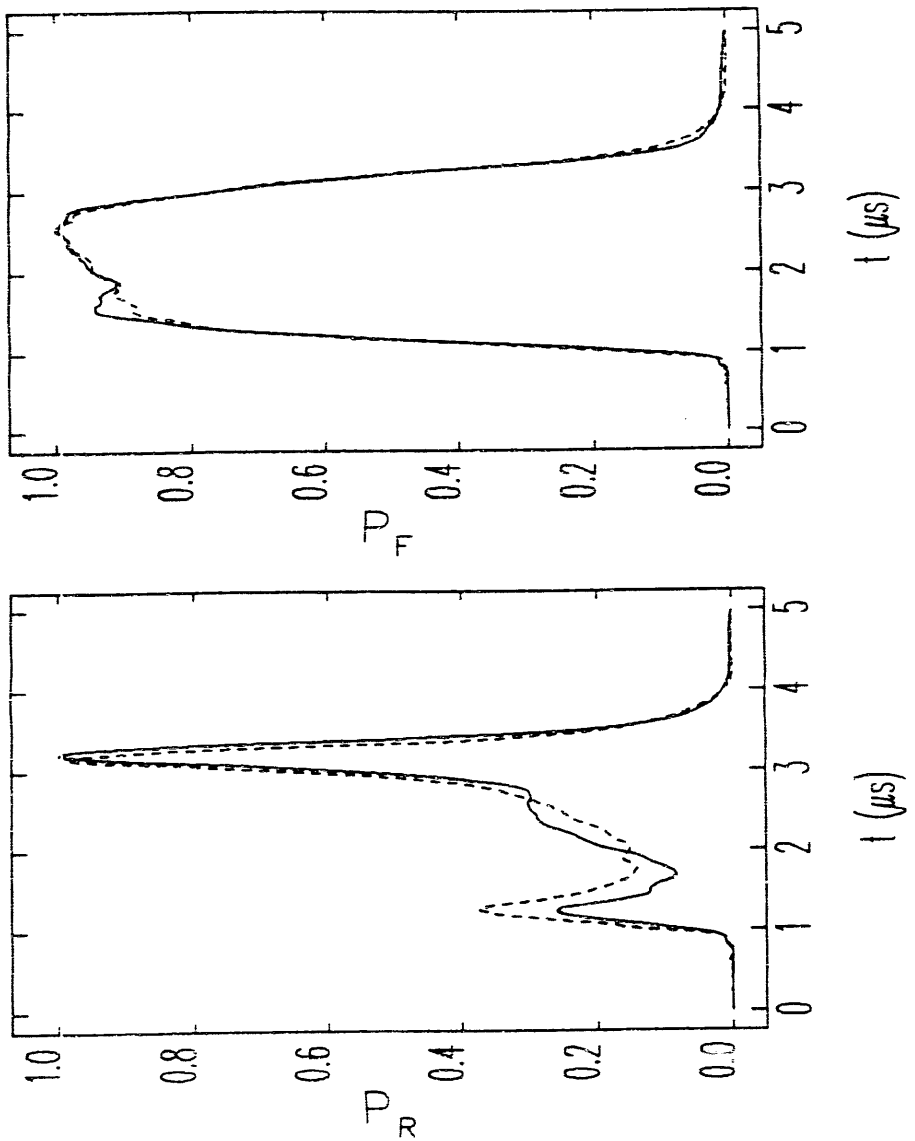

Figure 5.22: Normalized Forward and Reflected RF Power Waveforms for High and Low Forward Power

forward power pulse. For perfectly linearized detectors (and ignoring the other effects just mentioned), the data points should be scattered about the fit lines. The average reflected power appears quite linear in the average forward power, as does the peak forward power. The equilibrium reflected power appears to deviate at higher power levels, and to show some non-statistical wobble about the best-fit line: this is not surprising, since the shape of the reflected power pulse changes considerably over the range of forward power used. Normalized forward and reflected power waveforms for the highest and lowest power waveforms are shown in Figure 5.22.

One sees that the "emitted" power (at the second peak in the reflected power pulse) seems to deviate from linearity seriously in the lower ranges. This could explain the some of the discrepancies seen above, in that the value of P_r used in equation (5.35) is quite small when the gun is heavily beam loaded, as was the case in the second of the two series of measurements. I have not attempted to correct for this effect, as recalibration of the forward and reflected power detection circuitry, starting with the directional couplers in the waveguide, would have required partial disassembly of the waveguide leading to the gun. It is clear that this is not the only problem, since it would not explain the discrepancies for the first series of measurements.

5.4 Emittance

The emittance, like the bunch length, is one of the most important parameters of the gun beam, for it helps to determine the brightness of the beam and hence its suitability for applications, such as FELs, requiring high brightness. In this section, I discuss the measurements I did of the emittance of the RF gun beam, and compare these with expectations based on simulations. Ideally, one would like to measure the emittance as a function of beam current and momentum over the entire range of gun operation. Unfortunately, this did not prove possible, and emittance data is only available for a rather limited range of gun operating conditions. The reasons for this will become clear as I proceed. It is necessary first to discuss the principle of the measurements.

A number of methods are used for measuring emittance; a review of some of the commonly-used methods appears in one of the references [1]. The method I employed made use of existing hardware, and involved the variation of a quadrupole upstream of a beam-profile measuring device, in this case a phosphorescent screen. By measuring the beam-size at the screen as a function of the quadrupole strength, one can deduce the RMS geometric emittance, defined by equations (5.42) and (5.43) below. (Note that when I use the word "emittance" in this section, I mean the RMS *geometric* emittance, unless otherwise stated.)

5.4.1 Principle of the Emittance Measurements

The theory of the method as it is usually developed (see, for example, [83]) assumes a mono-energetic beam described by the so-called Σ-matrix[84], which for the x-plane is

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} x_{rms}^2 & \langle xx' \rangle \\ \langle x'x \rangle & x_{rms}'^2 \end{pmatrix}. \tag{5.42}$$

(Often, the beam is assumed to be gaussian, and hence completely characterized by the Σ-matrix, but this assumption is unnecessary, as discussed in subsection 5.4.6.) The RMS geometric emittance in terms of the Σ-matrix is simply

$$\epsilon = \sqrt{\det \Sigma} = \sqrt{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}. \tag{5.43}$$

where I have used the fact that $\Sigma_{12} = \Sigma_{21}$.

An arbitrary Σ-matrix, Σ_0, is transformed by a beamline with a transport matrix $\mathbf{r}$ according to:

$$\Sigma_1 = \mathbf{r}\Sigma_0\mathbf{r}^T, \tag{5.44}$$

where Σ_1 is the Σ-matrix at the end of the beamline and where $\mathbf{r}^T$ is the transpose of $\mathbf{r}$. The spatial beam sigma at the end of the beamline—e.g., at a phosphorescent screen—is given by

$$\sigma = \sqrt{(\Sigma_1)_{11}}. \tag{5.45}$$

Clearly, σ is a function of the original Σ-matrix, Σ_0, and the matrix $\mathbf{r}$, that is, of the initial beam phase-space and the properties of the beam-transport system between points 0 and 1.

In the experiment, one varies a beamline element—e.g., the strength of a quadrupole—thus obtaining a series of n matrices $\mathbf{r}_i$. Corresponding to each of these matrices is a spatial sigma, σ_i, at the end of the beamline. By explicitly writing out equations (5.44) and (5.45), one can express the results of the whole series as

$$\begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_n^2 \end{pmatrix} = \begin{pmatrix} r_{1,11}^2 & 2\cdot r_{1,11}\cdot r_{1,12} & r_{1,12}^2 \\ r_{2,11}^2 & 2\cdot r_{2,11}\cdot r_{2,12} & r_{2,12}^2 \\ \vdots & \vdots & \vdots \\ r_{n,11}^2 & 2\cdot r_{n,11}\cdot r_{n,12} & r_{n,12}^2 \end{pmatrix} \times \begin{pmatrix} \Sigma_{11} \\ \Sigma_{12} \\ \Sigma_{22} \end{pmatrix} \tag{5.46}$$

or

$$\mathbf{M} = \mathbf{T} \times \mathbf{S}. \tag{5.47}$$

From this, one obtains a solution for the elements of the Σ-matrix:

$$\mathbf{S} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{M} \tag{5.48}$$

from which the emittance is obtained by (see equation (5.43))

$$\epsilon = \sqrt{S_1 S_3 - S_2^2} \tag{5.49}$$

Note that this procedure gives not only the emittance, but also the beam-correlations.

5.4.2 Inclusion of Experimental Errors

For experimental data, the σ_i are known within some uncertainty $\Delta\sigma_i$. In this case, one can use a weighted least-squares fit [55, 61] instead of the equal-weights fit given by equation (5.48), obtaining

$$\mathbf{S} = (\mathbf{T}^T\mathbf{C}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{C}^{-1}\mathbf{M}. \tag{5.50}$$

where $\mathbf{C}$ is the covariance matrix of the experimentally measured quantities, defined by

$$\mathbf{C}_{ij} = \delta_{ij}\Delta\sigma_i\Delta\sigma_j. \tag{5.51}$$

with δ_{ij} being the Kronecker delta-function, and where I use Δ to indicate the uncertainties to avoid confusing multiple uses of the letter σ. The uncertainties in the elements of $\mathbf{S}$ (i.e., in the Σ-matrix derived from the fit) are found by computing the covariance matrix of the fit parameters:

$$\mathbf{K} = \mathbf{U}^T\mathbf{C}\mathbf{U}, \tag{5.52}$$

where $\mathbf{U}$ is defined as

$$\mathbf{U} = (\mathbf{T}^T\mathbf{C}^{-1}\mathbf{T})^{-1}\mathbf{T}^T\mathbf{C}^{-1}. \tag{5.53}$$

The uncertainties in the elements of $\mathbf{S}$ are related to the diagonal elements of $\mathbf{K}$, by

$$\Delta S_i = \sqrt{K_{ii}} \tag{5.54}$$

The uncertainty in the emittance is then given by propagation of errors through equation (5.49), which, when expressed in terms of the $\Sigma-$matrix yields

$$\begin{aligned}\Delta\epsilon \;=\; & \left(\frac{\Sigma_{11}}{2\epsilon}\right)^2 K_{33} + \left(\frac{\Sigma_{22}}{2\epsilon}\right)^2 K_{11} + \left(\frac{\Sigma_{12}}{2\epsilon}\right)^2 K_{22} - \\ & 2\frac{\Sigma_{11}\Sigma_{22}}{4\epsilon^2}K_{13} + 2\frac{\Sigma_{11}\Sigma_{12}}{2\epsilon^2}K_{12} + 2\frac{\Sigma_{12}\Sigma_{22}}{2\epsilon^2}K_{12}.\end{aligned} \tag{5.55}$$

Note that I have used the off-diagonal terms of the covariance matrix in an attempt to obtain better accuracy. As I will discuss below, this sort of error analysis gives dubious results and I have used a Monte-Carlo method [61] instead for my final results.

5.4.3 Thin Lens Treatment

It is instructive to write equation (5.48) out for the case of a thin lens quadrupole followed by a drift space of length L. In this case, the transport matrix is given by

$$\mathbf{r} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\kappa & 1 \end{pmatrix}, \tag{5.56}$$

where $1/\kappa$ is the focal length of the quadrupole. The beam-size at the end of the drift space as a function of the quadrupole focal-length is (from one line of equation (5.46))

$$\sigma^2 = (1-\kappa L)^2\Sigma_{1,11} + 2L(1-\kappa L)\Sigma_{1,12} + L^2\Sigma_{1,22}. \tag{5.57}$$

As pointed out by Ross, *et. al.*,[86], this can be written in a more transparent form as

$$\sigma^2 = \sigma_o^2L^2(\kappa-\kappa_m)^2 + \frac{L^2\epsilon^2}{\sigma_o^2}, \tag{5.58}$$

where σ_o is the beam-size at the entrance to the quadrupole and where

$$\kappa_m = \frac{\Sigma_{1,12}L + \Sigma_{1,11}}{\Sigma_{1,11}L} \tag{5.59}$$

is the value of κ that gives the minimum spot-size at the screen.

Since σ_o is presumably unknown, one must find both the second derivative and the minimum of σ^2 as a function of κ in order to find the emittance. If one rewrites equation (5.58) as

$$\sigma^2 = D^2(\kappa-\kappa_m)^2 + \sigma_m^2, \tag{5.60}$$

where

$$D = \sqrt{\frac{1}{2}\left(\frac{\partial^2}{\partial\kappa^2}\sigma^2\right)} \tag{5.61}$$

then the emittance is given by

$$\epsilon = \frac{\sigma_m D}{L^2}. \tag{5.62}$$

One can appreciate from this equation that in order to determine the emittance it is important to probe the minimum of the beam-size versus κ, in order to get σ_m, and that it is also important to vary κ over a sufficiently wide range to get a reliable value for D. For noiseless data, the emittance can be reliably extracted even without

probing the minimum. This shouldn't be surprising, since for ideal data one only needs three data points for any three different values of quadrupole strength in order to completely constrain the fit. For real (i.e., noisy) data, however, it is important to both probe the minimum beam-size and the "wings" where the beam-size becomes quite large.

5.4.4 Emittance Measurement Lattice and Procedure

The design of measurement systems to minimize the error in the emittance determination is discussed in the literature [86, 11, 87]. In order to minimize the effect of the finite resolution of the imaging system, one wants a large distance between the quadrupole being varied and a small beam-size at the quadrupole [86]. For a fixed emittance, this yields a larger value of σ_m relative to the fixed resolution of the imaging system. However, since the placement of the quadrupoles and phosphor screen was dictated by the requirements of the GTL as an operational—as opposed to experimental—system, these considerations are largely academic for the present work.

The emittance measurements for the RF gun made use of quadrupoles Q4 and Q5 and the chopper-screen. The alpha-magnet scraper was used to determine the momentum spectrum of the beam and to filter the momentum of the beam allowed to reach the screen. In this way, measurements could be made as a function of momentum spread, for comparison with the simulation results of Chapter 2.

The optics up to the alpha-magnet was designed so as to produce a reasonably small beam spot on the chopper-screen, even with Q4 and Q5 off. Q1, Q2, and Q3 were set at 98.9 m^{-2}, -90.4 m^{-2}, and 63.1 m^{-2}, respectively. These are considerably weaker quadrupole strengths than those for the running lattices listed in the earlier section of this chapter. The alpha-magnet strength was set so as to produce $\hat{q}_i$ = 14.431cm.

For the horizontal emittance measurements, Q4 was varied while keeping Q5 off. Following this, Q4 was degaussed and set to (nominally) $K_1 = 90m^{-2}$, and Q5 was varied. Ideally, the data from the variation of Q4 would have been used to obtain ε_x while that from the variation of Q5 would have been used to obtain ε_y. Unfortunately, saturation problems invalidated the Q5 data, as will be seen below.

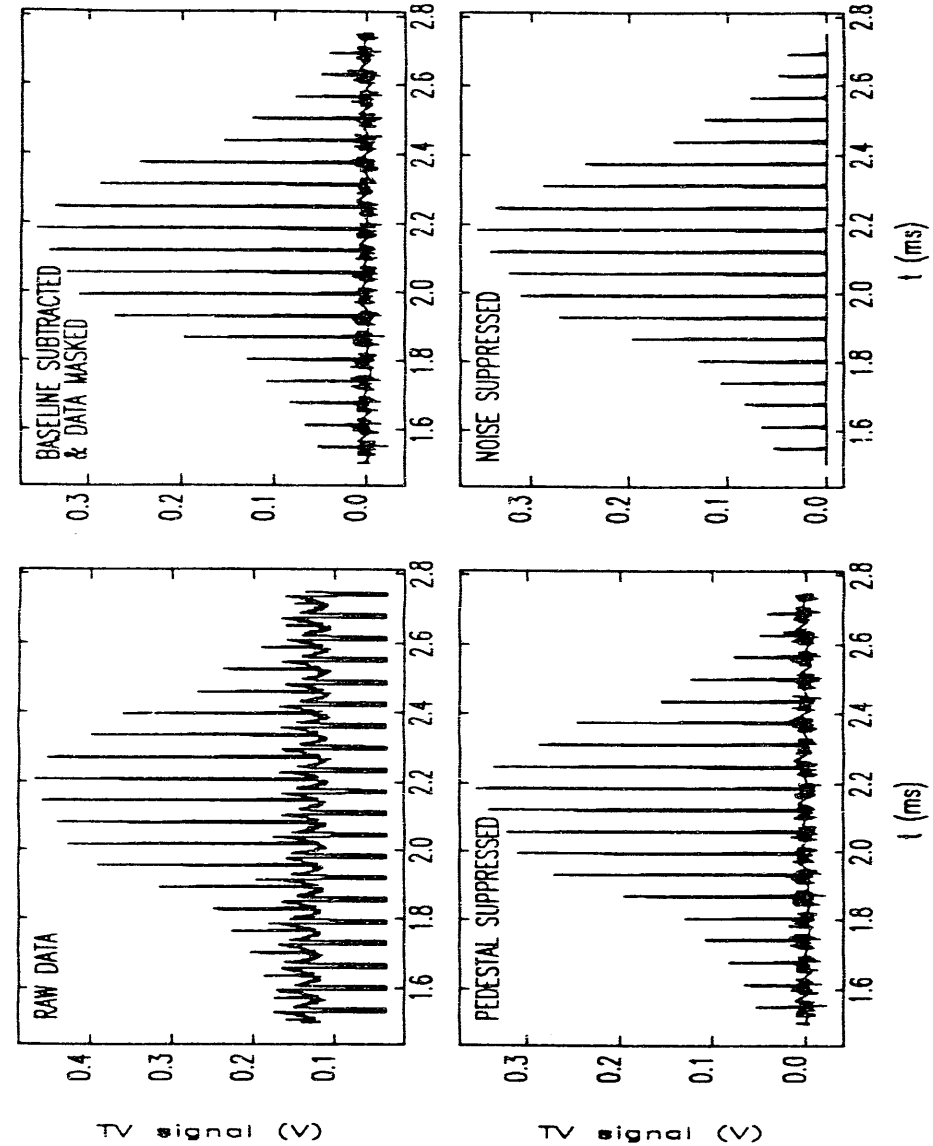

Figure 5.23: Signal Analysis for a Video Scan

For each quadrupole setting, the LeCroy 9450 oscilloscope was used to digitize a full screen sweep from the camera–i.e., a full two-dimensional intensity map of the beam—and the resultant waveform was transferred to the PC and stored there. Because of the large number of data points ($\sim$50K) in each waveform, on-line processing of the data was neither possible nor would it have been efficient. Rather, data was collected and stored for later analysis on SSRL's VAX 8810. Figure 5.23 shows a section of a raw video scan taken in the course of one of the emittance measurements. The envelope of this signal is essentially the vertical beam profile, while the "sub-scans" (separated by regions where the signal falls to < 0.05 V) are horizontal slices through the beam at successive vertical positions. This data has a clear baseline variation that largely repeats through successive horizontal scans.

5.4.5 Analysis of Digitized Beam-Spot Images

To analyse the large amount of data that came from even a single emittance measurement, it was desirable to have a computer program that would process the raw video scans and reconstruct the two-dimensional beam-profile. I have employed several stages in the processing of such scans, as illustrated in Figure 5.23, and as implemented in the program **hvscan**:

1. Organization—The beginning and end of each horizontal sub-scan was located, and the data organized into a two-dimensional intensity map (or "pixel map").

2. Baseline subtraction—The first horizontal sub-scan of every complete scan was used as the baseline for all subsequent horizontal sub-scans. (Even with no beam impinging on the screen, the signal from the camera was seen to vary over each horizontal sweep, rather than becoming the flat signal that one would expect in an ideal circumstance.) This involves assuming that this sub-scan is free of any beam-related signal, which is invariably true for the data I have taken. Since the height of the baselines for the sub-scans were seen to be vary slowly from sub-scan to sub-scan, the baseline scan was scaled by a smoothly-varying function of the scan number before subtraction.

In hindsight, a better procedure would probably have been to store a complete screen scan taken with no beam, and to subtract it in its entirety from all the other complete scans. This would have eliminated concerns about the variation of the baseline from sub-scan to sub-scan, assuming that the baseline scans were taken often enough to track variations in the camera performance with time.

3. Pedestal suppression—Baseline subtraction is sufficient to take out the variation of the baseline across each horizontal sub-scan, and to bring the resultant baseline within a few mV of zero (compared to several hundred mV for a typical beam-related signal). It was found, however, that sometimes there would be a clear pedestal, or voltage offset, remaining after the baseline subtraction. Hence, after baseline subtraction, **hvscan** can be directed to find the pedestal of each sub-scan. It then smooths these pedestals as a function of scan number (in order to prevent noise from causing abrupt variations), before subtracting the pedestals. In most cases, pedestal adjustment is unnecessary (the effect of pedestal adjustment cannot even be seen in the Figure), amounting to a few mV compared to a signal of several hundred mV.

4. Masking—the data was masked by a window of user-specified dimensions, to eliminate known non-beam-related features appearing on the images (e.g., reflections from the screen holder). For most of the scans, 10% of the data from around the edges of each scan was masked out. More drastic masking was also necessary for some scans in order to eliminate obviously spurious signals, such might be caused by RFI. Contour graphs of the intensity distributions where used to verify that the signals being masked out were indeed spurious (i.e., obviously not part of the beam).

5. Noise suppression—A sequence of noise-suppression stages was employed:

   (a) After subtraction of the baseline and pedestal, any pixel values below zero were assumed to be noise, and were set to zero.

   (b) Any non-zero pixel surrounded with zero pixels above and below, or to the left and right, was assumed to be noise, and set to zero. This proved

invaluable in eliminating spurious peaks due to RFI, though it may have slightly truncated the edges of some beam images.

(c) The peak pixel was found, which after the previous stages is almost guaranteed to be located in the image of the beam (i.e., to be unrelated to noise or baselines). The program then zeroed all pixels that were not connected to this pixel by a path going through other non-zero pixels.

6. Analysis—various analyses are available on the resultant pixel map:

(a) Fractional-widths, RMS-widths, intensities, and centroids can be computed for any vertical or horizontal slice through the pixel map. The various width measures are described below.

(b) The vertical and horizontal slices going through the peak of the pixel map can be written to disk.

(c) The pixel map can be collapsed vertically into a single horizontal scan, which can be written to disk and analyzed for width, intensity, and centroid.

(d) The pixel map can be collapsed horizontally into a single vertical scan, which can be written to disk and analyzed for width, intensity, and centroid.

(e) Diagnostic output on the baseline and pedestal subtraction is available.

(f) Data for making contour plots can be prepared.

An alternative to the somewhat intuition-based noise-suppression techniques used here would be to use digital filtering. Unfortunately, digital filtering is inapplicable because the number of samples per sub-scan is too small, meaning that the noise frequencies are not sufficiently above the signal frequencies to allow suppression of noise without obvious corruption of the signal. The danger of the noise suppression technique used here is that it may eliminate valid information around the edges of the beam image. However, there is no indication that this is a problem in the data I have analysed: examination of the resultant beam profiles show no indication of significant truncation of the actual beam image.

5.4.6 Measures of the Beam Size

The use of "collapsed" vertical and horizontal scans proved particularly valuable in improving the smoothness of the variation of the beam width with quadrupole strength. The use of slices through the beam peak is somewhat problematical in that the location of the peak is in some cases poorly defined and also susceptible to noise. Two measures of beam-width are computed by **hvscan**: the ±34.13% half-width (hereafter the "±34% half-width") about the median and the weighted RMS deviation from the median.

The ±34% beam-width, which I call ω, is identical to σ for a gaussian beam, but is well-behaved and easy to compute for a non-gaussian beam. It is also much less noise-sensitive than the weighted RMS deviation, and hence, initially at least, looked promising as a "robust" beam-size measure. ω for a scan V(t), whether a collapsed scan or a slice, was determined by first normalizing the area under the scan, such that

$$\int_0^T V(t)dt = 1, \tag{5.63}$$

where t=0 is the start of the scan and t=T the end. The median, T_m, was then found, defined by

$$\int_0^{T_m} V(t)dt = \frac{1}{2}. \tag{5.64}$$

The ±f fractional width, $\Delta T(f) = T_u - T_l$, is computed using

$$\int_{T_m}^{T_u} V(t)dt = f \tag{5.65}$$

and

$$\int_{T_l}^{T_m} V(t)dt = f. \tag{5.66}$$

where f=0.3413. The interval from $t = T_l$ to $t = T_m$ contains 34% of the intensity, as does the interval from $t = T_m$ to $t = T_u$. If V(t) where a gaussian function of t, then $\omega = \Delta T/2$ would be equal to the sigma of the gaussian.

The RMS-deviation, represented by τ, is also equal to σ for a gaussian beam, but not is necessarily equal to the best-fit σ that one might obtain by fitting a gaussian to an arbitrary distribution. τ for any scan V(t) is given by

$$\tau = \sqrt{\int_0^T V(t)(t - T_m)^2}. \tag{5.67}$$

where, as above, $V(t)$ is normalized to unit area.

The reader may wonder why I did not fit gaussians directly to the data. The reason is that the beam profiles are frequently so non-gaussian that this procedure would make little sense. Examples of beam profiles will be shown below.

Simulations of this method of measuring the emittance show that both ω and τ provide reasonable beam-size measures for use in place of σ in equation (5.48), even for very non-gaussian beams. Indeed, use of τ allows exact recovery of the RMS emittance even for non-gaussian beams. This should perhaps not be surprising, since the theory upon which the method is based deals explicitly with the transformation of the second moments of the particle distribution and since the RMS emittance depends only on the second moments of the particle distribution. The development of the Σ-matrix formalism in no way assumes that the beam being described has a gaussian phase-space distribution[68]. The use of ω, on the other hand, leads to errors for non-gaussian beams, as might have been anticipated. ω is, however, attractive for experimental work in that it is less susceptible to errors introduced by noisy "wings". Simulations with the lattice used for the experimental measurements shows that use of ω tends to over-estimate the emittance. Since **elegant** calculates τ and ω, it is possible to make direct comparisons of actual measurements and simulated measurements using the two beam-size measures, and I shall report on such comparisons below.

5.4.7 Imaging and Phosphor Resolution

The resolution of the imaging system[88] used to acquire the screen images will effect how accurately the emittance can be measured, primarily by affecting the ability to determine the minimum beam-size. (See the discussion above related to equation (5.62).) The resolution depends on a number of factors: the resolution of the screen material (i.e., how large a emitting area would be activated by a zero-emittance electron beam), the resolution of the camera tube (i.e., how many pixels), the magnification, and the contributions of diffraction in the camera lens.

As a step toward determining the resolution, I used a back-lit slit of variable gap (actually the gap of a set of precision calipers), and measured the FWHM of the

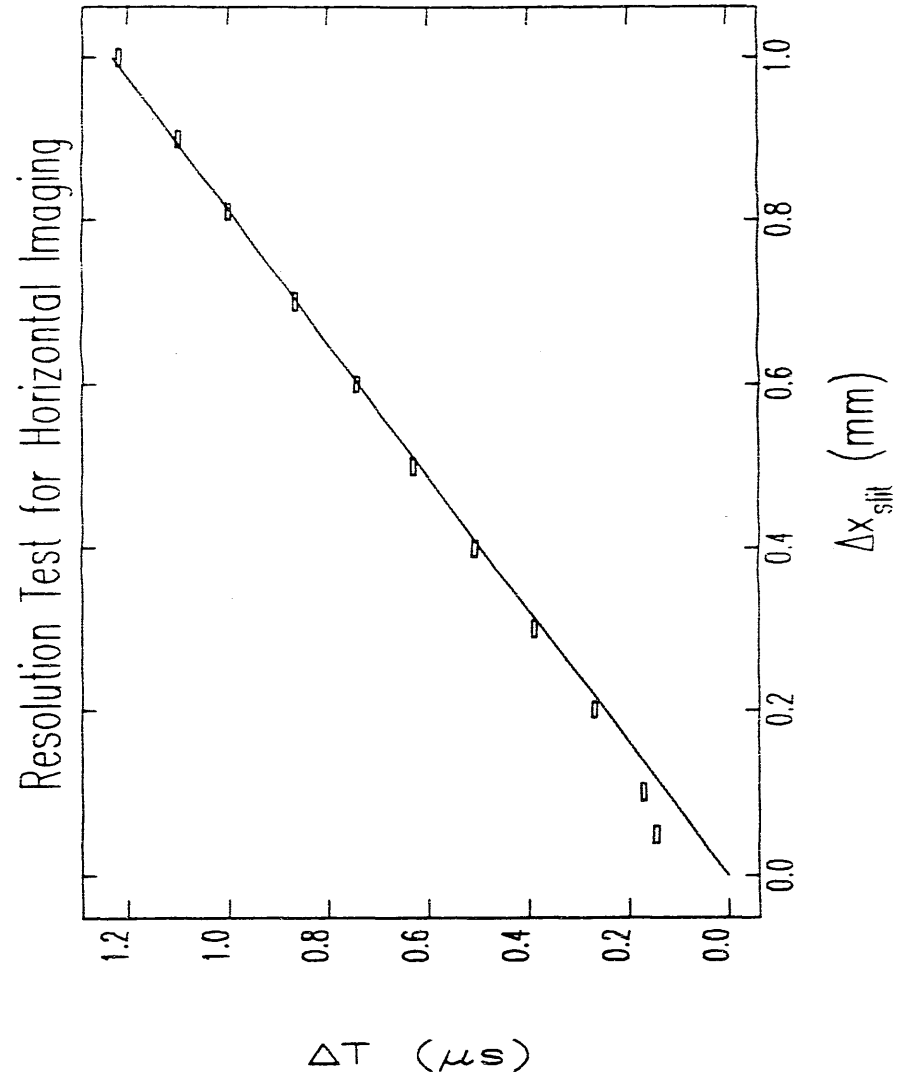

Figure 5.24: Resolution Test for Horizontal Imaging

digitized video signal as a function of the slit spacing. The results for the horizontal scan are shown in Figure 5.24, from which one sees that the resolution (the FWHM ΔT measurement for a hypothetical zero-width slit) is by extrapolation about 0.15 μs. This corresponds to a half-width resolution of 0.075 μs.

The corresponding experiment for the vertical scan is dubious, since the TV signal is broken into a sequence of horizontal scans, with each frame containing 240 such scans. The spacing between these scans is 64 μs, so that the resolution is clearly worse than this: the smallest vertical full width that can be measured is 64 μs, which I will take to be the vertical resolution. The corresponding half-width resolution is 32 μs.

For the beam measurements, as just discussed, I have used the parameters ω and τ to characterize the measured beam size. It is necessary that the resolution measurements be cast in an equivalent form, for comparison with ω and τ measurements for the beam.

The imaging resolutions above were given as half-width at half-max (HWHM) values. At the resolution limit, the video trace has a triangular shape, corresponding to the "lighting" of one pixel with the adjacent pixels being off:

$$V_\Delta(t) = \begin{cases} 0 & |t| > T \\ \frac{T-|t|}{T^2} & |t| \le T \end{cases}, \tag{5.68}$$

where the area is normalized to 1. The HWHM is, of course, simply $T/2$, while ω_Δ is defined by

$$0.3413 = \int_0^{\omega_\Delta} V_\Delta(t)dt. \tag{5.69}$$

from which $\omega_\Delta \approx 0.87*\text{HWHM}$. The RMS t value for $V_\Delta(t)$ is simply

$$\tau_\Delta^2 = \int_{-T}^{T} V(t)t^2 dt. \tag{5.70}$$

from which I obtain $\tau_\Delta \approx 0.82*\text{HWHM}$. These are close enough to allow simplification of the analysis by taking $\omega_\Delta = \tau_\Delta \approx 0.85*\text{HWHM}$.

Using the resolutions given above,

$$\omega_{x,\text{imaging}} = \tau_{x,\text{imaging}} = 0.85 \cdot 0.075\mu\text{s} \approx .064\mu\text{s} \tag{5.71}$$

and

$$\omega_{y,\text{imaging}} = \tau_{y,\text{imaging}} = 0.85 \cdot 0.032\text{ms} \approx .027\text{ms}. \tag{5.72}$$

For the magnification used in the experiments, these time resolutions in terms of correspond to spatial imaging resolutions of 58 μm horizontally and 55 μm vertically. The fact that these nearly agree is purely coincidental, but I shall take advantage of the coincidence and use 57 μm as the value for both planes.

In the simplest analysis, these resolutions add in quadrature to the actual beam width and the phosphor screen contributions, to produce

$$\omega_{\text{meas}}^2 = \omega_{\text{beam}}^2 + \omega_{\text{imaging}}^2 - \omega_{\text{phosphor}}^2 \tag{5.73}$$

$$= \omega_{\text{beam}}^2 - \omega_{\text{resol}}^2, \tag{5.74}$$

and similarly for τ in this and all subsequent equations. Rewriting this, one sees that

$$\omega_{\text{beam}} = \sqrt{\omega_{\text{meas}}^2 - \omega_{\text{resol}}^2}. \tag{5.75}$$

Thus, the actual beam width can be deduced from the measurements, with the uncertainty being

$$(\Delta\omega_{\text{beam}})^2 = \frac{(\omega_{\text{meas}}\Delta\omega_{\text{meas}})^2 - (\omega_{\text{resol}}\Delta\omega_{\text{resol}})^2}{\omega_{\text{meas}}^2 - \omega_{\text{resol}}^2}, \tag{5.76}$$

from which one sees that this procedure is highly uncertain unless $\omega_{\text{meas}}^2 \gg \omega_{\text{resol}}^2$, i.e., unless the use of equation (5.75) is a small correction. If this is the case, one sees that $\Delta\omega_{\text{beam}} \approx \Delta\omega_{\text{meas}}$. Estimation of $\Delta\omega_{\text{meas}}$ is discussed below.

The screen material used for the measurements I will analyse here (see below), is rather grainy and is reported to have a resolution of 50–100 μm[89]. (My own examination of the material with a microscope reveals pitting on the scale of 100–200 μm. If this is taken to define the "pixel" size, then the half-width resolution would be 50–100 μm, which is consistent with the published figure.) Adding this in quadrature with the imaging resolution gives $\omega_{\text{resol}}^2 = 75$–$115\mu$m. In order to simplify the analysis, it is convenient to use the average (95μm) of these two limits in what follows.

The minimum measured horizontal beam sizes are about 400 μm, whereas the minimum measured vertical beam sizes are about 220 μm. The use of equation (5.75) is thus acceptable for the horizontal and vertical planes.

5.4.8 Choice of Screen Material and Experimental Limitations

An additional source of uncertainty in the emittance measurements stems from the saturation and linearity of response of the screen material. Two different screen materials were installed for the chopper-screen. The first was $Gd_2O_2S:Tb$, which is the material chosen by SLAC for the SLC[90]. This material proved insufficiently robust for the beam intensities used in daily operation of the gun. As a result, the screen was damaged, making emittance measurements with it difficult (one had to try to steer the beam to an undamaged part of the screen, and keep it there while varying Q4 and Q5).

More recently, a more robust screen was installed, made of "SLAC chromate" ($Al_2O_3(Cr)$, material 4 in [90]). This material is grainy compared to $Gd_2O_2S:Tb$, but has good light output and (judging from the limited data presented in [90] and the statements made in [89]) linear response (i.e., the amount of light is proportional to the number of electrons). No experiments where done to specifically check the linearity (an oversight).

In an attempt to ensure that neither the phosphor nor the camera was saturating, I focused the electron beam to a very small spot and noted the signal level from the camera. I then decreased the beam intensity, noting when I had reached a beam intensity where the intensity of the video signal was indeed changing as the beam intensity decreased. I then maintained the beam intensity at a sufficiently low level to stay below about half of this signal level. As will be seen below, this precaution proved insufficient, and there is clear evidence of saturation in some of the measurements.

Another limitation on the emittance measurements was due to the thick substrate of the chromate screen, which resulted in unacceptable heating of the screen and out-gassing at beam currents (into the screen) greater than 100 mA. In a test, the vacuum increase for 300 mA was sufficient to cause an alarming increase in the GTL vacuum pressure, and an unacceptable increase in the gun vacuum pressure. As a result, no measurements could be made of the effect of beam current on emittance. The experiments reported on below were done for about 270 mA out of the gun, with the current out of the alpha-magnet further limited by the scraper. In retrospect,

data could have been taken for a larger range of gun currents for a sufficiently a small momentum spread. At the time, I was more interested in the emittance for large fractions of the beam, such as were intended to be used in actual operation.

5.4.9 Overview of the Experiments

Having used `hvscan` to process the scans for any set of measurements, the resultant beam-widths as a function of quadrupole strength were then available for calculation of the emittance. I wrote the program **`emitmeas`** to do this analysis, with **`elegant`** being used to calculate the matrices as a function of quadrupole strength. In addition, **`emitmeas`** and **`elegant`** can be used to do simulated emittance measurements, with various types of initial beam phase-space distributions (e.g., gaussian, uniform, or `rfgun`- or MASK-generated).

The experiments I shall report on here are by far the most complete set of results I have to date and give simultaneous horizontal and vertical emittance data as a function of momentum spread allowed through the alpha-magnet. Constant beam and RF conditions were maintained for the entire series, with $E_{p2} \approx 75$MV/m (estimated from the momentum peak of the beam) and $J \approx 5$A/cm^2 (estimated from E_{p2}, the current at GT1, and simulations of beam-losses in the transport line). The beam current into the alpha-magnet was 230 mA. The momentum peak was 5.026 ± 0.048 m$_e$c (this value is the average over the all spectrum measurements taken, one for each set of emittance scans).

As in previous chapters, I shall characterize the momentum spread in terms of the parameter f, defined as

$$p_{min} = p_{max}\frac{1-f}{1+f}, \tag{5.77}$$

where p_{min} is the minimum momentum allowed to pass through the alpha-magnet by the scraper, and p_{max} is the maximum momentum in the beam. f is essential the fractional half-momentum spread of the beam for which the emittance is measured. A total of eight emittance measurements were performed, four involving variation of Q4 and four involving variation of Q5. f was varied from about 0.08 to 0.03.

Unfortunately, the data involving variation of Q5 were seriously affected by camera saturation, as Figure 5.25 shows. The intensities plotted here are the signals

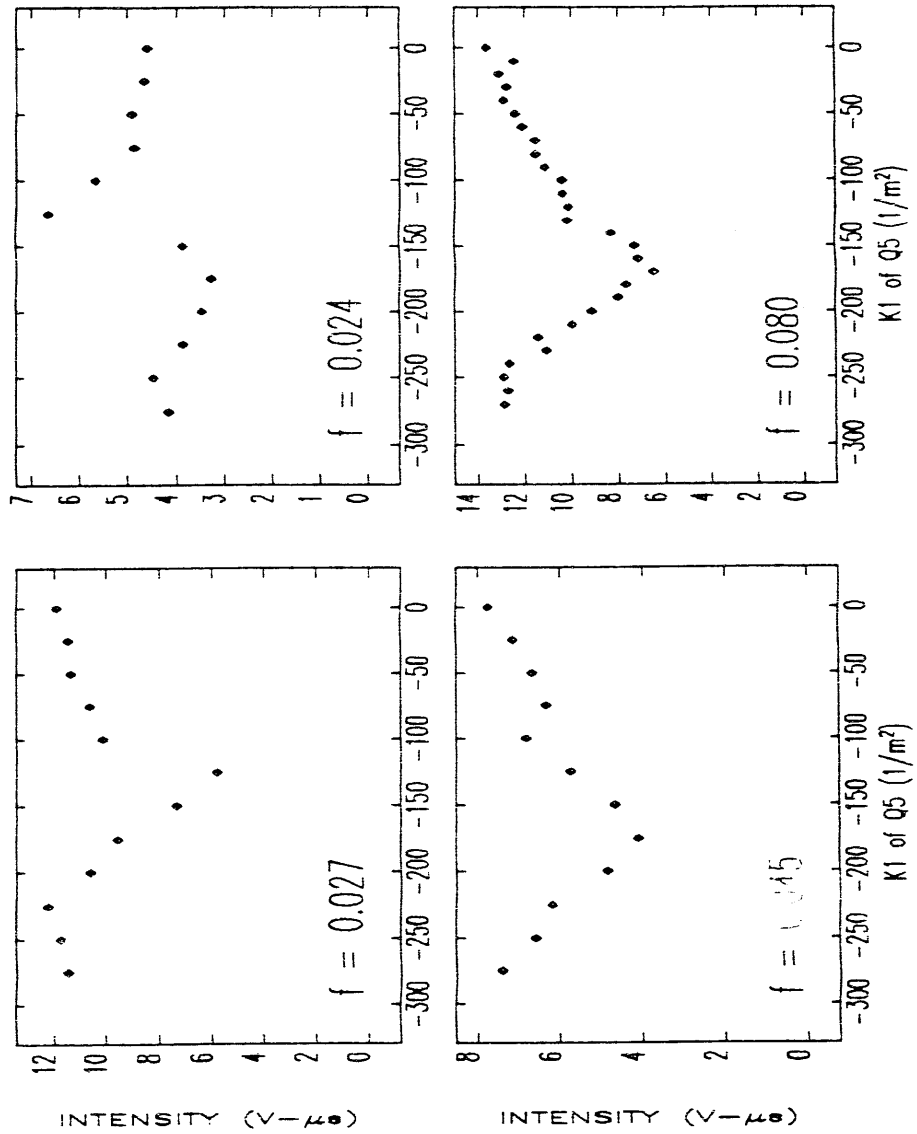

Figure 5.25: Integrated Video-Signal Intensity for Several Q5 Settings, for Various Values of f.

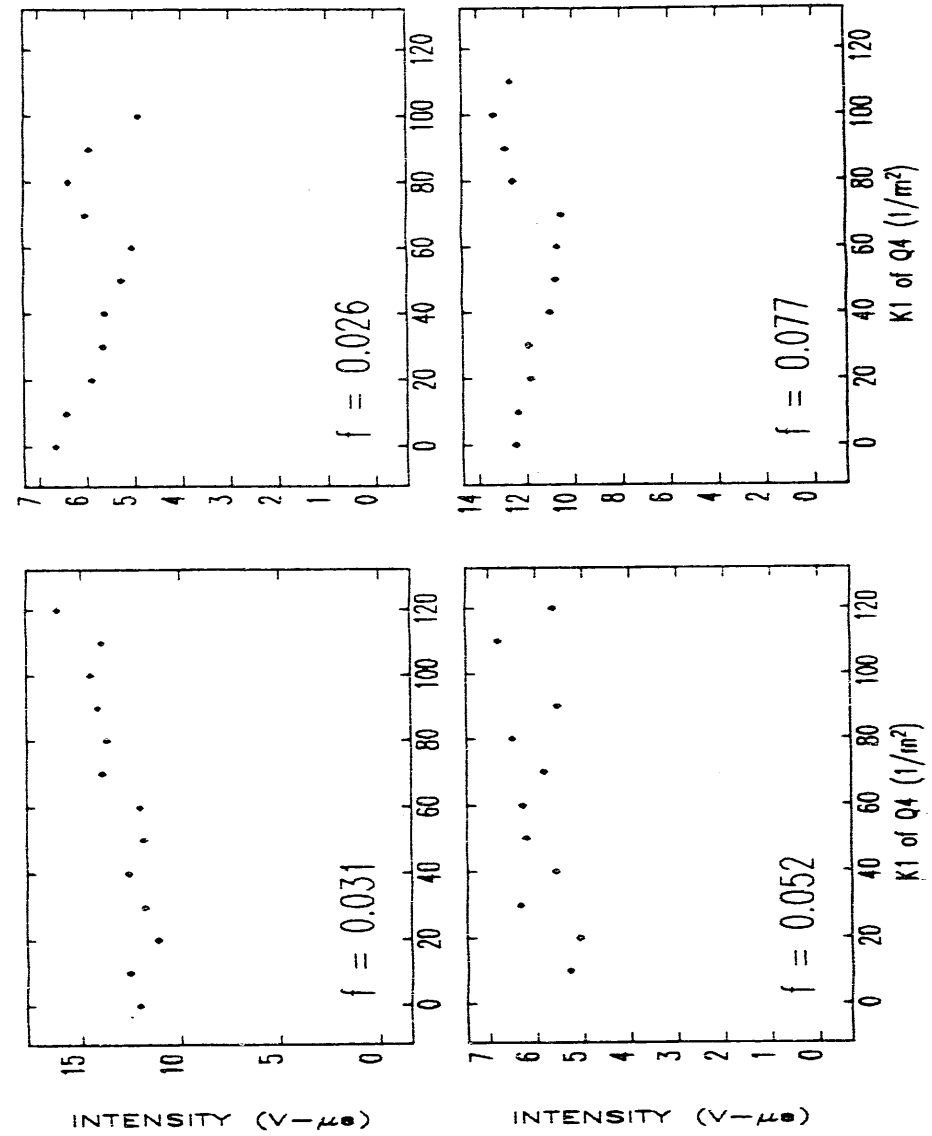

Figure 5.26: Integrated Video-Signal Intensity for Several Q4 Settings, for Various Values of f.

integrated over the pixel maps, and should be constant for a system with no saturation, since the beam current exiting the alpha-magnet was held constant. (Note that the intensity does not decrease monotonically with f, since the camera aperture was opened as f was decreased in order to get a better signal-to-noise ratio.) The dips in intensity in Figure 5.25 coincide with the minimum spot-sizes in the vertical plane, and clearly imply the presence of some saturation effect. While I will not take space to show the data, graphs of the horizontal and vertical beam-sizes versus the strength of Q5 give no indication that saturation is taking place, presumably because the saturation is a smooth phenomenon. Because of the dramatic saturation problem apparent for this data, I will not make further use of it.

Figure 5.26 shows intensity data for the measurements involving variation of Q4. The minimum horizontal beam-sizes occur around $70m^{-2}$, while the minimum vertical beam-sizes occur around $20m^{-2}$. In some cases, there is apparently systematic variation of the intensity, though the effect is mild compared to the previous Figure. Since the detailed nature of the saturation is not understood, I will simply use the data for Q4 "as is".

Since there is a great deal of data involved in these measurements, I will confine myself to presenting only a sample. In particular, Figure 5.27 shows a series of x-y contour graphs of beam-intensity. Similarly, Figures 5.28 and 5.29 show collapsed x and y beam intensity distributions corresponding the the contour graphs. One sees from these Figures that the intensity distributions are not well described by a gaussian, which is why I did not fit gaussians to the data in order to deduce the beam-size, as discussed above. One can also see how the increasing strength of the horizontally-focusing quadrupole first focuses then over-focuses the beam horizontally, and how the beam is defocused vertically. Note also that the beam is not symmetric in x: this assymetry is genuine, and was clearly visible in the TV images during the experiments. It is a result of second-order x-y coupling by the alpha-magnet, and is qualitatively reproduced in simulations.

The axes in these graphs were converted to distances from time intervals using $\Delta x(\text{mm}) = (0.91 \pm 0.02)\Delta T_x(\mu s)$ and $\Delta y(\text{mm}) = (2.05 \pm 0.03)\Delta T_y(\text{ms})$. The calibrations were obtained from noting the positions on the TV scans of known features on the screen. Note that the errors in the calibrations of spatial size in terms of

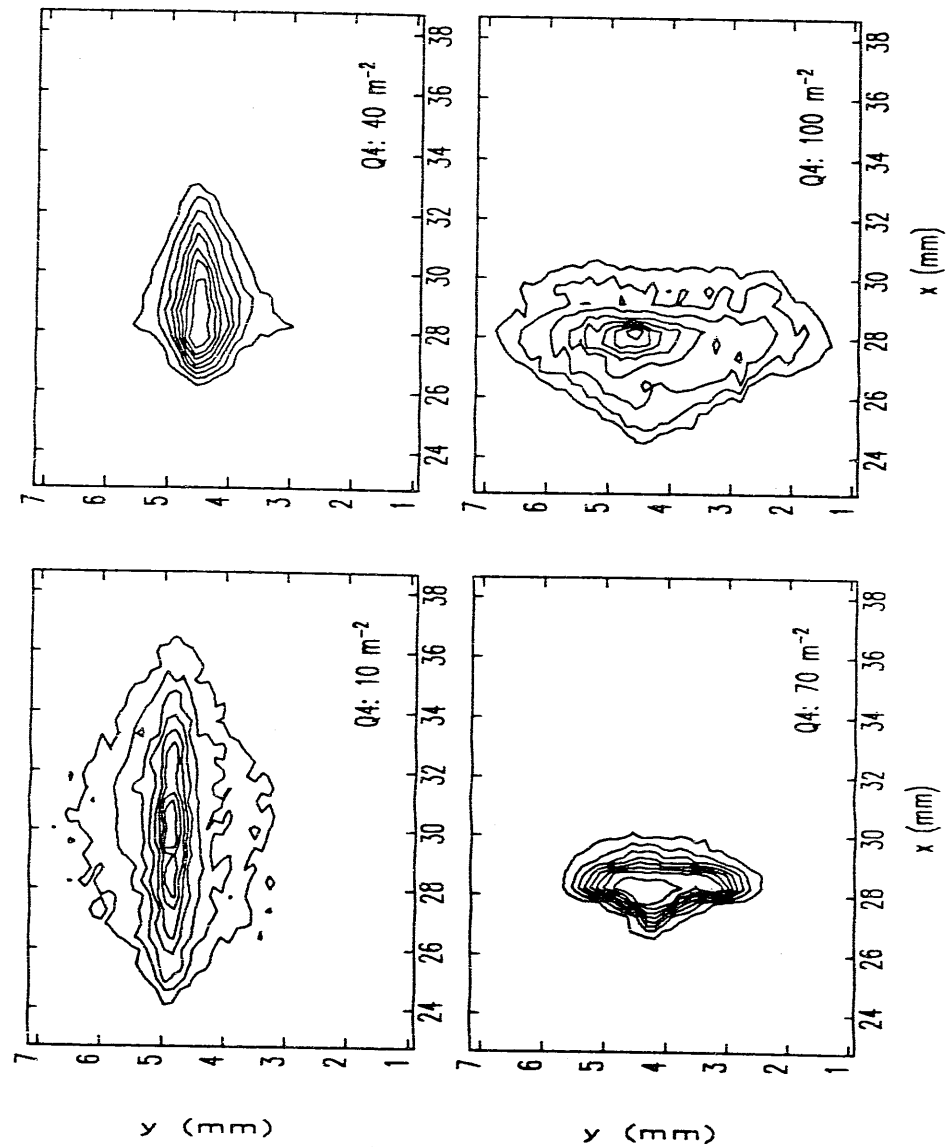

Figure 5.27: 10%-Contour Graphs of for Several Q4 Settings, from Experiments with f=0.08 .

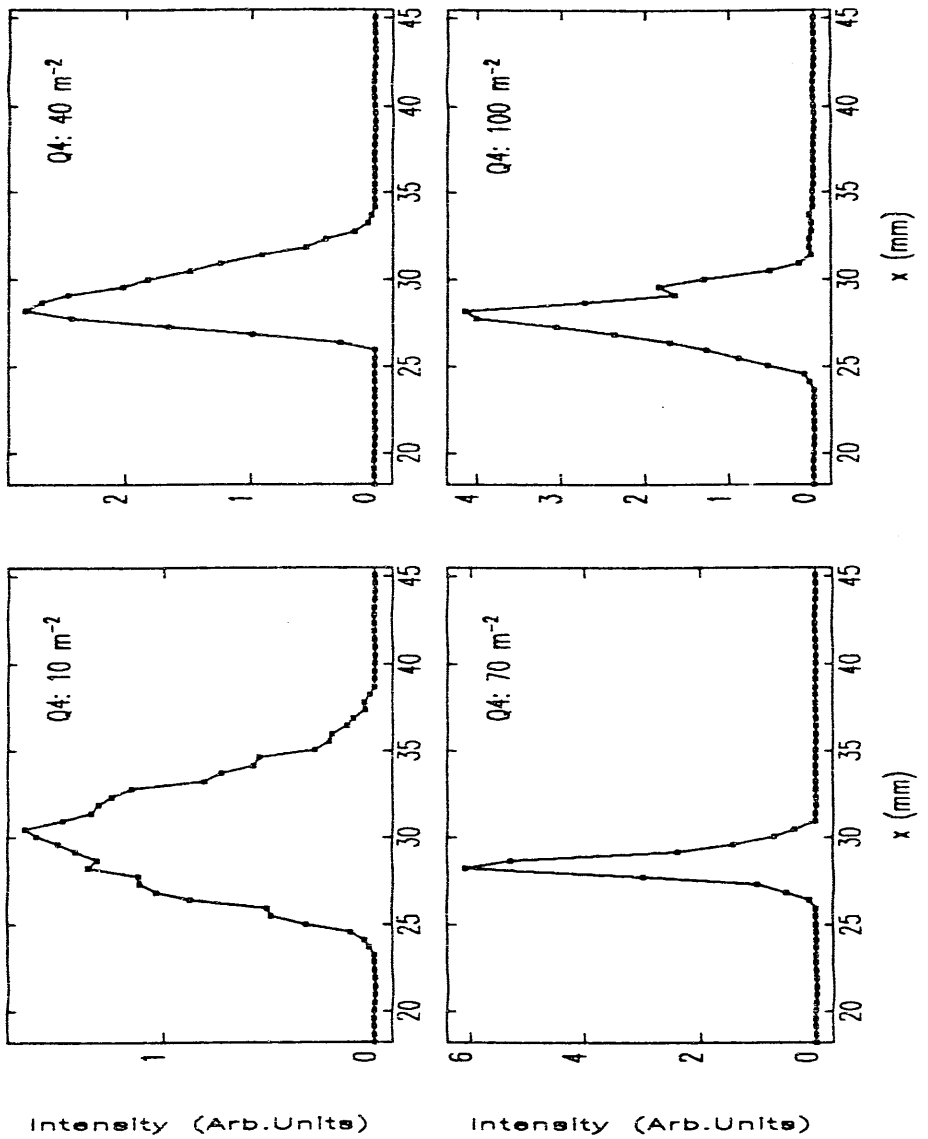

Figure 5.28: Collapsed Horizontal Beam-Intensity Profiles for Several Q4 Settings, from Experiments with f=0.08.

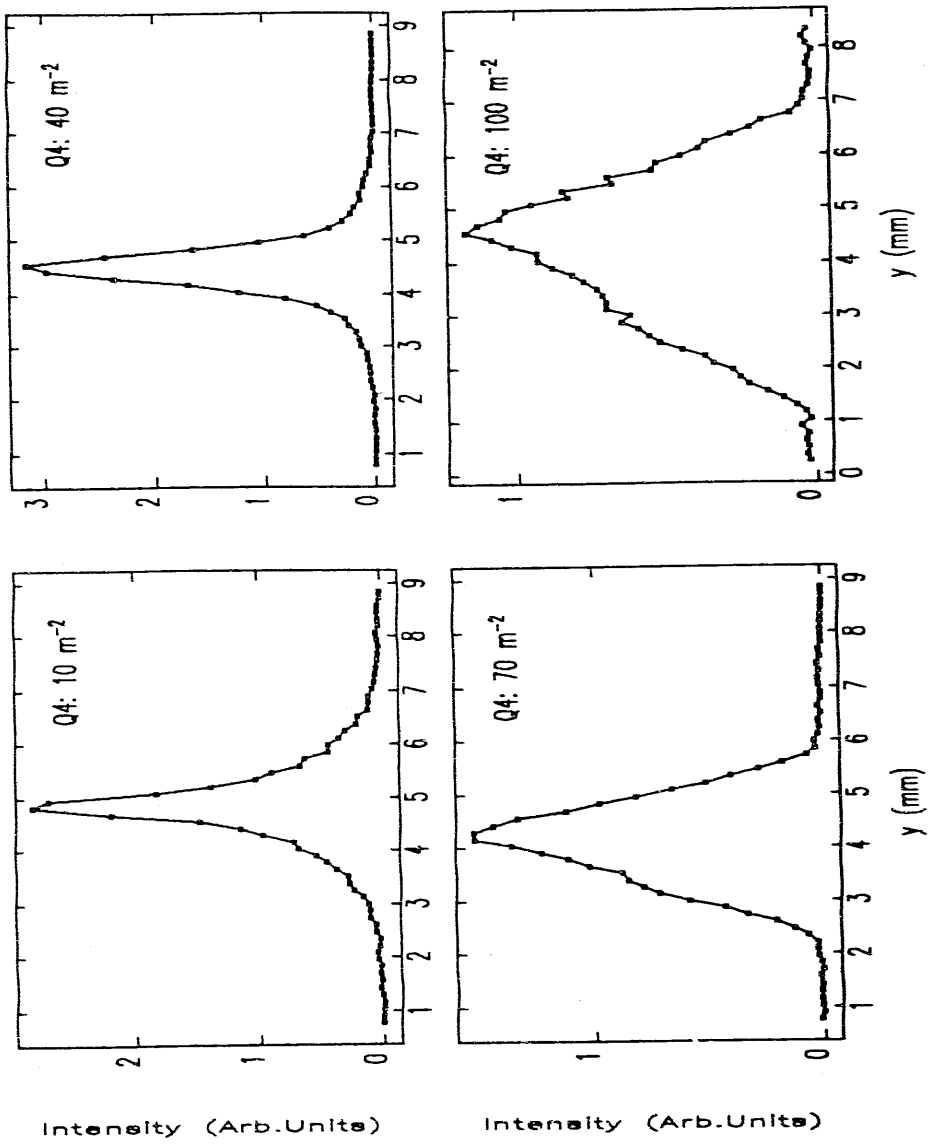

Figure 5.29: Collapsed Vertical Beam-Intensity Profiles for Several Q4 Settings, from Experiments with f=0.08.

time-interval imply errors in the measured beam-sizes and hence errors in the emittance derived from these. However, one cannot equate the fractional error in the calibrations with the (supposedly) normally-distributed fractional uncertainty of the measured beam-sizes, since the calibration error is systematic and implies no fluctuation of the measured beam-sizes about the fit.

The effect of calibration errors on the emittance can be estimated using the thin-lens solution given in equation (5.62). $\Delta\sigma_m/\sigma_m$ is equal to the fractional uncertainty in the calibration. Similarly, $\Delta D/D$ is equal to the fractional uncertainty in the calibration, since $D^2 \sim \sigma^2$. Hence, the fractional uncertainty in the emittance is $\sqrt{2}$ times the fractional uncertainty in the calibration. This amounts to about 3% for the horizontal plane and 2% for the vertical, which is insignificant compared to the effects of statistical errors in the beam-size measurements.

5.4.10 Estimation of Uncertainties

Knowledge of the normally-distributed errors in the measured beam-sizes can come only from statistical information gleaned from actual beam-size measurements. In addition, it is reasonable to assume that the resolution with which the beam-size can be measured is a lower limit on the uncertainty in any beam-size measurement.

In order to assess the uncertainty in the beam-sizes, I first fit the data using an assumed uncertainty equal to the resolution (i.e., 95 μm), in order to provide a covariance matrix for the weighted fit. This is important in obtaining good fits to the data near the minima, since the unweighted fit tends to fit the larger beam-sizes disproportionately well. To see why, note that

$$\Delta(\sigma_i^2) = 2\sigma_i\Delta\sigma_i. \tag{5.78}$$

The equal-weights fit to σ^2 assumes that $\Delta(\sigma_i^2)$ is the same for all points, which is equivalent to assuming larger uncertainties for the smaller σ_i. Hence, an equal weights fit should be expected to fit the smaller beam-sizes poorly, as is observed.

Having done the initial fit, I calculated the rms-deviation from the fit, defined by

$$r = \sqrt{\frac{1}{N-3}\sum_{i=1}^{N}(\sigma_i - \sigma_i')^2}, \tag{5.79}$$

where σ_i' represents the value returned by the fit, and where $N-3$ is used because there are 3 parameters in the fit, reducing the number of degrees of freedom. I then used the greater of r and the resolution as the actual uncertainty in the σ_i, and refit the data.

The errors in the Σ-matrix and the emittance can calculated using equations (5.54) and (5.56). As mentioned above, problems arose in doing this. In particular, the uncertainties in the y-plane emittance were very large, even though the fits were good, in the sense of having small deviations of the data points from the fit. The problem is related to the fact that the y-plane Σ-matrix is such that $\Sigma_{33}\Sigma_{44} \approx \Sigma_{34}^2$, so that small errors in the Σ-matrix can translate into large errors in the emittance. (This is not true for the x-plane, where Σ_{12}^2 is small compared to $\Sigma_{11}\Sigma_{22}$.) However, I am convinced that the error-propagation is inadequate and greatly over-estimates the actual effect. In equation (5.56), I included the effect of off-diagonal terms in the covariance matrix for the Σ-matrix, even though most analyses do not use these terms[85]. I found that the off-diagonal terms made large differences in the computed uncertainties, which is highly unusual.

To diagnose this further, I made use of a feature of `omitmeas` which allows the addition of normally-distributed random deviates to the beam-size data prior to fitting. In particular, I used `emitmeas` to add ensembles of normally-distributed random deviates to data that was known to be perfectly fit by equation (5.46) and to compute the Σ-matrix and emittance for each "tweeked" data set. (I used 500 error ensembles with 4-σ gaussian distributions.) Statistics where kept to permit computation of the average Σ-matrix, the rms deviation of the Σ-matrix elements from their averages, the average emittance, and the rms deviation of the emittance from the average. If the first-order error-propagation is adequate, then the average Σ-matrix together with the rms deviation of the elements of the Σ-matrix should, when used in equations (5.49) and (5.56), reproduce the emittance and its rms deviation as computed directly from the results for each ensemble.

For the x-plane data, the two methods agreed very well on the emittance itself, and within 20% on the uncertainty. For the y-plane data, however, the uncertainty in the emittance as computed via error propagation was wildly different from that computed directly from the emittance for the ensembles. Differences of several orders

of magnitude were frequently obtained even though the emittances themselves agreed to 10^{-6}.

Additional skepticism about the error-propagation when applied to the y-plane data comes from observing the effects of eliminating points from the fits. If the calculated emittance is in fact highly uncertain, then elimination of poorly-fit points should reduce the uncertainty and also change the emittance. Neither of these effects was observed. In one case, half of the data was eliminated with little change either in the calculated emittance or the uncertainty computed via error propagation.

I believe that the reason for this failure in the error propagation is the fact that the elements of the Σ-matrix are not independent, being physically constrained by the emittance to have a certain relationship. Hence, in the Monte-Carlo error analysis (to be described presently), the addition of errors causes all of the Σ-matrix elements to change in a constrained way, keeping the emittance approximately unchanged. In the error-propagation analysis, however, the elements of the Σ-matrix are assumed to be independent, which causes over-estimation of the effect of an error in one element of the Σ-matrix by ignoring necessary correlated errors in the other elements. While I have not attempted higher-order error-analysis or pursued this problem beyond what is stated here, it is clear that the Monte-Carlo method is the more reliable method, since it is, in effect, accurate to all orders.

Based on these conclusions, the analysis proceeded by first doing the initial fit and computing the rms deviation, r, as described above. To apply the Monte-Carlo method for the error analysis, I then computed the emittance and its uncertainty from statistics for 100 sets of 4-σ normally-distributed errors added to the x or τ data values, with the sigma of the normal distribution being given by the larger of r and the resolution. This procedure is performed automatically by `emitmeas`.

5.4.11 Experimental Results

Figures 5.30 and 5.31 show the horizontal and vertical beam-sizes as a function of the strength of Q4, for f = 0.08. Also shown are the fits to the data, as computed by `emitmeas` (with no errors added to the data). The data for the horizontal plane is quite well matched by the fits, with the results for the collapsed scans being the

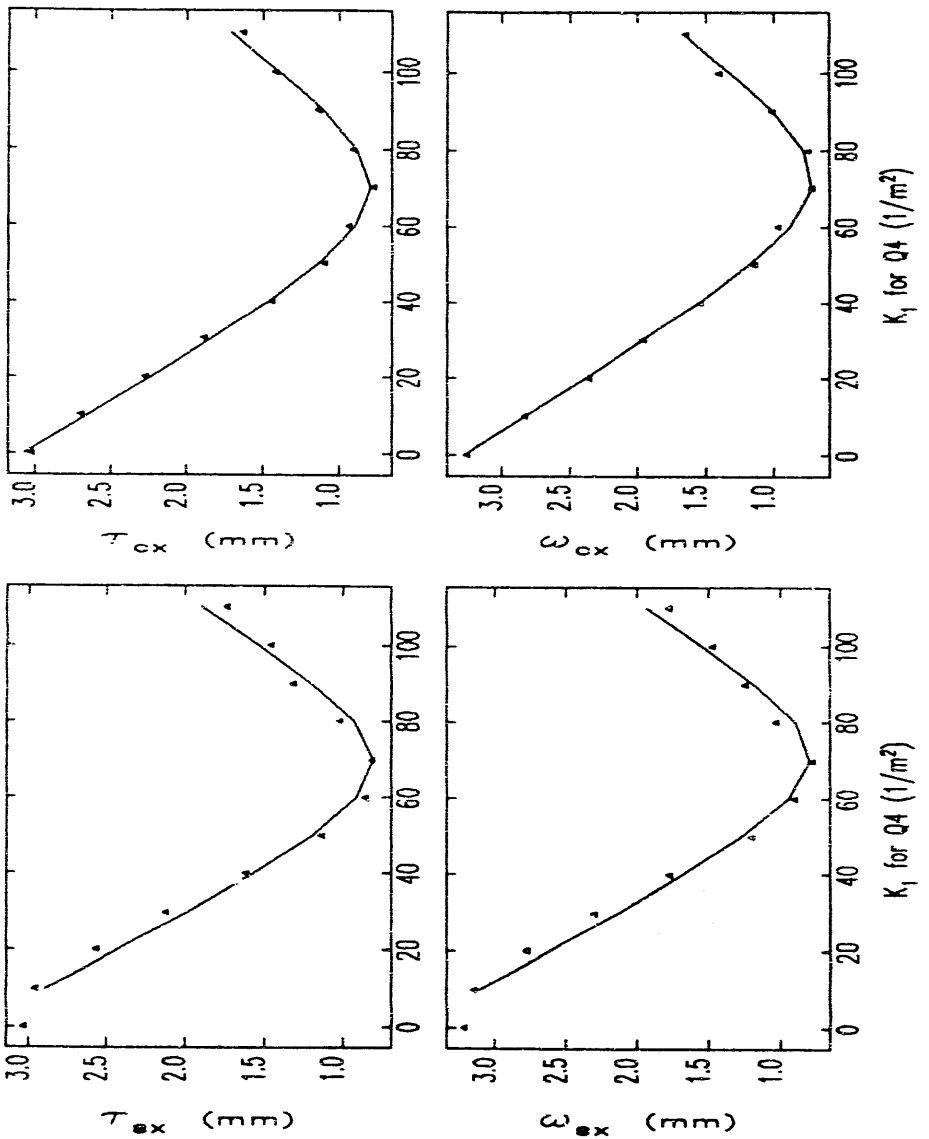

Figure 5.30: Horizontal Beam-Size versus Q4 Strength, for f = 0.08 . (Points are experimental data, lines are fits.)

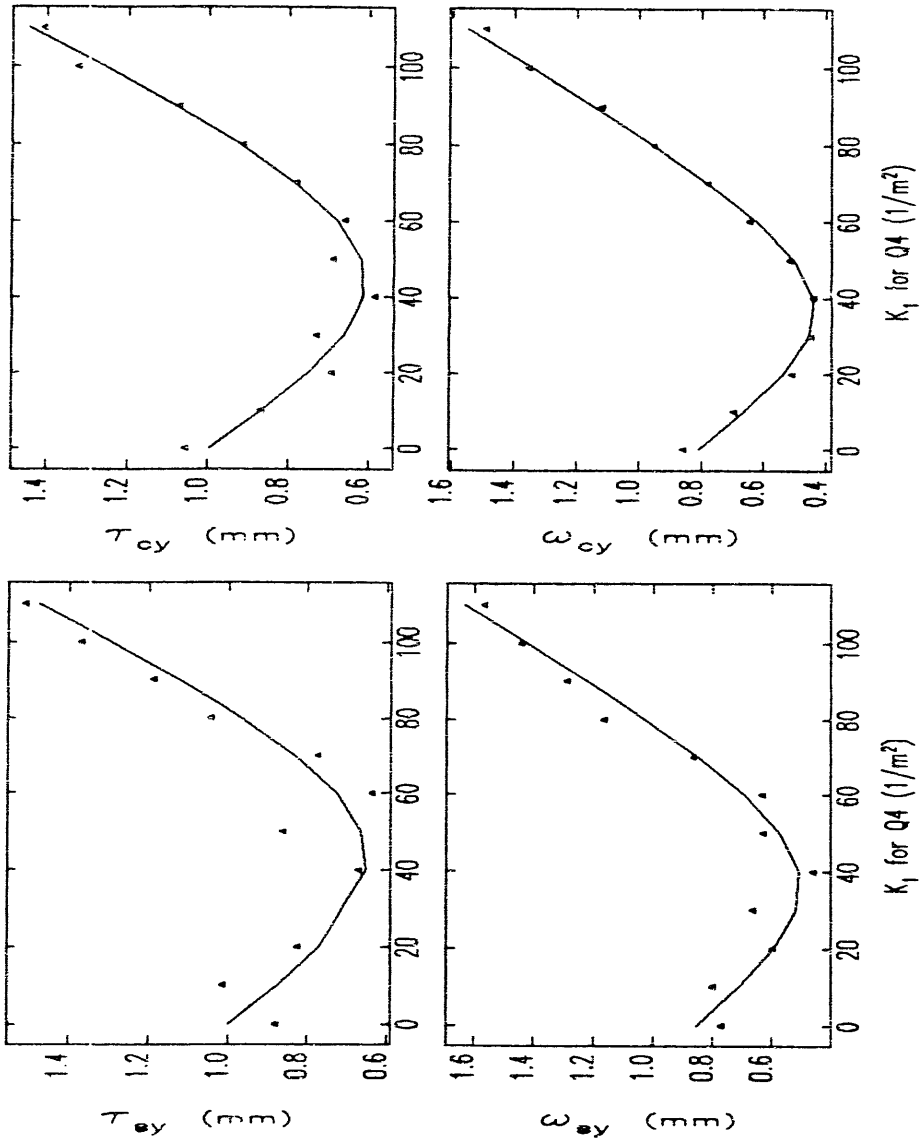

Figure 5.31: Vertical Beam-Size versus Q4 Strength, for f = 0.08 . (Points are experimental data, lines are fits.)

smoothest. The different beam-size measures, ω and τ, appear to agree reasonably well for the horizontal plane. Similar observations hold for the vertical plane, although the data appears to be noiser because the range of variation of the beam-size is smaller. The vertical τ data is considerably noisier than the vertical ω data, a pattern which holds in general (though not for the data in Figure 5.30).

The reader may wonder why the vertical beam size should decrease as the strength of the horizontally focusing quadrupole is increased. The reason is that the beam is vertically converging when it exits the alpha-magnet. With Q4 turned off, there is a beam-waist in the vertical plane before the beam reaches the screen. As the strength of Q4 is increased, the vertical defocusing causes the vertical waist to move toward the screen, resulting in smaller beam-sizes at the screen.

Figures 5.32 and 5.33 show the measured emittance as a function of the momentum spread, f. For each of the x and y planes, there is a set of four graphs. Each Figure shows the emittance as deduced from ω and τ beam-sizes for the central slice through the beam and for the collapsed scan. One sees that there is good general agreement between the emittances as deduced from the four different measures, with a clear trend toward smaller emittances at smaller momentum spreads. (In some cases, only three data points appear, because the beam-size data was too noisy to fit.)

5.4.12 Comparison of Experiments and Simulations

The mathematical basis for the emittance measurements, given in equations (5.42) through (5.50), assumes a mono-energetic beam. Until now, I have not dealt with the question of the validity of the method in the presence of significant momentum spread in the beam. At issue here is the effect of chromatic aberrations in the quadrupole that is varied.

The MASK data I shall use for all simulations in the remainder of this section is for $E_{p2} = 75MV/m$ and $J = 10A/cm^2$, which is close to the experimental conditions of $E_{p2} \approx 75MV/m$ (determined from the position of the momentum peak) and $J \approx 5A/cm^2$. Since the emittance is a weak function of J at low current densities (see Chapter 2), the difference between 5 and 10 A/cm^2 is negligible.

As a preliminary illustration of the importance of chromatic aberrations in the

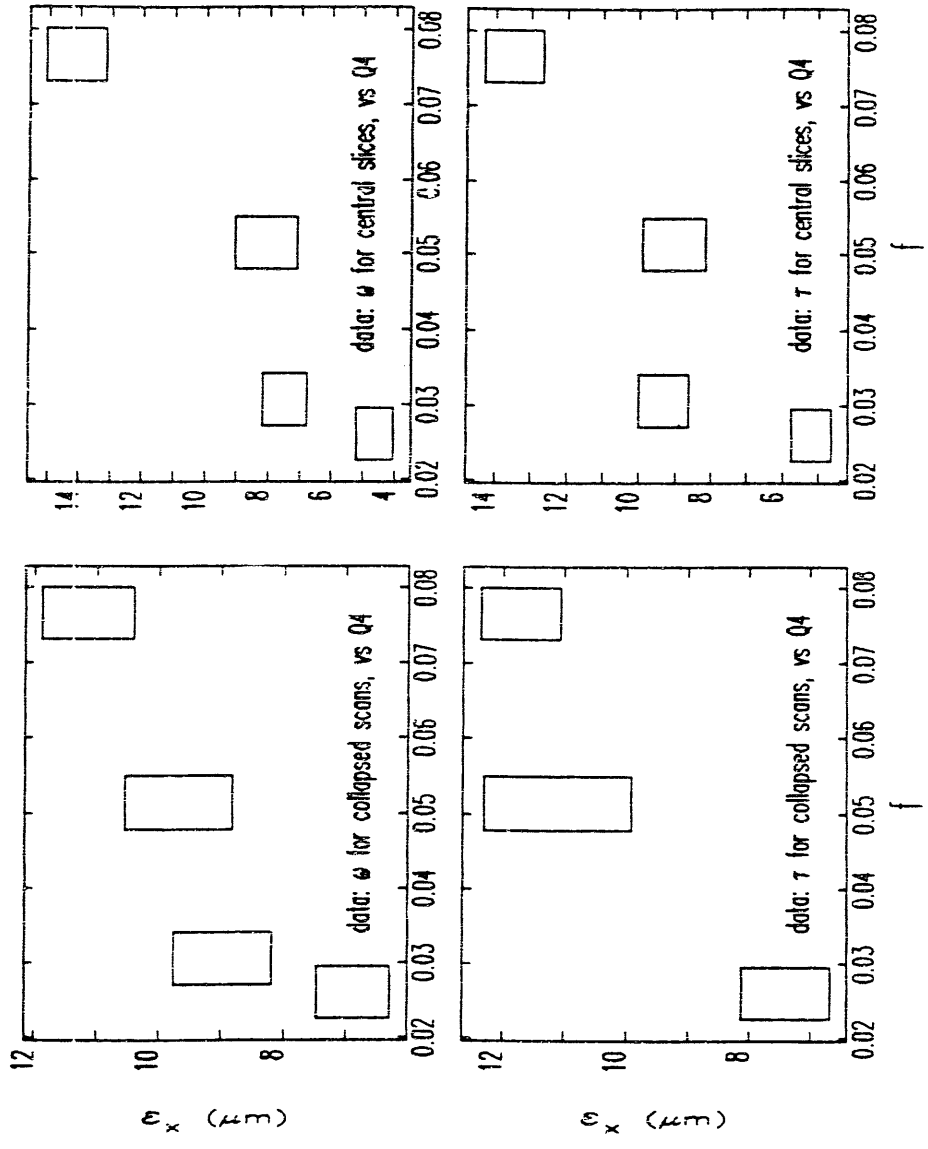

Figure 5.32: Horizontal Emittance versus f (Momentum Spread) as Inferred from Variation of Q4.

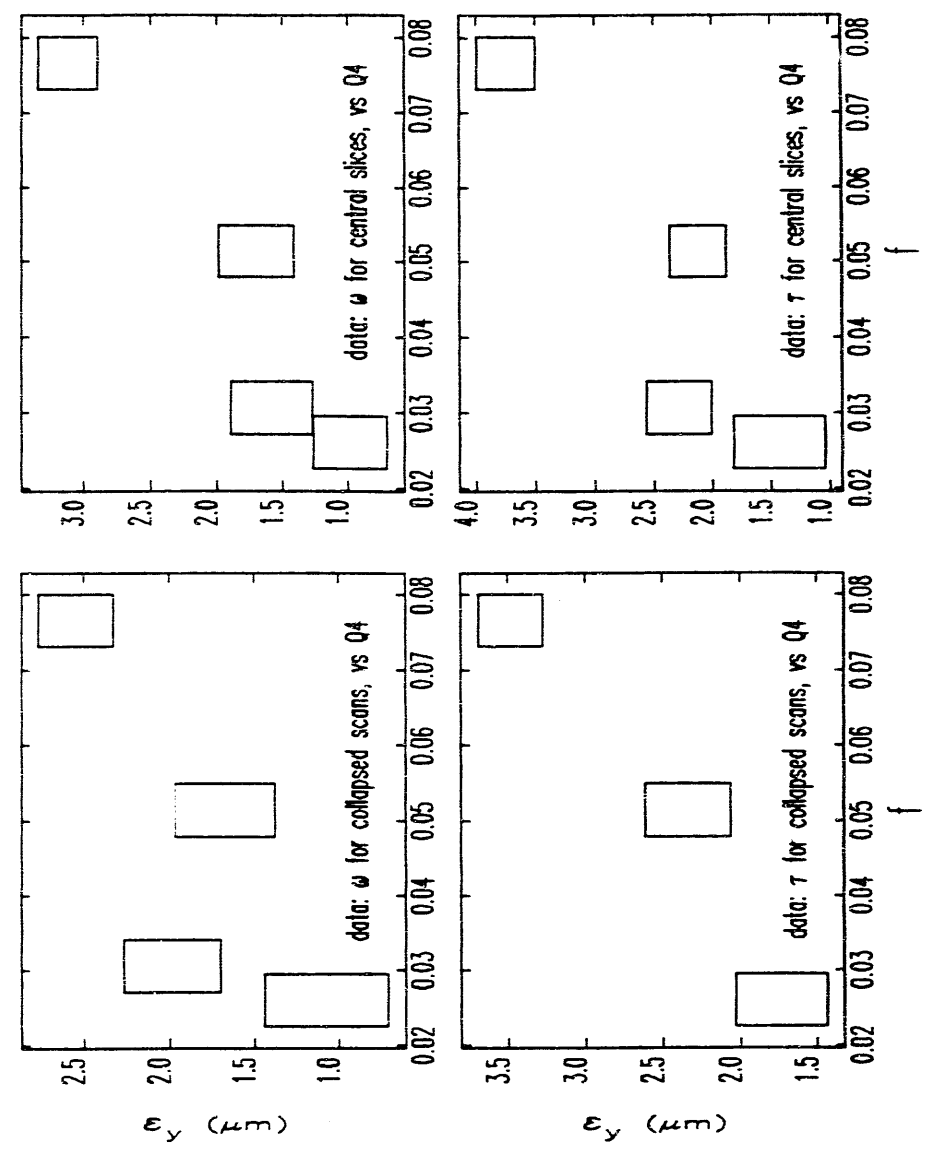

Figure 5.33: Vertical Emittance versus f (Momentum Spread) as Inferred from Variation of Q4.

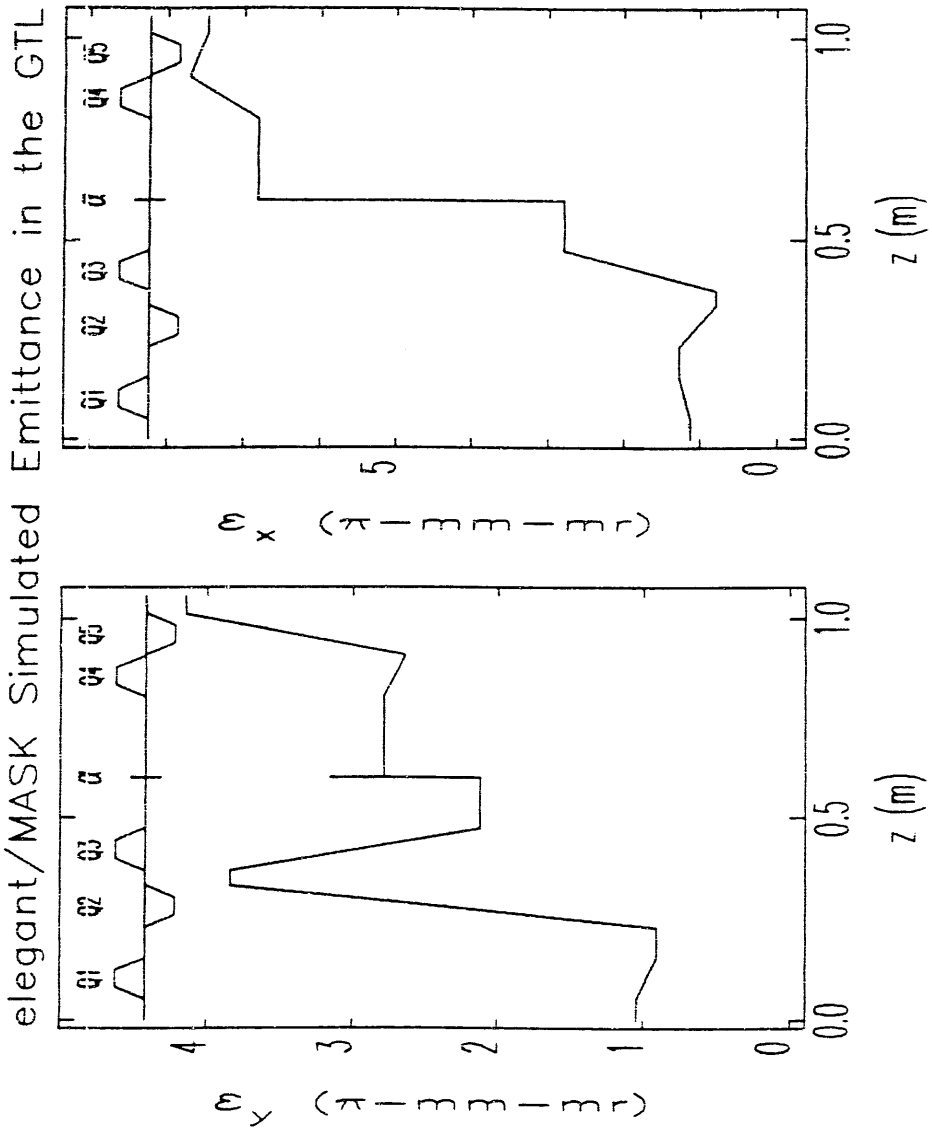

Figure 5.34: Variation of the RMS Geometric Emittance Along the Beamline for the Emittance Measurement Lattice with Q4 at 90 m^{-2}, from MASK/**elegant** Simulations.

experiments. Figure 5.34 shows an **elegant** simulation of the RMS geometric emittance in the x and y planes as a function of position along the beamline, where I have used the same optics as was used for the measurements (with Q4 set at 90 m^{-2}). The momenta of the simulation particles was pre-filtered in order to show the effect only of those particles that would actually exit the alpha-magnet, for f = 0.08 . In addition, the data point at the center of the alpha-magnet is not shown, since the x emittance at this point is greatly inflated due to the dispersion. As discussed earlier in this chapter, one sees that chromatic aberrations have a dramatic effect on the emittance as the beam travels along the beamline.

Next, I show the effect of that variation of Q4 has on the emittance at the chopper screen (again, this is for f = 0.08). Specifically, Figure 5.35 shows the emittance for horizontal and vertical planes as a function of Q4, with Q5 set to zero, and in addition shows the emittance for Q4 set to 90 m^{-2} while Q5 varies. The emittances in these graphs are computed directly from the coordinates of the simulation particles after tracking through the beamline with **elegant**. Since the emittance is not constant as the quadrupoles are varied, it is a valid question whether obtaining the emittance by measuring the beam-size as a function of quadrupole strength is applicable.

To answer this question, I have run **elegant** for a series of alpha-magnet slit positions corresponding to those used in the experiment. **elegant** tracked simulated electron bunches for a series of values of Q4 (or Q5), dumping the transport matrix and the collapsed beam-sizes (both σ and τ) to disk for subsequent analysis with `emitmeas`.

Figure 5.36 shows the predicted variation of horizontal beam-size σ_x as a function of Q4 for f = 0.08, along with the experimental data to which the simulations correspond. (These beam-sizes are for the collapsed scans.) One sees that in general the simulation and experiment agree very poorly, with substantial agreement occurring only for σ_x versus Q4.

The differences are particularly dramatic for the vertical plane. In the experiments, the vertical beam-size went though a clear minimum as Q4 was varied. In the simulation, however, the beam-size does not go through a minimum in either case. This may well be related to the uncertainty in the alpha-magnet optics for the vertical plane. As discussed in Chapter 3, the effect of field errors in the alpha-magnet

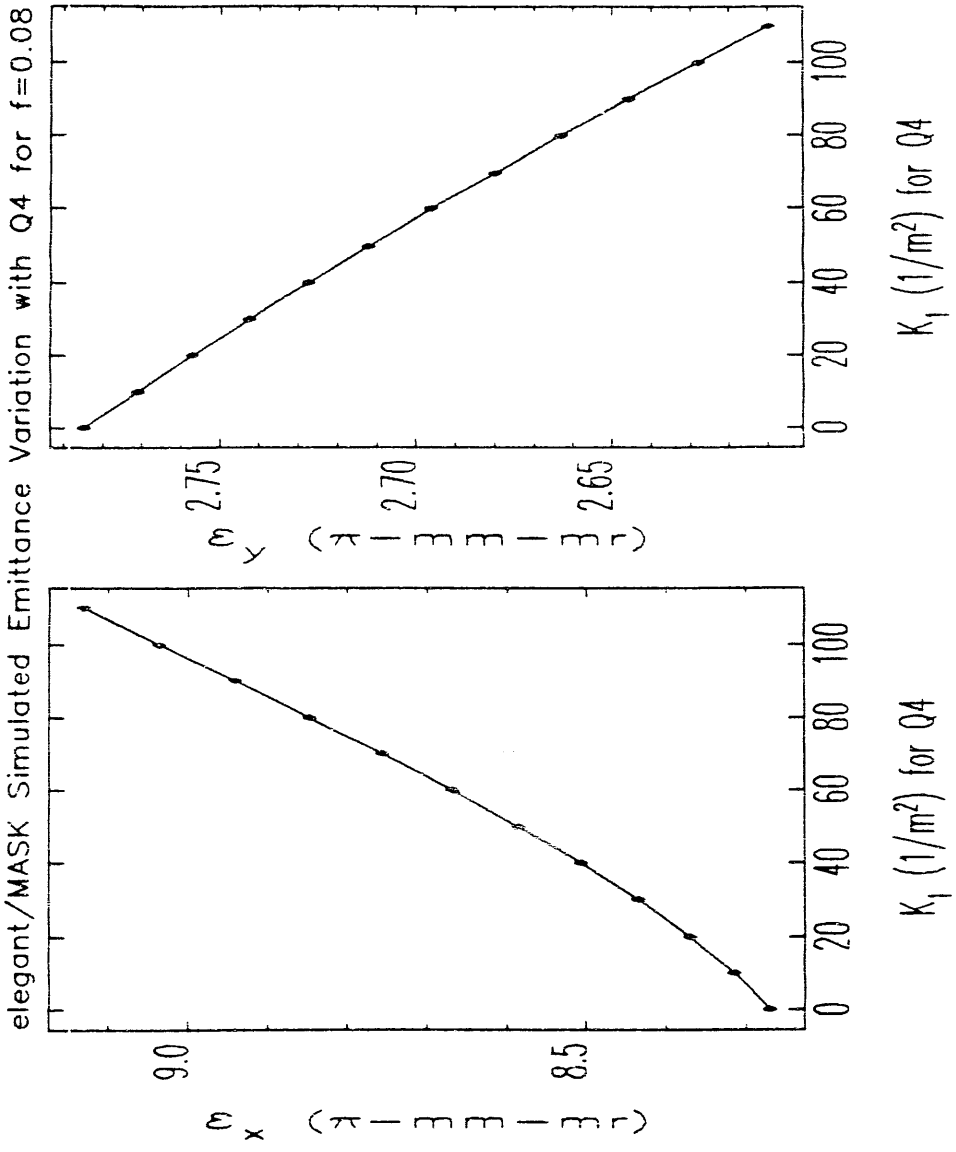

Figure 5.35: Emittance at the Chopper Screen as Altered by Q4, for f = 0.08, from `MASK elegant` Simulations.

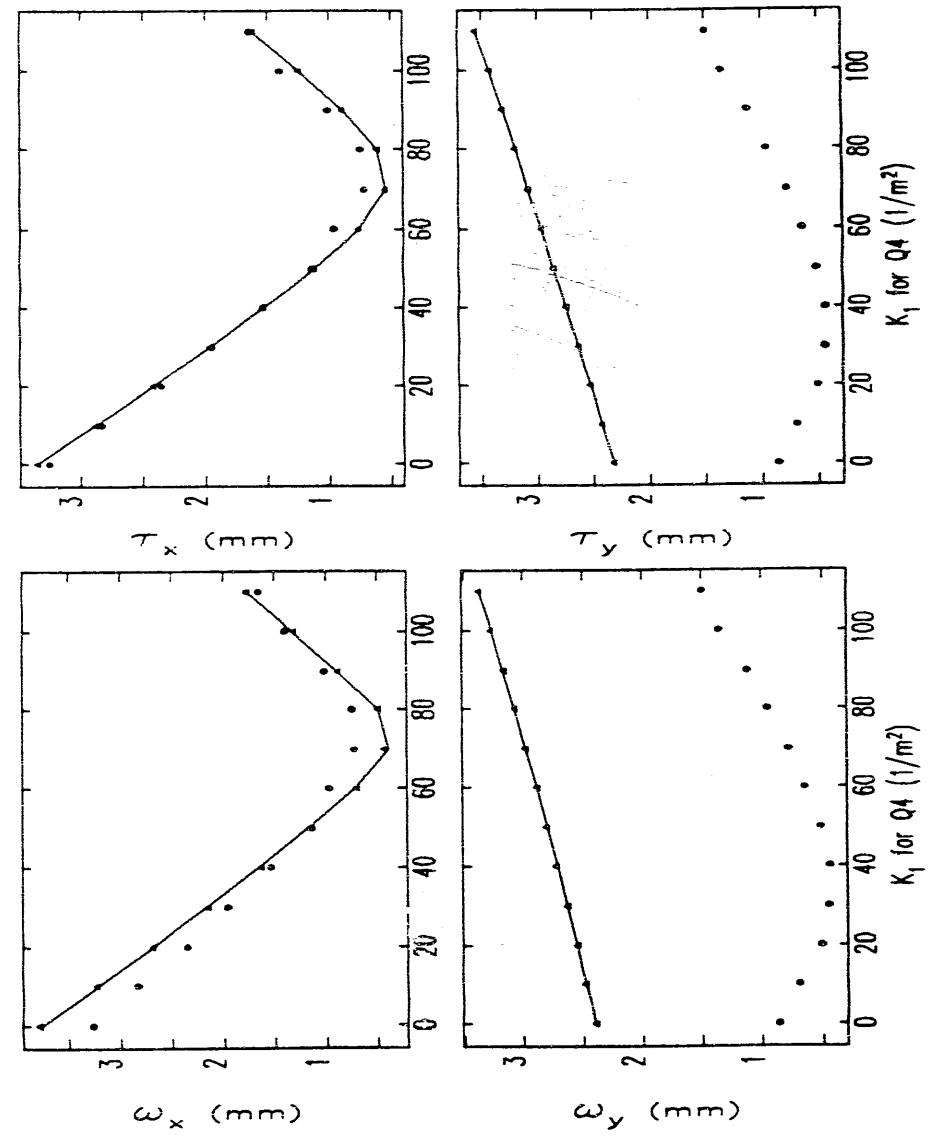

Figure 5.36: Beam-Size as a Function of Q4 Strength, from **`MASK/elegant` simulations** with f = 0.08

is much greater for the optics in the vertical plane than for the horizontal plane, and hence the correspondence between simulation and experiment may be expected to be worse for the vertical plane than for the horizontal plane. More work is needed to ascertain whether the known field errors in the alpha-magnet are sufficient to account for the observed discrepancy.

This does not mean, however, that the emittance measurements for either plane are invalid, since the emittance measurement depends only on the matrix from the beginning of the first varied element to the point of observation. Assuming that the errors in the alpha-magnet are only significant for the first-order matrix (an assumption I have not proved), the emittance would be unchanged, though the Σ-matrix would be altered. This would change the variation of beam-size with either Q4 as well, without changing the emittance that would be deduced from that variation.

Figures 5.37 and 5.38 show the emittance measurements for the collapsed scans (previously shown in Figures 5.32 through 5.33) along with simulation results obtained from **elegant**, **rfgun**, and MASK. As in the earlier Figures, the boxes represent the experimental measurements. The solid lines are the emittance as inferred from simulated measurements, using **rfgun**-generated initial particle distributions tracked by **elegant**, then analysed by **emitmeas**. (On some of the graphs, the solid line is missing because the beam-size did not go through a minimum, and the fitting failed.) The dotted lines are the upper and lower limit of the emittance in these simulations, indicating the variation of the emittance due to changing chromatic aberrations. Similarly, the dashed lines are the upper and lower limit of the emittance when MASK-generated initial particle distributions were used.

From Figure 5.37, one sees that use of ω as a beam-size measure is problematic, despite the appeal of the measure because of its robustness. I now refer specifically to simulation results, and the discrepancy between the emittance deduced from ω and upper and lower bounds on the RMS emittance computed by explicit averaging over the coordinates of the simulation particles. By contrast, in Figure 5.38 one sees that the emittance deduced from the RMS beam-size is generally quite close to the range through which the emittance actually varies in the simulation.

One would expect that as the momentum fraction was decreased, the inferred

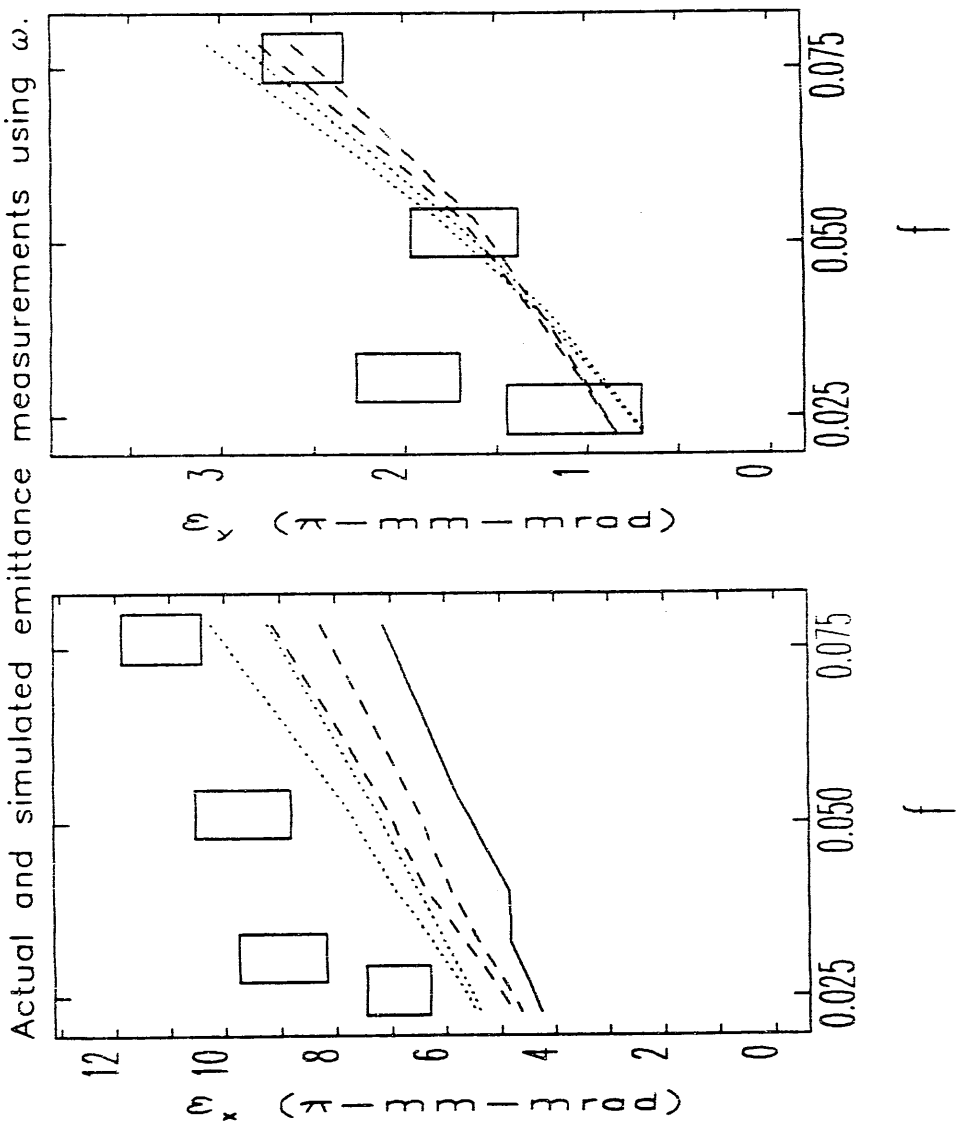

Figure 5.37: Emittance from ω vs. Quadrupole Strength, as a Function of Half-Momentum Spread, f, from Simulations and Experiment. (See text for explanation.)

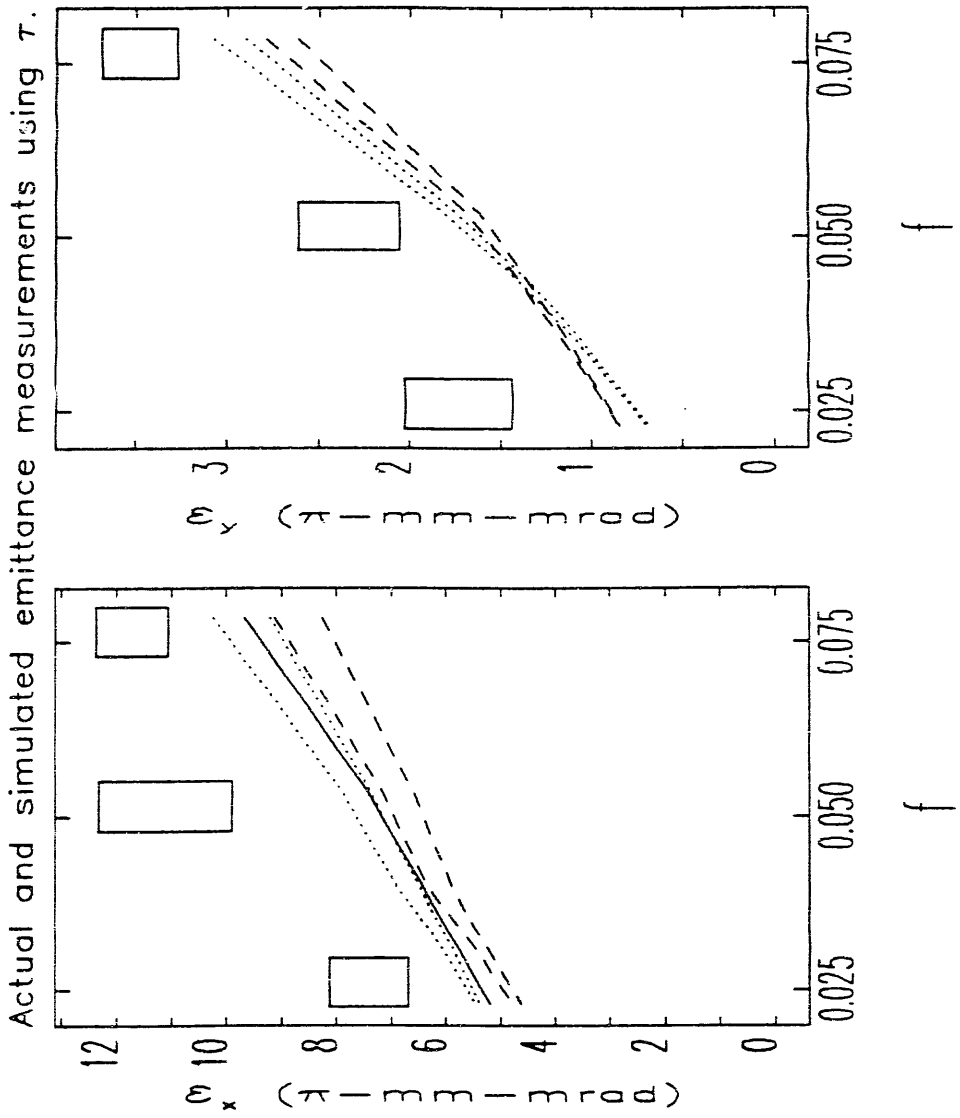

Figure 5.38: Emittance from τ vs. Quadrupole Strength, as a Function of Half-Momentum Spread, f, from Simulations and Experiment. (See text for explanation.)

emittance would approach the boundaries set by the variation of the explicitly computed emittance. From the Figures, one sees that this is not always true, even when τ is used as the beam-size measure. The reason is that, even for comparatively low momentum spreads (as low as $\pm 2.3\%$ in the simulations), there is still some dependence of the emittance on quadrupole strength. This is apparently sufficient to "throw off" the fits and make the emittance appear to be something different from what it actually is. I am confident that these results are not spurious, since **emitmeas** has been tested thoroughly using zero-momentum-spread simulation beams; in tests, the emittance for such a beam proved recoverable to the very high accuracy (limited by the fact that the data delivered to **emitmeas** was only good to single-precision accuracy). Note that the effects of momentum spread on the the emittance measurements are quite unlike the effects of noise, in that the latter do not cause systematic variations.

5.4.13 Possible Sources of Discrepancies

While there is some agreement between the simulations and the experimental results, the measurements on average give emittances that are 50% higher than the simulations predict. Possible sources of error in the emittance results stem from possible uncertainty in the position of the quadrupoles and in the calibration of the quadrupole strength (as a function of driving current). Simulations with **elegant** of the effect of errors in quadrupole strengths and positions show that these are relatively minor issues:

1. Motion of Q4 by $\pm$ 5 mm produces less than $\pm$ 5 % variation in the inferred emittance for either plane.
2. Motion of Q5 by $\pm$ 5 mm produces less than $\pm$ 10 % variation in the inferred emittance for either plane.
3. $\pm$ 10 % strength errors in Q4 produces less than $\pm$ 10 % variation in the inferred emittance for either plane.
4. $\pm$ 10 % strength errors in Q5 produces less than $\pm$ 13 % variation in the inferred emittance for the x plane, and less than $\pm$ 8% variation for the inferred emittance for the y plane.

The effects of such errors are less than indicated by these results, since the positions are known to within a few mm and since the strength calibration is good to at least ±5%. (Recall, however, that the quadrupole calibration was adjusted by 8% as a result of beam-optics tests, as discussed in a previous section of this chapter.)

In Chapter 2, I discussed the strong effect that non-linear fields in the gun have on the beam emittance. It is possible that if the cathode is not properly positioned, the character of these fields could be changed, thus leading to discrepancies between simulation and experiment. In particular, the non-linear fields would be expected to increase rapidly as the cathode is pulled back into the ceramic annulus. This issue requires further investigation.

Another possible explanation for the discrepancies, and one that I consider likely, is the excitation of other structure modes in the gun, leading to beating of the fields in the cells. Such beating (occurring at 50 MHz) has been observed in actual operation. In addition, the beam-sizes for the measurements are from averages over hundreds of bunches, so that the variation in current and bunch momentum during the RF pulse will effect the measured beam-sizes. That is, the phosphor responds to all electrons that reach it, and during any RF pulse this includes electrons from bunches emitted over several hundred nano-seconds (just how long depends on the experimental conditions, in particular the setting of the alpha-scraper). Both of these effects would increase the measured emittance through variation of the beam momentum (which would lead to different focusing in the quadrupoles and alpha-magnet) and through variation of the RF focusing in the gun. I have not attempted to quantify these effects.

While the experiments disagree with the simulations, "bottom-line" is that the emittance, in the presence of chromatic aberrations and other errors, has been found to be less than $12\ \pi \cdot \text{mm} \cdot \text{mrad}$ in the horizontal plane, and less than $3.5\ \pi \cdot \text{mm} \cdot \text{mrad}$ for the vertical plane. Depending on the momentum filtration employed, substantially smaller emittances can be obtained. The geometric mean of these upper bounds is about $6.5\pi \cdot \text{mm} \cdot \text{mrad}$, corresponding to a normalized emittance of $32\ \pi \cdot m_e c \cdot \mu\text{m}$.

5.5 Bunch Length Measurement

In Chapter 4, I discussed the principle of magnetic compression, and made predictions of how effective the compression was expected to be. In particular, I predicted that bunch lengths of order 1 ps should be possible, with corresponding peak currents of several hundred amperes, depending on the current drawn from the gun. The SSRL preinjector consists of the GTL (described in an earlier section), followed by three SLAC-type linear accelerator sections and the linac-to-booster (LTB) transport line. The LTB features several bending magnets, the first of which (B1) is followed by an insertable phosphor screen ("the B1 screen"), allowing measurement of the momentum spread at the end of the linac with resolution as good as 0.02 %, depending on the adjustment of the quadrupoles between linac sections 1 and 2. (This resolution was determined empirically from the minimum beam-size achieved at the B1 screen.)

For the amount of RF power (2.6 MW) that was available for the gun as of this writing, the current reaching GT1 is limited to about 800 mA, with the momentum peak occurring at about 4 $m_e c$. This corresponds to $E_{p2} \approx 60 \text{MV/m}$ and hence an effective cathode area of about 5 mm^2, from which I infer that the current density is about 20 A/cm^2. Comparison with results of Chapter 4 would lead one to conclude that peak currents of about 200 A should be possible. However, the results of Chapter 4 were for $E_{p2} = 75 \text{MV/m}$, which results in a higher-momentum beam ($p_{peak} > 5$) that is easier to compress than the lower momentum beams obtained when the gun is heavily beam-loaded.

Routine operation of the SSRL preinjector involves operating the gun at full RF power and with GT1 currents of 600-800 mA. About 2/3 of this current is allowed out of the alpha-magnet, and the full momentum spread at the end of the linac is typically less than 0.2 %. The momentum of the beam at the end of the linac is typically 200 $m_e c$, giving $\Delta P \leq 0.4 m_e c$. This suggests that the bunch compression is working quite well, even for the initially low momentum beam from the fully beam-loaded gun. Ignoring the contribution of initial momentum spread and assuming $\beta = 1$ for the initial beam, I estimate the bunch length as 7 ps. This clearly over-estimates the bunch length, since it is known that the beam has significant initial momentum spread. I have not attempted to perform detailed simulations to investigate this

particular case further.

5.5.1 Principle of Bunch-Length Measurements

In order to get experimental verification of the bunch compression, I have made some bunch length measurements by employing the third accelerator section phased so that the center of the bunch receives no acceleration. Since the beam reaching the third section is already highly relativistic ($\approx 150 m_e c$), one can approximate the effect of the third section by

$$\Delta p = \Delta p_3 \sin(\phi), \tag{5.80}$$

where Δp_3 is the momentum change that would be imparted to a relativistic electron passing through section 3 at the accelerating crest, and δp is the momentum change imparted to a particle at phase ϕ (relative to the null). (This is very similar to the equations of the first section of Chapter 4 for acceleration of an already relativistic particle, except that in Chapter 4, I referenced the phase to the RF crest.)

If the phase-length of the bunch is $\delta\phi$, centered about $\phi = 0$, then the resultant momentum spread for an initially mono-energetic beam is

$$\delta p = 2\Delta p_3 \sin(\delta\phi/2) \tag{5.81}$$

If the beam is not initially mono-energetic, the analysis is more complicated, and may depend on the detailed initial phase-space distribution of the bunch. However, if momentum spread imparted to the beam is large compared to the initial momentum spread, such effects can be ignored. By measuring δp as a function of Δp_3, one can verify that the initial momentum spread is (or is not) inconsequential by verifying that δp depends (or does not depend) linearly on Δp_3.

Because of this issue, for the experiments I have done and will report on here, I first phased sections 1 and 2 to obtain the minimum final momentum spread before varying Δp_3. As seen in Chapter 4, this does not achieve the minimum bunch length, though it will produce a very short bunch. The bunch length also depends strongly on the alpha-magnet gradient, and hence it is necessary to select a gradient that is matched to the anticipated gun longitudinal phase-space.

It is convenient to rewrite equation (5.81) in terms of fraction momentum spread by dividing by the central momentum, p_o:

$$\frac{\delta p}{p_o} = 2\frac{\Delta p_3}{p_o}\sin(\delta\phi/2). \tag{5.82}$$

Δp_3 is varied, of course, by varying the RF power level in section 3. This is quantified (see Chapter 4) as

$$\Delta p_3 = 20.4\sqrt{P_3(\mathrm{MW})}. \tag{5.83}$$

where P_3 is the RF power to section 3.

5.5.2 Practical Considerations

While the principle of the measurement is simple, its execution is less so on the SSRL preinjector, due to a number of factors:

1. At the time of this writing, there are still problems with pulse-to-pulse jitter in the RF power level, which can cause the beam on the B1 screen to jitter by more than the width produced by the momentum spread. This problem was circumvented by using the LeCroy 9450 oscilloscope to grab frames from the TV camera, each of which contains the image produced by a single beam pulse.

2. There is no way to determine the RF phase of the individual sections, or to ensure that the phases are held constant during the experiment. Long-term drifts in the modulator HV can cause drifts in the individual RF power output levels and phases of the individual klystrons. The only solution to this at present is to do each measurement as quickly as possible.

3. The RF phases are referenced to the phase of the second section, as discussed earlier in this chapter. Hence, optimization for the best alpha-magnet gradient to compress a given beam from the gun would involve rephasing the RF for the gun and sections 1 and 3, which can be very time-consuming. Ideally, a single control would be available to simultaneously vary the RF of all three sections relative to the gun. (The reason there must be any RF re-phasing at all is that changing the alpha-magnet gradient varies the path-length, and hence the time-of-flight, between the gun and linac.)

4. The phase-shifter for section 3 has a significant phase-dependent insertion loss, so that phasing the section to the phase null can be difficult. To verify that the section is phased at the null, one must vary the power into the section and verify that the central beam momentum is unchanged. If the section is not at a null, the phase must be varied until the null is found. Unfortunately, upon changing the phase, the insertion loss changes, which can cause the beam to jump in momentum, requiring the experimenter to vary B1 in order to find the beam again. In addition, the phase shifter shows significant hysteresis, further complicating the optimization.

Time permitting, one would vary the gradient and rephase the accelerator sections, and measure the bunch-length as a function of gradient. There was insufficient experimental time to do this (the experiments took place during the commissioning of the 3 Gev Booster, which required dedicated operation of the linac for injection into the booster). The results I will report on here thus do not represent the optimized compression of the bunch, but only the best that I achieved in the few experiments that were conducted.

5.5.3 Experimental Results

Figure 5.39 shows the experimental data for the most optimum conditions achieved. Plotted here is the fractional momentum spread versus $\Delta p_3/p_o$, where p_3 is again the maximum available momentum gain from section 3, as deduced from RF power measurements. The three sets of data are for three different positions of the alpha-scraper, so that different fractions of the beam are let through for each set, as indicated by the value of Q printed with the data. The Figure also gives the bunch-lengths deduced from the data, using the fits shown with the data points. The momentum spreads plotted here are FWHM values, which, because of the shape of the momentum distribution at the end of the linac, contains more than 90% of the beam.

The peak-current, defined as

$$\hat{I} = \frac{Q}{\delta t}, \tag{5.84}$$

is a measure commonly employed for rating the suitability of a beam to applications requiring high-brightness. From the data given in the Figure, one sees that the highest

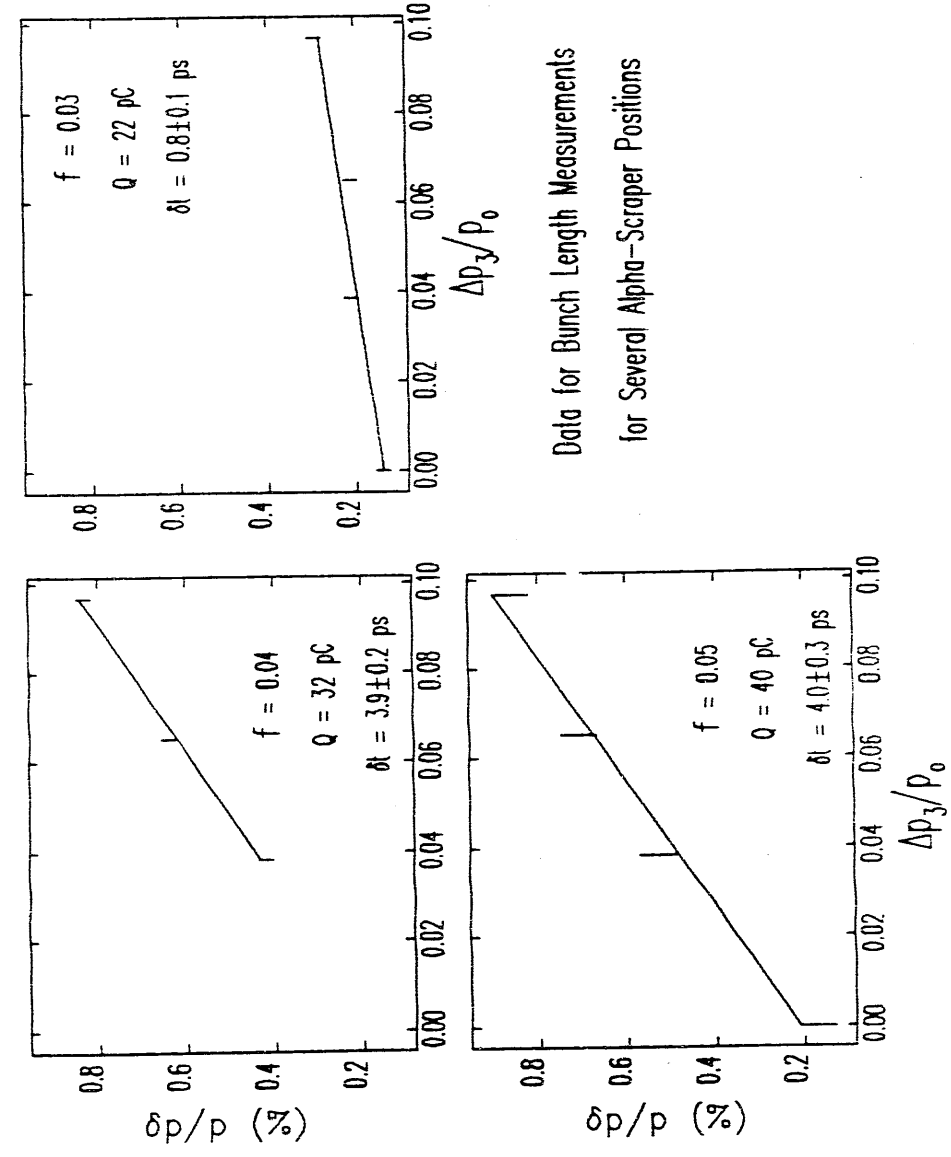

Figure 5.39: Data for Bunch-Length Measurements for Various Momentum Fractions Allowed Through the Alpha-Magnet

measured peak current is 28 ± 3 A. The current to GT1 in this case was 395 mA, which corresponds to 180 pC per bunch. Hence, in order to obtain this high peak-current, it was necessary to allow only 12% of the beam through the alpha magnet, corresponding to an initial momentum spread of ±3 % (i.e., f = 0.03).

The peak-currents for the other two data sets are about 10 A, which is a value that was reproduced in several other measurements as well. While these peak-currents, even the 28 A, are significantly less than the predicted optimum performance of the system, it is clear that too little work has been done to date to conclude that the predictions are inaccurate. It is hoped that others will pursue this work with additional experiments.

5.5.4 Comparison with Simulations

The peak momentum of the bunch from the gun, as measured with the alpha-magnet scraper, was 4.4 m_ec, which implies $E_{p2} \approx 65$MV/m. Beam-loss simulations using **elegant** and a MASK-generated initial particle distribution predict 85% transmission from the gun to GT1, implying that the current out of the gun was 464 mA. Since the effective cathode area for $E_{p2} = 65$MV/m is 4.7 mm^2, the current density is inferred to be about 10 A/cm^2. Hence, simulations of the bunch compression for comparison with experiment should use MASK-generated data for $E_{p2} = 65$MV/m and J = 10A/cm^2.

In order to ascertain how close the experimental results are to what could have been achieved, and also to see how sensitive the bunch-length is to the alpha-magnet gradient, I have used **elegant** to compute the bunch length and other parameters at the end of the first section for a series of alpha-magnet gradients, using the just-mentioned MASK data. Figure 5.40 shows the results for a momentum fraction of 0.03. The alpha-magnet gradient for the data shown in Figure 5.39 was 189 G/cm. The conclusion is that the compression is roughly consistent with simulations for f = 0.03. One sees from the Figure that the bunch length is quite sensitive to the gradient, and in particular that a 5% increase in the gradient from 190 G/cm would have decreased the bunch length by a factor of about 2.

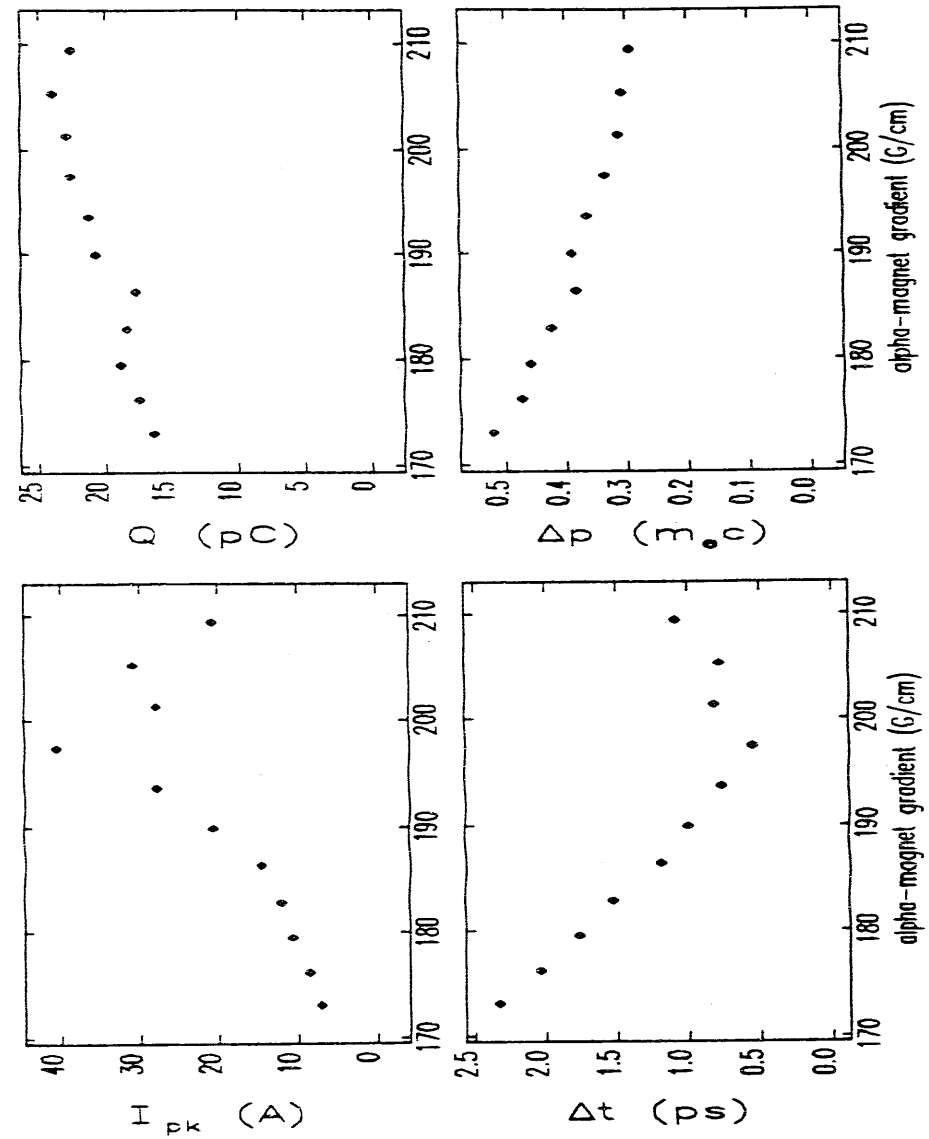

Figure 5.40: **elegant**/MASK-Simulated Beam Parameters After First Linac Section vs Alpha-Magnet Gradient for f=0.03, ($E_{p2} = 65$MV/m, J = 10 A/cm^2).

Appendix A

Computational Issues and Techniques

As the reader has no doubt noticed, I have made extensive use of computers in the research reported on in this thesis. I wrote much of the software that I used while at SSRL, and, since it has not been described in any publications, I am adding this appendix to acquaint the reader with some of this software. Much of it is of very general application, and this appendix will act as partial documentation. In the course of my descriptions, I will also comment on several issues pertinent to the use of computers in physics and to the state of the programs used by the accelerator physics community.

Part of my philosophy in doing research is that it is better to spend time writing a program that does a task than to do that task manually (using a hand-calculator, say), since the program will then be available to immediately solve the problem the next time. Though writing the program may take more time than solving the original problem, in the long run one gains in productivity through repeated use of the program. Because my goal in creating programs is to produce a tool for future use, I attempt to make each program solve a much more general form of my particular problem. In addition, I try to make the program as convenient and flexible as possible, to minimize the need for changes.

Some physicists have disdain for programming, an attitude I do not share. Programming is no different from solving differential equations, both being tools for solving problems. It would be as silly to look down on solving differential equations as it is to look down on programming. Computer programs differ from analytic approaches in that they are more easily misused and more easily used by those with little understanding, i.e., they are more easily used or misused by those who don't understand physics. This is not, however, an indictment of computer methods, but of those who abuse them.

One of the worst aspects of many of the programs used by accelerator physicists today is the user-interface. Many programs still require the user to type in rows of numbers, often in a rigid format, and still spit out row after row of numbers in tables spanning pages of fan-fold paper. These programs make little use of the computer's ability to talk to the user on *his* terms and to present the results in a way that aids the physicist's understanding. The user is left on his own to produce graphical or alternative representations of the results.

There are, of course, many programs that serve the physicist's needs much more satisfactorily. For the most part, these programs make use of a custom-designed user interface, so that the data from one program must be "post-processed" (e.g., plotted) with the aid of an auxiliary program written just for the physics program in question (or for some small group of programs). Examples of such codes are the standard version of MASK (which does all graphics internally and writes all other data in a single output file) and the POISSON group programs (which share a common, though incomplete, graphics program TEKPLOT, along with common pre-processors and some program-specific post-processors).

I consider this duplication of effort to be wasteful and to unnecessarily complicate the use of programs. In response to this problem, I have created three standardized data formats that are suitable for use in almost all applications. The simplest format—called `mpl` format after one of the primary programs that uses the format—is used by a large group of inter-communicating programs called the `mpl` Scientific Toolkit. The other two formats, the `dpp` and `column` formats, are self-describing data formats that are used to store and manipulate data, selected elements of which can then be written to `mpl` data sets for use with the Toolkit.

Physicists should not disdain to be involved in the design of the user interfaces for the programs they will use, for it is the requirements of doing physics that should dictate the design of the user interface. A program with an inflexible, inconvenient user interface forces the physicist to deal with tedious matters rather than the physics of the suituation; writing such a program is bad physics as well as bad computer science. I believe that the ideas described in this appendix, and the programs that implement them, represent a significant step toward a more universal user interface for scientific programs.

`mpl` and `awe` formats and the principle programs that use them are described in more detail in a following sections. For now, I will describe each briefly. `mpl` format is essentially a series of (x, y) pairs, with optional uncertainties, and hence is not appropriate for storing all types of data. For a great many applications, however, this is not a serious limitation, and the simplicity of the format makes it a good choice.

The most sophisticated format, `dpp` format, is used for the output from many simulation programs. It consists of a header that gives the names and units of the

quantities in the file, together with the names and units of "auxiliary quantities". After the header, there follows a series of "tables", each consisting of values for the auxiliary quantities and an arbitrary number of rows of data, with each column of the table containing values for one of the named quantities. (For example, the quantities might be phase-space coordinates (x, x', y, y'), with the auxiliary quantities being the time and z coordinate.) A sophisticated post-processing program, `dpp`, is available to access and manipulate the data in `dpp` format files (which are not human-readable), allowing the user to select the data of interest, create new quantities by defining them in terms of existing quantities, and much more. `dpp` can create human-readable tables and, in its most common application, `mpl` data sets for use with the `mpl` Toolkit.

`column` format is intermediate between `mpl` and `dpp` formats. Like `mpl` format, it is human-readable, but like `dpp` format, it has a self-describing data structure that supports any number of data elements in a table. The programs `column` and `sheet` (a spread-sheet program) are available to manipulate data in `column` format and to create `mpl` data sets of selected sub-sets of the data in a `column` data set. `column` format has the disadvantage that it is cumbersome and inefficient for very large data sets, which are best stored in `dpp` format.

At this time, the programs that make use of these formats are available only on SSRL's VAX 8810 computer. I anticipate making the first release of the `mpl` Toolkit, `dpp`, and other of my programs (such as `rfgun` and `elegant`) in the near future.

A.1 The mpl Scientific Toolkit

The mpl Scientific Toolkit is a group of programs, written in the C programming language [9], sharing a common data file format. The Toolkit is a powerful aid to a scientist's use of computers, alleviating much of the tedium associated with the analysis and interpretation of data. While I developed the Toolkit in the course of the research reported on in this thesis, the programs are of much more general application.

Each of the programs in the Toolkit was developed in response to a specific need in my research. However, my philosophy in implementing each program was to try to generalize my particular problem, then to write a program to deal with this more general problem. While this involves more effort at the outset, I have found that it produced great productivity benefits in the long run, since the Toolkit is now sufficiently complete that a great many tasks can be performed using some combination of programs from the Toolkit.

In order to enhance the productivity of the Toolkit, the programs not only use mpl format (described below) for their input data, but they also create new data sets in the same format, when it is at all meaningful to do so. Hence, one can perform a sequence of operations on one's data by invoking a sequence of Toolkit programs.

An alternative to the Toolkit concept would be a single, integrated program that incorporated all of the functions of the individual Toolkit programs. I believe such a system is unwise, in that it locks the user into the capabilities that are put into the program by the programmer. With the Toolkit concept, the user can readily write his own programs to augment the Toolkit as needed, without any change in the way he uses the member programs. In addition, individual Toolkit programs can be improved or replaced, and new Toolkit programs can be added, without altering other parts of the Toolkit or recompiling and redistributing a massive new version of the complete code. Finally, the user is not forced to abandon the system command environment, so he is left free to interleave Toolkit program calls with calls to system programs, as his needs dictate.

The user-interface for the programs is command-line based, rather than menu-based or query-based. That is, with some exceptions among older or more complicated

programs, the programs in the Toolkit are executed by typing the program name followed by a list of command-line options. This makes the programs more amenable to use in batch-files and allows argument substitution through the command language. It also caters to the experienced user and his productivity, at the expense of making things slightly harder on the novice.

An mpl data set is an ordinary text file, such as might be created using a text editor. The lines of the file are read and processed one at a time, and no line may be longer than 1024 characters. The first four lines are descriptive text (as opposed to numeric data). The first line should be the name of the abscissa (or "x" variable), while the second line should be the name of the ordinate (or "y" variable). The third and fourth lines will be placed at the bottom and top of the plot, respectively, when the data set is plotted with mpl or dpl.

The fifth line contains the number of data points in the data set. It should be equal to or larger than the number actually in the data set. A warning will appear if this number is different from the actual number of points. The program fixcount is available to count the points for you and put the right number on the fifth line of the data file.

Subsequent lines each contain one data point, which consists of two through four numbers. The first two values are x and y, respectively. If there are three values, the third is interpreted as the uncertainty (or sigma) for the y value. If there are four values, the third is interpreted as the x uncertainty, and the fourth as the y uncertainty. The number of values that are taken from each line is established by the number of values given for the first data point. The data set ends when the file ends, or when the number of points specified has been reached, whichever comes first.

The format of the data is very free. Any non-numeric ascii character (i.e., any character other than a digit, a plus or minus sign, or a period) can be used as a separator between data elements on a single line. Spaces and tabs are preferred, but solely for aesthetic reasons. Any line beginning with an exclamation point is treated as a comment, and ignored.

What follows are short descriptions of the programs in the mpl toolkit at this time. A few of these programs are for very specific applications, though most are of general application. Some are quite simple, a few are downright trivial, while others

are very sophisticated. What really matters is that the are all useful, easy to use, and interlinked by the common data format.

This list does not constitute a user's manual for these programs. Rather, the programs are intended to be self-documenting, since each program will respond with help information if it is run improperly. Hence, the user may obtain help for any program simply by running the program without any command-line input. The general form of the "usage message" that each program responds with is

```
usage: program_name required_argument1 required_argument2 ...
       [optional_argument_1] [optional_argument_2] ...
       {choice1.1 | choice1.2 | ...}
       {choice2.1 | choice2.2 | ...}
       [{optional_choice1.1 | optional_choice1.2 | ...}]
       [{optional_choice2.1 | optional_choice2.2 | ...}] .
```

In words: required arguments (e.g., the name of the input data set) are listed without delimiters. Optional arguments (usually a program-control "switch", or an optional output filename), are delimited by square brackets. Sets of arguments that the user *must* choose one and only one of are grouped by curly braces and separated by vertical bars. Optional sets of arguments that the user may choose one (and only one) of are grouped by curly braces inside square brackets. An example of a usage message is given below.

Options, or "switches", are of the general form

```
-keyword[=value1[,value2...]],
```

with several alternative forms recognized (you may use / instead of -, and : or , instead of =). Any keyword may be abbreviated when enter on the command line, so long as enough characters are supplied to may the keyword uniquely identifiable. If the value listed in the usage message has single quotes around it, then the value must be typed literally (i.e., as in `-average=rms`); you should not type the single quotes yourself in running the program. Such values may, like the keywords, be abbreviated. In contrast, a value listed in double quotes represents a string that *may* need to be enclosed double quotes (if, for example, it contains spaces, as in `-title="x vs y"`). Illustrations of the calling syntax appear below. Note that dpp and the other programs described in this appendix employ the same command-line interface.

In what follows, the phrase "data set" refers to data stored in an `mpl` format file, which consists of a set of data points (x_n, y_n) with optional uncertainties σ_{y_n} and σ_{x_n}. Unless otherwise stated, each program writes its results as a new data set in a file of the user's choice.

`add`: Adds (or subtracts) two data sets, with error propagation. Normally, the ordinates are added assuming that the abscissae match. Optionally, the abscissa can be added. There is also an option to match the abscissae before adding the ordinates, or vice-versa.

`average`: Performs averages of either column of data in a set of data sets, creating a new data set with the ordinates given by the averages, and the abscissae given either by the user or by incrementing a counter.

`column`: Extracts data from generic tables, such as might be output by any number of programs, and creates data sets. Hence, a program that prints rows of quantities in many columns can be interfaced to the Toolkit via `column`. (A better way to accomplish this is to modify the program to write its data in either `mpl` or `dpp` format.) Also, `column` extracts data `column`-format files into `mpl`-format files.

`combine`: Combines any number of data sets into a new data set, sorting the data sets by the first abscissa value in each set.

`conpair`: Performs the discretized convolution of a pair of data sets.

`conap`: Convolves a data set with a square aperture.

`deriv`: Takes the derivative of a data set, using

$$\left(\frac{\partial F}{\partial x}\right)_{x=x_a} \approx \frac{F(x_{n+m}) - F(x_{n-m})}{x_{n+m} - x_{n-m}}, \tag{1.1}$$

where $x_a = (x_{n+m} - x_{n-m})/2$ and where m is specified by the user. Strictly, the points should be equispaced (if not, use m=1).

`dpl`: Simple plotting program for a single `mpl`-format data set. See `mpl` for a more sophisticated program.

`ellipse`: Generates an `mpl`-format data set containing (x, y) points on an ellipse, given the Twiss parameters and the emittance.

`filter`: Performs digital filtering of a data set by doing an FFT, applying the filter, then doing an inverse FFT. Supports low-pass and high-pass filters with linear roll-off, as well as Parzen windowing[61] of the data.

`fixcount`: Fixes the point count of data sets. `fixcount` will attempt to simply overwrite the existing point count, if the fifth line of the data set has a sufficient number of characters to allow this without over-running onto the next line stored in the file. E.g., if the fifth line contains ''`1<return>`'' and the actual point count is 100, fixcount will have to re-write the entire file to make room for the two zeroes. If, however, the fifth line had contained ''`1 <return>`'', fixcount would over-write two of the trailing spaces with zeroes. This is not a trivial concern for data sets with thousands of points, and hence it is a good idea to put trailing spaces on the fifth line of each data set.

`fft`: Performs the FFT of a data set, producing a new data set with the magnitude vs. the frequency. Also does optional Parzen windowing[61] and provides output of real and imaginary components. Will work in either single or double precision, and will optionally pad or truncate your data to achieve 2^n data points (necessary for the FFT).

`fwhm`: Computes the full-width at half-maximum, with optional smoothing.

`gfit`: Fits a gaussian to a data set, finding the sigma, mean, baseline, and height, with the option of generating a new data set containing the fit evaluated at a series of points.

`ggen`: Generates a data set from evaluation of a gaussian at equispaced points.

`hist`: Makes histograms and cumulative distributions of one or more data sets, with optionally weighting of the histogrammed variable by the other variable.

`integ`: Computes the integral of a data set using the trapizoid rule with error propagation, creating a new data set with the integral as a function of the abscissa.

`interp`: Does polynomial interpolation on a data set, using whatever order of polynomial the user requests, with options for creating a new data set with interpolated values at equispaced points and for transforming one variable of a data set via interpolation on the first data set.

`lsf`: Does error-weighted least-squares polynomial fits to any order . Options for fitting only even or odd polynomials, or for fitting only specified powers. Also provides difference data sets and data sets from evaluation of fitted polynomials.

`mpl`: Versatile plotting program for multiple data sets. Allows zooming, point plotting, symbol plotting, and much more. Supports many common graphic output devices through use of SSRL's GHOST package. Has Greek and scientific character sets, with subscripting, superscripting, and in-line control of character attributes. The following "escape sequences" are recognized in character strings (these will be illustrated below):

1. $g, $r : switch to Greek or Roman characters.
2. $a, $b, $n: go to Above (superscript), Below (subscript), or Normal script.
3. $s, $e: Start and End Special (mathematical) symbols.
4. $i, $d: Increase (×1.5) or Decrease (÷1.5) character size.
5. $t, $f: go to Taller or Fatter letters (these are inverses of each other).
6. $u, $v: displace text vertically Upward or downward (respectively) by one-half character-height.
7. $h: back-space one-half character width.

(`mpl`-format data sets and these character sequences are also recognized by my `draw` program, which was used to create all but a few of the graphs in this thesis. `mpl` is actually called as a subroutine in `draw`.)

`mult`: Does pair-wise multiplication (or division) of the ordinates (or abscissae, or both) of two data sets, with error propagation. If the data sets are of unequal

length, the program will optionally try to line up the abscissae before multiplying.

merge: Merges two data sets, in the sense of taking the abscissae and ordinates of the new data set from the abscissae and ordinates of two other data sets, as specified by the user. (The strange spelling is to avoid confusion with the VAX VMS MERGE utility.)

params: Computes many parameters of a series of data sets, including positions, heights and sharpnesses of peaks, positions and values of minima and maxima, averages, and medians.

peakfind: Finds peaks in a data set, with optional smoothing and a user-defined threshold.

qsort: Sorts a data set into ascending (or descending order) by the abscissa, subsorting by the ordinate, with optional elimination of duplicate points.

rescale: Arguably the most powerful program in the toolkit. Performs very versatile transformations of a data set, including normalization, centering, scaling and offsetting, taking the logarithm, and more. Data sets may be sparsed, windowed, re-ordered, and the abscissae and ordinates may be swapped. Propagation of errors can be performed if sigmas are given in the data set. **rescale** will also accept user-defined transformations, specified as equations (this was done by incorporating my programming language calculator rpn into **rescale**, just as it was incorporated into dpp).

rndgen: Generates random-number pairs, with gaussian and uniform distributions supported. The random number generator uses a random shuffling routine with two linear congruential generators [61].

smooth: Smooths a data set by multi-pass averaging over adjacent points.

setlog: Performs set logic on two data sets. For example, **setlog** will find all the data points in one data set but not the other, or all the data points in both

sets. User-specified tolerances are accepted to determine when two points are "the same".

stats: Does statistical analyses of a data-set, computing various moments, widths, and cumulative distribution parameters. Also gives the minimum, maximum, and spread.

total: Creates a new data set with

$$\tilde{y}_n = \sum_{i=1}^{n} y_i \tag{1.2}$$

zerofind: Finds locations of zeros in a data set, with interpolation between points.

Next, I present a simple example session making use of some programs in the **mpl** Toolkit. In particular, this is the sequence I used to create the data for the comparison of linear and non-linear field terms in Chapter 2. Recall that I compared $E_z(z)$ (the on-axis longitudinal field) with

$$T_1(z) = \frac{R_c^2}{8}\left[\partial_z^2 - k^2\right]\frac{E_z(z)}{E_{p2}}, \tag{1.3}$$

where $R_c^2 = 0.3$ cm. The starting point is a data set containing $(z_i, E_{z,i})$ with z in meters, as previously extracted from a SUPERFISH-generated output file (which could be done using **column**). Lines preceded by exclamation points are comments. Lines in italic type are output from a program, while those in type-writer type are typed by the user:

```
! The initial data set is z_ez.sfish. Normalize it to a peak value of 1 and convert
! the abscissae from cm to m. The names of the data columns are also changed,
! using a subscript escape sequence to obtain E_norm. Note the use of the
! ampersand as a command-line continuation character:
$ rescale z_ez.sfish z_ez.out -normalize -x_divisor=100 &
$ -y_label="E$bnorm$n" -x_label="z (m)"
! Find out how to use the deriv program:
$ deriv
```

usage: deriv inputfile outputfile [-interval=number] [-x_sigma=value] [-y_sigma=value]

```
! Take its derivative, using a derivative interval of 2 points:
$ deriv z_ez.out z_dez1.out -interval=2
! Take the second derivative as well. Abbreviate the command somewhat:
$ deriv z_dez1.out z_dez2.out -int=2
! Compute k²E_z, where k = 2π/λ and λ = 0.105m.
! The y-axis label is changed to "k²E_norm(1/m²)".
$ rescale z_ez.out z_k2_ez.out -y_transform="y k2 *" &
-rpn_expression="2 pi * 0.105 / sqr sto k2" &
-y_label="k$a2$n E$bnorm$ (1/m$a2$n)"
! Add ∂²E_z and k²E_z
$ add z_dez2.out z_k2_ez.out sum.out
! multiply by R_c²/8, with abbreviation of command-line options
$ rescale sum.out t1_nonlin.out -y_t="y C *" &
-rpn_ex=".003 sqr 8 / sto C" &
-y_label="T$b1$n" &
! An alternative form of this command, that is quicker
$ rescale sum.out t1_nonlin.out -y_multiplier=1.125e-6 &
-y_label="T$b1$n"
! Compare E_z and T_1 by graphing on a VT100-series terminal (GHOST device 3),
! using different intensities for the two data sets.
$ mpl t1_nonlin.out,z_ez.out -multipen
! Repeat, but expand around the end of cell 1 (z = 0.026 m), with auto-scaling of
! the vertical extent of the plot.
$ mpl t1_nonlin.out,z_ez.out -multipen -scale=0.024,.028,0,0
```

Special note to FORTRAN *users:* VAX FORTRAN (and perhaps other implementations as well) uses, by default, the FORTRAN-carriage-control record structure for output files. C uses stream files, which by definition have no record structure. The VAX C library is supposed to provide translation between the different record formats in an invisible way, but does not do so properly for FORTRAN-carriage-control files. Hence, if you wish to generate mpl data sets from FORTRAN programs, you should

open the output file using the following syntax:

```
open(unit=unit_number, name=filename, status='new',
.     carriagecontrol='none', recordtype='stream_cr')
```

A.2 The awe Self-Describing Data Format

I created **awe** format from the realization that most output from scientific programs can be organized into tables. The columns of a table contain quantities having the same name and units, while each row contains quantities that are grouped together. For example, the columns might be z, E_x, E_r, and B_θ, so that a row would contain the latter three quantities as a function of the first. **awe** format consists of a header that gives the names and units of the columns, plus additional information (names and units of "auxiliary quantities") that pertains to all the data in the table. After the table-format is defined for a particular file, any number of tables in that format may follow, distinguished by a counter and (optionally) by the values of the auxiliary quantities. There are no internal limitations on the number of rows, columns, or tables that may be present in the data file; this is limited only by the memory capacity of the machine.

The program **awe** (for Dump Post-Processor) provides a universal user-interface to any program that creates **awe**-format output files. By using the unformatted representation of numbers, **awe**-format escapes the problem that many programs fall prey to, namely, the necessity of having to print fewer and fewer significant figures as more columns are added to the table, with the attendant problems whenever any numerical quantity exceeds the allowed format width. In addition, unformatted output is much faster and more efficient in terms of data storage, an important consideration for for programs (like MASK and **elegant**) that dump very large amounts of data. This has the disadvantage that **awe**-format files are not human-readable, as common text format output is. Hence, among other things, **awe** provides the user with the means to create readable, text files in very flexible ways. **awe** permits the user to create his own text files with tables of just the data he is interested in. **awe** also produces **mpl**-format files containing data of the user's choice; this is the most common use of **awe**, since it is by this route that graphical output can be obtained. **awe** performs a variety of data manipulations, such as selecting subsets of a data set, calculating new quantities using user-supplied equations, and taking averages, to name a few.

For example, as discussed in Chapter 2, MASK dumps the quantities $(r, z, \beta_r, \beta_z, t, \gamma)$ for any macro-electron that crosses certain user-defined z planes. If one wishes to plot

t vs. γ (or any other pair of quantities) the program **awe** will create an **mpl**-format file from MASK's **awe**-format output. This same **mpl**-format file can be used to create histograms, graphs, as input to statistical analysis programs, and so forth, using programs from the **mpl** Toolkit.

awe also allows one to define new quantities that are functions of the quantities stored in a **awe**-format file. Hence, if one was more interested in the phase than the time-of-exit, this could be defined in terms of t. New quantities (calculated from equations given by the user) are on an equal footing with the quantities stored in the file. Hence, the physicist does not need to pore over computer printouts, calculator in hand, in order to put a program's output into the form he needs, nor is it necessary to write a new program to perform this task. **awe** supports a full range of scientific functions (from simple arithmetic operations to integer-order Bessel functions and complete elliptic integrals), and will also accept user-defined functions stored in a function library.

I have implemented **awe**-format output for MASK, **rfgun**, PARMELA, SUPERFISH, POISSON, **elegant**, **emitmeas**, and **serrors**, among other programs. Using **awe**-format for MASK also allows direct use of MASK output in my tracking code **elegant** and several other codes. Since **rfgun** and my version of PARMELA (see chapter 3) both use **awe** format as well, it is possible to write post-processing procedures and programs that work equally well with data from any of these programs.

The following is the usage message printed by the current version (December 1990) of **awe**, using the same syntax conventions as was used for the **mpl** programs in the last section:

```
usage: awe data_file {output_file | outputfile_root}
    [-dumps=dump_numbers,...] [-separate_dumps] [-verbose]
    [-sample_interval=number] [-average_over_dump[={rms|sum}]]
    [-log_file=filename] [-format="C-style format string"]
    [-define_quantity=quantity_name,quantity_unit,"rpn-expression"]
    [-filter=quan_to_filter,lo_limit,up_limit,pair_or_seq_numbers_to_filter]
    [-sort=quan_to_sort_by,number_to_keep[,''descending'']]
    [-rpn_expression="rpn-expression-to-execute-first"]
    [-rpn_defns_file=filename]
```

```
{  -first=quantity_name[=alias],...
   -second=quantity_name[=alias],...
 | -sequence[=auxiliary_name[=alias]]
   -quantity=quantity_name[=alias],...
 | -contour_grid=min1,max1,min2,max2,n1,n2[,'combine'][,'simple']
   [-n_contours=number] -first=quantity_name
   -second=quantity_name -quantity=quantity_name[,...]
 | -table=filename,quantity_name[,quantity_name,...]
 | -list_quantities
}
```

Description of non-self-explanatory arguments:

- {outputfile | outputfile_root}: The name of an output file to be created, or a partial output filename from which output filenames will be constructed.
- [-dumps=dump_numbers,...] [-separate_dumps] [-verbose]: Allows specification of the "dumps" (or tables) from the file that are to be used, whether these dumps should be output to separate files, and whether informational messages will be printed for each dump.
- [-sample_interval=number]: Specifies sparsing of the data, so that only every number-th data point is used.
- -first=quantity_name[=alias] -second=quantity_name[=alias]: Specifies output of pairs of data (to the first and second columns of mpl-format files). The filename for each pair will be constructed from the outputfile_root and the names of the quantities (or their aliases, if any).
- -sequence[=auxiliary_name[=alias]] -quantity=quantity_name[=alias]: Specifies output to mpl-format files of data as a function of an auxiliary variable or the dump "time" counter (which is a mandatory auxiliary quantity, and a left-over from the early days of awe). The filename for each pair will be constructed from the outputfile_root and the names of the quantities (or their aliases, if any).

- -contour_grid=min1,max1,min2,max2,n1,n2[,'combine'][,'simple'] : Specifies that the user wishes to create data suitable for input to the contour program for contour plotting, computation of path integrals, etc. The -first and -second arguments give the names of the independent variables (call them (x1, x2)) of the contour map, with the -quantity arguments giving the names of any number of dependent variables (for separate output). The items for use with the -contour_grid keyword are the extent of the contour map in the x1 and x2, and the number of grid points in each dimension. The 'combine' and 'simple' options specify the type of data interpretation. If 'combine' is given, then data from all dumps is combined before processing; otherwise, dumps are processed separately. If 'simple' is given, then the program assumes that (x1, x2) pairs lie on the grid specified by the arguments, which reduces the processing to simply putting each data point in array location of the closest grid point to any given (x1, x2) pair. If 'simple' is not given, then the program interpolate in two dimensions to find the value of the dependent variables at each grid point, which is a much more time-consuming process, but necessary for processing data from programs that use irregular meshes (like SUPERFISH). This latter procedure is not entirely reliable.
- -table=filename,quantity_name[,quantity_name...] : Specifies creation of a text table in the named file, in column format (which is human-readable).

As a simple example of the use of awe, suppose that one has just run elegant with a variable quadrupole strength in order to do a simulated emittance measurement (see Chapter 5). The beam-size and quadrupole strength, along with many other parameters, are stored in the "final parameters" output file, which I will refer to as emitmeas.fin in the example. I also show the use of mpl to plot the data:

```
! Extract the horizontal RMS beam-size (called "Sx" by elegant) vs the quadrupole
! strength (called "Q4[K1]"):
$ awe emitmeas.fin em -first:"Q4[K1]=K1" -second="Sx"
file opened: em_K1_Sx.out
! Plot this data using symbol 28, with points connected by lines:
```

```
$ mpl em_k1_sx.out -plot_code=0,28 -connect
! Extract the emittance for both planes as a function of Q4 strength. elegant
! uses "ex" and "ey" for the emittances in the x and y planes, respectively
$ awe emitmeas.fin em -first:"Q4[K1]=K1","Q4[K1]" -second="ex","ey"
file opened: em_K1_ex.out
file opened: em_K1_ey.out
! Plot the emittances for both planes, using different symbols:
$ mpl em_k1_ex.out,em_k1_ey.out -plot_code=0,28,1,29 -connect
! Extract the geometric mean of the emittances vs Q4 strength:
$ awe emitmeas.fin em -first:"Q4[K1]" -second=eg &
-define=eg,m-rad,"ex ey * sqrt"
1 new quantity was defined:
    eg (m-rad): ex ey * sqrt
file opened: em_K1_eg.out
! Plot this along with x and y emittances, changing plotter pens between data sets.
$ mpl em_K1_e%.out -multi_pen
! Make a table containing the quadrupole strength, horizontal beta function, and
! the elements of the horizontal transport matrix, R, along with 1 - det(R):
$ awe emitmeas.fin &
-define=odetR," ","1 R11 R22 * R12 R21 * - -" &
-define=betax,m,"Sx 2 pow ex /" &
-table=horiz.tab,Q4[K1],Sx,R11,R12,R21,R22,betax,odetR
2 new quantities were defined:
    odetR ( ): 1 R11 R22 * R12 R21 * - -
    betax (m): Sx 2 pow ex /
```

The file `emitmeas.fin` contains 46 quantities as a function of the quadrupole strength, including beam-sizes, emittances, transport-matrix elements, and more. I leave it to the reader to imagine how unwieldy it would be to have these quantities printed to a text file in the fashion typical of physics programs. With **awe**, the user selects just the data he is interested in, and manipulates it into the desired form with a minimum of effort. This would be powerful enough as the interface to a single physics program (**elegant**, say), but it is all the more powerful in providing a uniform

interface to many different programs, as well as an interface *between* different physics programs (e.g., between MASK and **elegant**, or **elegant** and **emitmeas**).

Note that since **awe** format is self-describing, upgrades to **elegant** that result in more data being sent to the "final parameters" file (or any other file) will not change the user interface, beyond making more data available to the user via **awe**. (In fact, the final parameters file contains a variable number of quantities, depending on how many quantities the user chooses to vary in his **elegant** simulation.) This applies to the interface between programs as well, since a program that accepts **awe**-format input obtains its data by "asking for it" by name. For example, the program **emitmeas** (see Chapter 5) finds the data for the R-matrix and beam-sizes by finding quantities with the appropriate names (the same names used in the examples above). A set of subroutines is available for use in any program that needs to utilize data stored in **awe** format.

I will not give a detailed description (i.e., at the byte level) of **awe** format here beyond what has been given above, since the format is going to undergo modifications in the near future to further improve its functionality. The upcoming changes will allow further expansions of the format without invalidating data files stored in a previous version of the format, by the addition of a format version number. This will be done before any release of the program is made, to forestall problems with multiple versions that use different formats. Also, the format will be expanded to include text descriptions for each variable, format-specifiers for each variable (to specify how the data should be printed when ascii output is requested), and text descriptions of each individual table in a file (in addition to the existing text description of the file as a whole), in addition to other changes.

This issue of changing formats should not be confused with the issue of whether **awe** can accept data tables other than those it is current working with. There is no limitation (except at the hardware level) on the creation of new data tables in the existing **awe** format. Any new application can describe a new data-table in the format **awe** currently accepts . The upcoming changes in **awe** format refer to improvements in the way programs can describe their data to **awe** or to other programs that call the **awe** subroutine library to read **awe** format files.

Bibliography

[1] Y.Minowa (Mitsubishi Electric). Japanese Government Patent, Number SHO-WA58-20180, 1983. (Applied for: 1974, Public Notice: 1975).

[2] G. A. Westenskow and J. M. J. Madey. Microwave electron gun. *Laser and Particle Beams*, 2:223–225, 1984.

[3] H. Goldstein. *Classical Mechanics*. Addison Wesley, Reading, Massachusetts, second edition, 1981.

[4] M. Weiss. A Short Demonstration of Liouville's Theorem. In *Proceedings of the 1986 CERN Accelerator School*, July 1987.

[5] L. D. Landau, E. M. Lifshitz. *Statistical Physics, Part 1*. Pergammon Press, Oxford, third edition, 1980.

[6] H. Wiedemann. Physics of Particle Accelerators. Lecture notes of course given in Fall 1987 in the Stanford University Applied Physics Department, 1987.

[7] M. Reiser. Theory and Design of Charged Particle Beams. Draft of a forthcoming book, Spring, 1989.

[8] M. Sands. The Physics of Electron Storage Rings. Technical Report SLAC-121, SLAC, November 1970.

[9] E. D. Courant and H. S. Snyder. Theory of the Alternating Gradient Synchrotron. *Annals of Physics*, 3:1–48, 1958.

[10] K. L. Brown. A First- and Second-Order Matrix Theory for the Design of Beam Transport Systems and Charged Particle Spectrometers. SLAC Report SLAC-75, SLAC, June 1982.

[11] K. T. McDonald and D. P. Russell. Methods of emittance measurement. Technical Report DOE/ER/3072-51, Princeton University, March 7 1989.

[12] B. E. Carlsten and R. L. Sheffield. Photoelectric Injector Design Considerations. In *Proceedings of the 1988 Linear Accelerator Conference*, October 1988.

[13] R.L.Sheffield. Photocathode RF Guns. In *Physics of Particle Accelerators, AIP Conference Proceedings* 184, 1987-1988.

[14] C. Travier. RF Guns: A Review. RFG Note SERA/90-219/RFG, Laboratoire De L'Accelerateur Lineaire, August 1990.

[15] J. C. Gallardo and R. B. Palmer. Preliminary Study of Gun Emittance Correction. *IEEE Journal of Quantum Electronics*, 26(8):1328–1331, 1990.

[16] M. Reiser. Physics of High-Brightness Sources. In *Proceedings of the 1989 Workshop on Advanced Accelerator Concepts*, January 1989.

[17] R. D. Ruth. Report on the International Workshop on Next Generation Linear Colliders. In *Proceedings of the IEEE Particle Accelerator Conference*, pages 716–720, March 1989.

[18] P. B. Wilson. Wakefield Acceleration: Concepts and Machines. In *High-Brightness Accelerators*, pages 129–155. NATO Advanced Study Institute, 1986.

[19] A. M. Sessler and D. B. Hopkins. The Two-Beam Accelerator. In *1986 Linear Accelerator Conference Proceedings*, pages 385–390, 1986.

[20] R. H. Pantell. Free-Electron Lasers. In *Physics of Particle Accelerators*, volume 184 of *AIP Conference Proceedings*, pages 1707–1728, 1989.

[21] C. A. Brau. Free-Electron Lasers. In *Physics of Particle Accelerators*, volume 184 of *AIP Conference Proceedings*, pages 1615–1706, 1989.

[22] A. Hofmann. Theory of Synchrotron Radiation. ACD Note 38. SSRL. September 1986.

[23] S. Krinsky. Undulators as Sources of Synchrotron Radiation. *IEEE Transactions on Nuclear Science*. NS-30(4):3078-3082. 1983.

[24] K.-J. Kim. Angular Distribution of Undulator Power for an Arbitrary Deflection Parameter K. *Nuclear Instruments and Methods*. A246:67-70. 1986.

[25] H. Motz. Applications of the Radiation from Fast Electron Beams. *Journal of Applied Physics*. 22(5):527-535. 1951.

[26] H.Wiedemann. *et. al.* 3 GeV SPEAR Injector Design Handbook. Technical report. SSRL. March 1988.

[27] Kwang-Je Kim. Brightness and Coherence of Radiation from Undulators and High-Gain Free Electron Lasers. In *IEEE Particle Accelerator Conference Proceedings*. 1987.

[28] M. Borland. Calculating Emittance from Images of Undulator Radiation. In *IEEE Particle Accelerator Conference Proceedings*. 1989.

[29] L. Mandel and E. Wolf. Coherence Properties of Optical Fields. *Reviews of Modern Physics*. 37(2):231-284. 1965.

[30] E. Hecht and A. Zajac. *Optics*. Addison-Wesley. Reading. Mass.. 1979.

[31] J. D. Jackson. *Classical Electrodynamics*. J.Wiley. New York. second edition. 1975.

[32] G. A. Westenskow. J. M. J. Madey. L. C. Vintro. S. V. Benson. Owner's Manual for the Microwave Electron Gun. HEPL Technical Note TN-86-1. HEPL. February 1986.

[33] W. C. Turner. *et. al.* High-Brightness. High-Current Density Cathode for Induction Linac FELs. In *Proceedings of the 1988 Linear Accelerator Conference*. October 1988.

[34] J.S.Fraser. R.L.Sheffield. E.R. Gray. G. W. Rodenz. High-Brightness Photoemitter Injector for Electron Accelerators. In *IEEE Particle Accelerator Conference Proceedings*. 1985.

[35] B.E.Carlsten and R. L. Sheffield. Photoelectric Injector Design Considerations. In *Proceedings of the Linear Accelerator Conference*. 1988.

[36] K. T. McDonald. Design of the Laser-Driven RF Electron Gun for the BNL Accelerator Test Facility. *IEEE Transactions on Electron Devices*. 35(11). 1988.

[37] J.S.Fraser. R.L.Sheffield. E.R. Gray. P.M.Giles. R.W.Springer. V.A.Loebs. Photocathodes in Accelerator Applications. In *IEEE Particle Accelerator Conference Proceedings*. 1987.

[38] P. E. Oettinger. R. E. Shefer. D. L. Birx. M. C. Green. Photoelectron Sources: Selection and Analysis. *Nuclear Instruments and Methods*. A272:264-267. 1988.

[39] S. V. Benson. J. Schultz. B. A. Hooper. R. Crane. J.M.J.Madey. Status Report on the Stanford Mark III Infrared Free Electron Laser. *Nuclear Instruments and Methods*. A272:22-28. 1988.

[40] J.S.Fraser. R.L.Sheffield. E.R. Gray. P.M.Giles. R.W.Springer. V.A.Loebs. Photocathodes in Accelerator Applications. In *IEEE Particle Accelerator Conference Proceedings*. 1987.

[41] M. B. James. Production of High-Intensity Electron Bunches for the SLAC Linear Collider. SLAC Report SLAC-Report-319. SLAC. August 1987.

[42] S. P. Kapitza and V. N. Melekhin. *The Microtron*. harwood. London. 1969.

[43] R. B. Neal. A High Energy Linear Electron Accelerator. M. L. Report 185. Stanford Microwave Laboratory. February 1953.

[44] E. Tanabe. Private communication.

[45] H.A.Enge. Achromatic Magnetic Mirror for Ion Beams. *The Review of Scientific Instruments*. 34(4). 1963.

[46] T. I. Smith. Intense Low-Emittance Linac Beams for Free Electron Lasers. In *Proceedings of the Linear Accelerator Conference*, pages 421–425, June 1986.

[47] G. A. Loew, R. H. Miller, C. K. Sinclair. The SLAC Low-Emittance Accelerator Test Facility. In *Proceedings of the Linear Accelerator Conference*, pages 144–147, June 1986.

[48] R. B. Neal, editor. *The Stanford Two-Mile Accelerator. R.B.Neal. editor.* W.A.Benjamin Inc., 1968.

[49] M. Borland. elegant: A Matrix Integrating Code for Third-Order Tracking. ACD-Note, SSRL. To be published.

[50] M. Borland. `rfgun`: A Program to Simulate RF Guns. ACD-Note, SSRL. To be published.

[51] A.T.Drobot, *et. al.* Numerical Simulation of High Power Microwave Sources. *IEEE Trans.*, 32:2733–7, 1985.

[52] K.Halbach, R.F.Holsinger. SUPERFISH—A Computer Program for Evaluation of RF Cavities with Cylindrical Symmetry. *Particle Accelerators*, 7:213–222, 1976.

[53] E.Tanabe, M.Borland, M.C.Green, R.H.Miller, L.V.Nelson, J.N.Weaver, H.Wiedemann. A 2-MeV Microwave Thermionic Gun. SLAC Publication SLAC-PUB-5054, SLAC, August 1989.

[54] M.Borland, M.C.Green, R.H.Miller, L.V.Nelson, E.Tanabe, J.N.Weaver, H.Wiedemann. Performance of the 2 MeV Microwave Gun for the SSRL 150 MeV Linac. In *Proceedings of the Linear Accelerator Conference*, September 1990.

[55] G. A. Loew, R.Talman. Elementary Principles of Linear Accelerators. SLAC Publication SLAC-PUB-3221, SLAC, September 1983.

[56] P. B. Wilson. High-Energy Electron Linacs: Applications to Storage Ring RF Systems and Linear Colliders. SLAC Publication SLAC-PUB-2884, SLAC, February 1982.

[57] J. N. Weaver. Formulas for Measuring Q_o, β, E_{max}, and B_{max} in a Superconducting Cavity. Technical Report TN-68-21, SLAC, July 1968.

[58] G. A. Loew. Charging and Discharging Superconducting Cavities. Technical Report TN-68-25, SLAC, November 1968.

[59] M.Bartsch, *et. al.* MAFIA Release 3.X. In *Proceedings of the Linear Accelerator Conference*, Albuquerque, NM, 1990.

[60] J.Gao. The Precise Measurement of Electric and Magnetic Fields in a Resonant Cavity. In *Proceedings of the Linear Accelerator Conference*, Albuquerque, NM, 1990.

[61] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Pres, Cambridge, 1988.

[62] C. Kittel and H. Kroemer. *Thermal Physics*. W.H.Freeman, San Francisco, second edition, 1980.

[63] R.H.Miller. Private communication.

[64] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washingston, D.C., 1964.

[65] U. Laustroer, U. Van Rienen, T. Weiland. URMEL and URMEL-T User Guide. Technical Report DESY-M-83-03, DESY, February 1987.

[66] K. Halbach. A Program for the Inversion of System Analysis and Its Application to the Design of Magnets. In *Proceedings of the 2nd Conference on Magnet Technology*, Oxford, England, 1967.

[67] K. G. Steffen. *High Energy Beam Optics*. Wiley, New York, 1965.

[68] K. L. Brown, F. Rothacker, D.C.Carey, Ch.Iselin. TRANSPORT: A Computer Program for Designing Charged Particle Beam Transport Systems. SLAC Report SLAC-91, SLAC, May 1977.

[69] R.H.Helm, R.H.Miller. *The Stanford Two-Mile Accelerator*, R.B.Neal, editor. W.A.Benjamin Inc., 1968.

[70] H.Wiedemann, *et. al.* 3 GeV SPEAR Injector Design Handbook. Technical report, SSRL, March 1988.

[71] F. C. Iselin, J. Niederer. The MAD Program (Methodical Accelerator Design) Version 7.2 User's Reference Manual. Technical Report CERN LEP-TH 88-38, CERN, July 1988.

[72] B.E.Carlsten, B.D.McVey, E.M.Svaton, G.R.Magelssen, L.M.Young. Magnetic Bunchers for the Generation of High Peak Current, Low Emittance Electron Pulses at Medium Energy. In *Proceedings of the Linear Accelerator Conference*, September 1990.

[73] R. I. Cutler, E. R. Lindstrom, S. Penner. Conceptual Design of a High Current Injector for the NIST-NRL Free Electron Laser. In *Proceedings of the IEEE Particle Accelerator Conference*, pages 328–330, March 1989.

[74] C. Bourat, G. Meyrand, D. T. Tran, D. Tronc, A. Massarotti, A. Wrulich. The 100 MeV Preinjector for the Trieste Synchrotron. In *Proceedings of the IEEE Particle Accelerator Conference*, pages 935–937, March 1989.

[75] U. Bizzarri, *et. al.* Status of the Free Electron Laser Experiment at the ENEA Frascati Center. *Nuclear Instruments and Methods*, A237:213–219, 1985.

[76] J. Rees. A Progress Report on the SLAC Linear Collider. In *Proceedings of the Linear Accelerator Conference*, pages 209–213, June 1986.

[77] M. Borland, J.N. Weaver, R. Anderson, M. Baltay, L. Emery, A.S. Fisher, R. Hettel, R.H. Miller, H. Morales, J. Sebek, B. Youngman. Design and Performance of the Traveling-Wave Beam Chopper for the SSRL Injector. In *To be published in IEEE Particle Accelerator Conference Proceedings*, San Francisco, CA, 1991.

[78] J. Sebek, M. Baltay, M. Borland, J. Cerino, L. Emery, R. Hettel, H. Morales, D. Mostowfi, J.-L. Pellegrin, M. Rowen, J. Safranek, V. Smith, J. Voss, J.N. Weaver, H. Wiedemann, Y. Yin, B. Youngman. Diagnostic Instrumentation for the SSRL 3 GeV Injector for SPEAR. In *To be published in IEEE Particle Accelerator Conference Proceedings*, San Francisco, CA, 1991.

[79] J.N Weaver, S. Baird, M. Baltay, M. Borland, C. Chavis, L. Emery, R. Hettel, S.J. Mello, R.H. Miller, H. Morales, H.-D. Nuhn, J. Safranek, J. Sebek, J. Voss, D. Wang, H. Wiedemann, B. Youngman. The Linac and Booster RF System for a Dedicated Injector for SPEAR. In *To be published in IEEE Particle Accelerator Conference Proceedings*, San Francisco, CA, 1991.

[80] J. Sebek. Private communication.

[81] E. Benson, J. N Weaver. Private communication.

[82] R. L. Sheffield. Private communication.

[83] M. C. Ross, N. Phinney, G. Quickfall, H. Shoaee, J. C. Sheppard. Automated Emittance Measurements in the SLC. SLAC Publication SLAC-PUB-4278, SLAC, March 1987.

[84] K. L. Brown, R. V. Servranckx. First- and Second-Order Charged Particle Optics. SLAC Publication SLAC-PUB-3381, SLAC, July 1984.

[85] P. R. Bevington. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, New York, 1969.

[86] M. C. Ross, J. T. Seeman, R. K. Jobe, J. C. Sheppard, R. F. Stiening. High Resolution Beam Profile Monitors in the SLC. *IEEE Transactions on Nuclear Science*, NS-32:2003–2005, 1985.

[87] K. D. Jacobs, J. B. Flanz, T. Russ. Emittance Measurements at the Bates Linac. In *IEEE Particle Accelerator Conference Proceedings*, Chicago, IL, 1989.

[88] M. C. Ross. Using Television Cameras to Measure Emittance. Single Pass Collider Memo CN-280, SLAC, September 1984.

[89] S. Yencho and D. R. Walz. A High-Resolution Phosphor Screen Beam Profile Monitor. *IEEE Trans.*, 32:2009–2011, 1985.

90 J. T. Seeman, V. Luth, M. Ross, J. Sheppard. "Beam Tests of Phosphorescent Screens. Single Pass Collider Memo CN-290. SLAC. March 1985.

91 B. W. Kernighan and D. M. Ritchie. *The C Programming Language*. Prentice Hall. Englewood Cliffs. N.J., second edition, 1988.

Errata
19 November 1991

Errata are listed by section number.

- 3.3.2
  - I've used the prime symbol inconsistently in this section. In particular, x′ and y′ are slopes relative to the central trajectory, while q'_i ($i = 1,2,3$) are velocities (normalized to βc).
  - In equations 3.64 and 3.75, the fourth entry of the right hand vector should be $\frac{q'_3}{\sqrt{1-q'^2_1-q'^2_2}}$.
  - Just before equations 3.66, 3.71, and 3.76, replace the word "slopes" by "velocities".
  - Equation 3.67 should have a minus sign in front of the square root sign.
  - In equation 3.70, replace q'_3 by q'_1/q'_2.
- 3.3.3
  - In equation 3.82, the second entry of the right hand vector should be Q'_1/Q'_2, and the fourth entry should be Q'_3/Q'_2.
- 3.5.2
  - In Table 3.5, the following matrix elements should be changed to the values given:
    $T_{1122} = 7.202$
    $T_{3412} = -2.728 \cdot 10^{-3}$
    $T_{3444} = -1.113 \cdot 10^{2}$
    $T_{4122} = 1.282 \cdot 10^{1}$
    $T_{4333} = -4.324 \cdot 10^{-1}$
    $T_{4432} = -1.439 \cdot 10^{-3}$
    $T_{4433} = -1.445$
    $T_{4443} = 8.683 \cdot 10^{1}$
    $T_{4444} = -3.154 \cdot 10^{2}$
  - In Table 3.7, the following matrix elements should be changed to the values given:
    $T_{3422} = -6.979 \cdot 10^{-1}$
    $T_{3444} = 3.661 \cdot 10^{-1}$
    $T_{4311} = -1.491$
    $T_{4321} = -6.846$
    $T_{4322} = -6.318$
    $T_{4333} = 2.389 \cdot 10^{-1}$
    $T_{4411} = 6.622$
    $T_{4421} = 1.452 \cdot 10^{1}$
    $T_{4422} = 6.[illegible]$
    $T_{4433} = -5.745 \cdot 10^{-1}$
    $T_{4443} = -1.683$
    $T_{4444} = 2.312$
    $T_{4641} = 1.269$

– In Table 3.9, the following matrix elements should be changed to the values given:

$U_{1641} = 6.914 \cdot 10^{-4}$
$U_{3422} = -1.376$
$U_{3444} = -3.287$
$U_{4311} = -1.021 \cdot 10^{1}$
$U_{4333} = -3.022$
$U_{4411} = -1.434 \cdot 10^{1}$
$U_{4422} = -1.762$
$U_{4431} = -1.087 \cdot 10^{-3}$
$U_{4433} = -1.321 \cdot 10^{1}$
$U_{4444} = -8.426$
$U_{4631} = -2.404 \cdot 10^{1}$
$U_{4641} = -2.923 \cdot 10^{1}$
$U_{4643} = -1.008 \cdot 10^{-3}$
$U_{4663} = -1.481 \cdot 10^{1}$
$U_{4664} = -1.615 \cdot 10^{1}$