SLAC-R-799

Longitudinal Phase Space Measurements and Application to Beam-Plasma Physics

C. Barnes

Stanford Linear Accelerator Center
Stanford University
Stanford, CA 94309

SLAC-Report-799

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

LONGITUDINAL PHASE SPACE MEASUREMENTS AND APPLICATION TO BEAM-PLASMA PHYSICS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Christopher Dwight Barnes
December 2005

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Robert H. Siemann
(Principal Advisor)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Todd I. Smith

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Alexander W. Chao

Approved for the University Committee on Graduate Studies.

Abstract

Beam driven plasma wakes show great promise for meter scale accelerators with high gradients. Plasma wakefield theory indicates that the achievable gradient is proportional to N/σ_z^2, and the bunches as short as $12\,\mu\text{m} \approx 40\,\text{fsec}$ in RMS length which are now possible at the Stanford Linear Accelerator Center (SLAC) are predicted to allow gradients in the tens to hundreds of GeV/m. We discuss the three stages of compression needed to achieve such short bunches.

No technique currently available can measure these longitudinal profiles directly shot by shot, requiring an indirect method. We added a magnetic chicane near the end of SLAC's 3 km main accelerator to measure the energy spread of each bunch in a non-destructive manner. Additionally, we performed a series of detailed simulations of the main accelerator in LiTrack, a code developed at SLAC. By comparing each measured spectrum against the library of spectra from simulations, we can find the best match to determine the input conditions to the accelerator and the total longitudinal phase space of every shot in the machine.

We discuss several methods employed to verify that the longitudinal profiles coming from simulations are accurate. We can use this information to understand which particles are accelerated in each bunch, and by how much. Additionally, we use the longitudinal information to choose a subset of shots that always have the same incoming profiles to see the differing acceleration experienced by those shots as we vary the plasma density and length. This allows a more robust calculation of achieved gradient, as well as illuminating the effect of transverse deflections on that acceleration.

Finally, we discuss other applications, as the technique for measuring the energy spectra and for matching to simulations is quite general.

Acknowledgements

The completion of one's thesis is a good time to reflect on all of the people that have made the journey not only possible, but worthwhile. My (eight!) years in graduate school have been great, and I would first like to thank Leland and Jane Stanford for placing their university in such an idyllic setting. Thanks also to Stanford University itself for becoming such a center of scientific learning - and for admitting me.

Keeping me sane outside of work were many friends and family; being part of a support network is critical in surviving the inevitable tough times. I can always talk to my parents Linda and Virgil, and my brothers Jeff and Dan. Their insight, caring and the many wide-ranging conversations with them have been central to my success.

I have been fortunate to know many great people during graduate school. Here is a partial list, in no particular order. Alex Panchula and I did homework together our first year and still love to discuss anything technical while on bike rides. I've learned a lot of things from him and can't wait to see what he does. Hiro Tanaka, Jeremy Bricker and Clarence Chang are physics friends whose intelligence and relaxed demeanor are a pleasure to be around. Erica Mueller and Christine Moyle take me outside of physics and give me perspective. I met James Geraci of MIT while spending the summer in Japan in 2000, and hope we'll get to live closer to one another in future. *Itsudemo tomodachi.* Nick Leavy has been a friend for 14 years, a roommate, my inspiration to take up triathlon and generally a deeply knowledgeable person I hope to know always. Joel Brown, Brucek Khailany, Eric Rossetter, Greg DeFouw, Scott Krueger, and Chris Stolte have been friends and roommates since early in grad school. Knowing them has been the cornerstone of my graduate school experience.

I am not married, but have had the good fortune to date several amazing women over the years, despite the numerical handicap of living in Silicon Valley during the height of the tech boom! Their influence has made me a better person.

Several activities gave structure to my time here, first was Shotokan Karate, where I progressed to first degree brown belt. I plan to get my black belt someday, I promise! The triathlon team has been a joy the last year and a half. I'm only sorry I didn't join earlier, but plan to continue training. Being a tour guide at SLAC was also very rewarding. Thanks to Emily, Barb, Neil, Maura, Elizabeth and Melinda.

At SLAC, I have had the opportunity to work with a wide variety of fascinating people. I would like to thank my entire research group and colleagues. In administration, Angie Seymour and my friend Stephanie Santo have made life infinitely easier. Their skills with the practical aspects of getting things done in a big laboratory have saved my hide on a number of occasions.

SLAC has a deep talent pool and I wish to thank the technical staff and all of ARDB and ARDA for their shared expertise. A very talented peer group is what makes SLAC so special. In particular, I would like to thank Paul Emma and Karl Bane for writing LiTrack and for many stimulating discussions. Also, Alex Chao has been very helpful, not least by reading my thesis.

My graduate career has been more varied than some because I had the opportunity to spend significant time on two separate experiments where I met great people. First was the group focusing on the LEAP experiment: Ben Cowan, Jim Spencer, Tomas Plettner, Bob Byer and Eric Colby, all of whom have been inspirational. At HEPL, I got to work with Todd Smith, who kept our experiment going on several occasions and has kindly agreed to read this thesis. Thanks!

After switching to the E164 experiment, I had the chance to interact with another large group of talented people: Tom Katsouleas, Suzhi Deng, Erdem Oz and Patric Muggli of USC; Chris Clayton, Ken Marsh, Chan Joshi and Devon "Keeping it Real" Johnson of UCLA; Caolionn O'Connell, Dieter Walz, Rasmus Ischebeck and Mark Hogan here at SLAC. All of these are not only colleagues, but friends, and made the last three years a great experience.

Lastly, I wish to thank my advisor Robert Siemann. I chose well, and his guidance, tough love, and expertise have made me grow as a scientist and as a person.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Reasons for Advanced Accelerators

Progress in high energy physics for the last hundred years has been intimately linked with continued developments in methods for accelerating particles to higher and higher energies. As we seek to produce ever more massive particles, a means of continually increasing the energy from accelerators is required. Originally, particles were accelerated in a variety of machines that relied on very large electrostatic fields. Beyond about 10 MeV, however, radiofrequency waves are required for devices such as the early cyclotrons and modern synchrotrons [1]. The most powerful accelerators in the world today all use microwaves inside conducting cavities to accelerate charged particles such as electrons here at the Stanford Linear Accelerator Center (SLAC), LEP and KEK, protons at Fermilab and LHC, and heavy ions at RHIC [2].

Current machines can accelerate particles by at most a few tens of MeV per meter of accelerating structure. The world's longest linear accelerator, the SLC at SLAC, has a maximum gradient of approximately 17 MeV/m. Going to shorter wavelength power sources has allowed the highest demonstrated gradient to date: 65 MeV/m [3]. There are a variety of proposals to extend conventional microwave technology, but all appear limited in the ability to increase gradient to perhaps 150 MeV/m, such as at CLIC [4]. In the nearer term, the gradient in the International Linear Collider (ILC), is projected to be only about twice that of the 40 year old SLC.

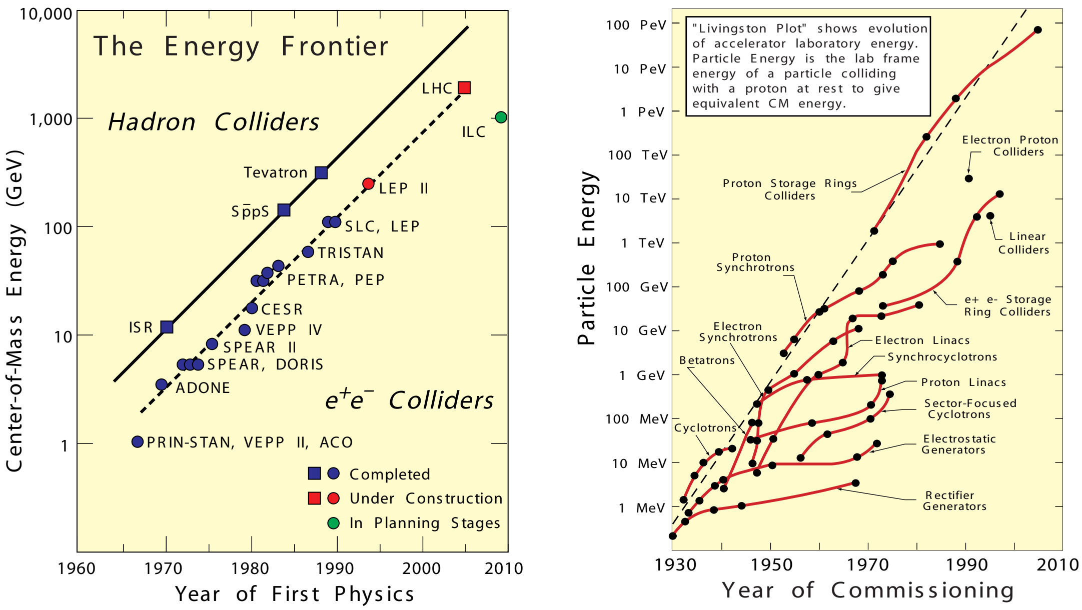

Figure 1.1: Two views of progress in collision energy over time, from [1]

The fundamental limits on current accelerators derive from the properties of the materials from which they are constructed. When the electric fields inside a cavity or the power deposited into that cavity get beyond a certain point, damage occurs. There is a variety of such mechanisms which fall under the rubric of "damage threshold," and these have been investigated widely [5–10].

One approach to achieving stronger acceleration is to use dielectric surfaces to contain laser pulses. Such surfaces have substantially higher damage thresholds from incident laser pulses than any metal [11,12]. This ability to withstand stronger pulses allows a substantial increase in the fields supported in the structure, and shows real promise [13]. Nevertheless, any design which contains strong electromagnetic fields inside a solid piece of material faces damage above some intensity.

For truly dramatic increases in accelerating gradient, the best solution is to do away with solid boundaries altogether by using plasmas instead. Plasmas have demonstrated that they can support electric fields of over 100 GV/m [14] and thus accelerate particles with enormous gradients. The drawback is that the accelerating cavity is no longer a static object such as a machined piece of metal or a lithographically produced dielectric surface, but rather something that must be created dynamically each time a particle bunch is to be accelerated.

The promise of enormously strong accelerating fields has led to much theoretical work, and a number of experiments. We give an overview of two of the main classes of plasma accelerators, and then discuss the basis of this thesis, the E164 experiment at the Stanford Linear Accelerator Center. This electron beam driven plasma wakefield experiment was performed during 2003 and 2004 in the Final Focus Test Beam facility at SLAC, and is a collaboration between scientists from SLAC, the University of California at Los Angeles, and the University of Southern California.

1.2 Plasma Acceleration Overview

The basic idea of a plasma wakefield accelerator is relatively straightforward and seems to have been first proposed by Fainberg in 1956 [15]. Using a laser pulse or an electron beam, one creates a wave inside a plasma by driving the electrons radially out from the position of the beam. This leaves a positively charged column of ions, which do not move significantly on the time frame of plasma electron oscillations. After the passage of the driving beam, the field of the ion column causes the electrons to rush back in to the center and temporarily create an excess of negative charge on axis. The excess of charge has very strong fields in the longitudinal direction which can accelerate any free electrons in that area. This wave of electrons rushing back in propagates behind the driver beam at whatever velocity that beam goes through the plasma, by direct analogy with the wakes behind boats. Figure 1.2 shows a simulation of the electric fields in a plasma from a passing electron beam.

Plasma wakes have been driven both by intense lasers and by electron beams, as in the E164 experiment at SLAC which uses the powerful electron bunches available from the main accelerator. Both laser and electron beam driven schemes have several advantages, so we give a brief introduction to laser wakefield accelerators, partly for background, and partly to give motivation for the E164 experiment.

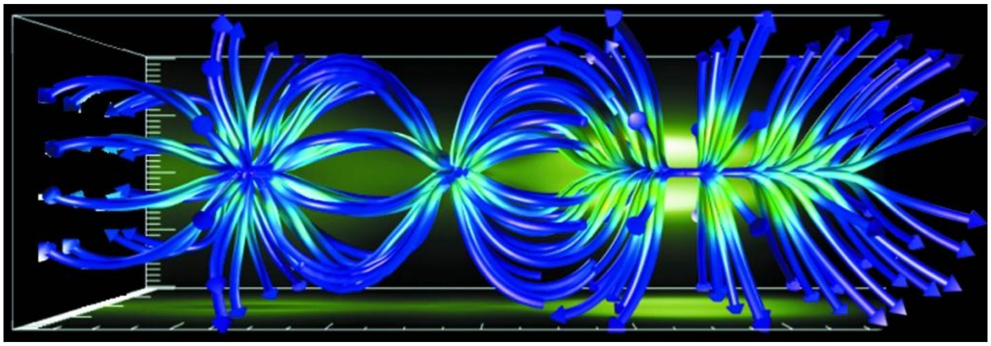

Figure 1.2: Simulation of electric fields due to plasma blowout from an electron beam moving to the right. Acceleration of electrons happens inside the left "bubble" created by the electric field lines. Graphic by Ricardo Fonseca.

1.2.1 Laser Driven Plasma Accelerators

Ultrafast and very powerful lasers were first developed in the 1970's and achieved greater power with the introduction of solid state lasers such as the Ti:Sapphire system now commonly in use for producing very short laser pulses.

Since first being proposed in 1979 by Tajima and Dawson [16], many groups have sought to drive plasma waves using modern intense laser pulses. As indicated above, when laser pulses have ultrahigh intensities, ($\gtrsim 10^{18}$ W/cm^2), they can drive plasma wakes by ponderomotively expelling all of the electrons from a channel in the plasma [17]. A way to visualize the mechanism is that the electric fields are so intense that plasma electrons move significantly during the course of a single optical cycle and are removed from the region of high laser intensity before the opposite phase of the laser oscillation has a chance to arrest their motion. This is very similar to the blow out caused by the simple space charge fields of an electron beam.

The group velocity of a laser pulse through typical plasmas is close to the speed of light, $v_g \lesssim c$, and so provides a synchronous accelerating bucket for electron beams. Most experiments so far have trapped plasma electrons and accelerated them with large energy spread. Fortunately, some recent successes have reduced this particular problem, although with complex arrangements [18–20].

A real concern for all laser wakefield schemes is that a laser pulse focused to a small spot–into a plasma or otherwise–diffracts strongly away from the point of focus unless

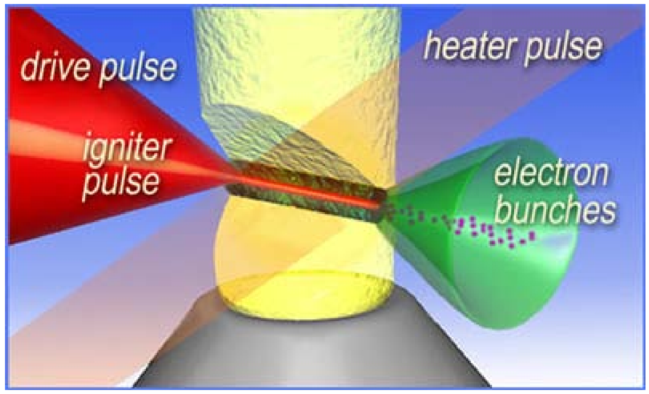

Figure 1.3: Laser acceleration scheme. An igniter pulse forms a "plasma wire" in a hydrogen jet. The heater pulse expands this, making a plasma channel to guide the following drive pulse, which accelerates electron bunches to relatively uniform energy. The green cone is Terahertz radiation from the plasma wake. [21]

specific measures are taken to try to guide the laser in the plasma. This normally limits the distance over which one can create a wake to a few millimeters.

By analogy with optical fibers, some real progress in using plasmas with a spatially varying density profile to guide the laser has begun to solve these concerns, but this is at least a source of significant complexity in using lasers to drive wakefields [19]. For a schematic, see Figure 1.3.

Although laser experiments have had a number of exciting results recently, beams from the SLAC linac have several desirable properties, especially as we seek to sustain the impressive gradients of plasma wakefield accelerators to larger distances.

1.2.2 Beam Driven Plasma Accelerators

A major reason for performing the E164 experiment and its two predecessors, E157 and E162, is that electron beams can propagate long distances in a plasma without requiring complicated measures. Additionally, the SLAC main accelerator produces electron bunches with very advantageous properties for driving a plasma wake. Figure 1.4 shows a schematic of the E164 experiment at the end of the SLAC accelerator.

Electron bunches in the 3 km linac have 2×10^{10} particles, with an average energy of 28.5 GeV. When compressed in the Final Focus Test Beam (FFTB) at the end

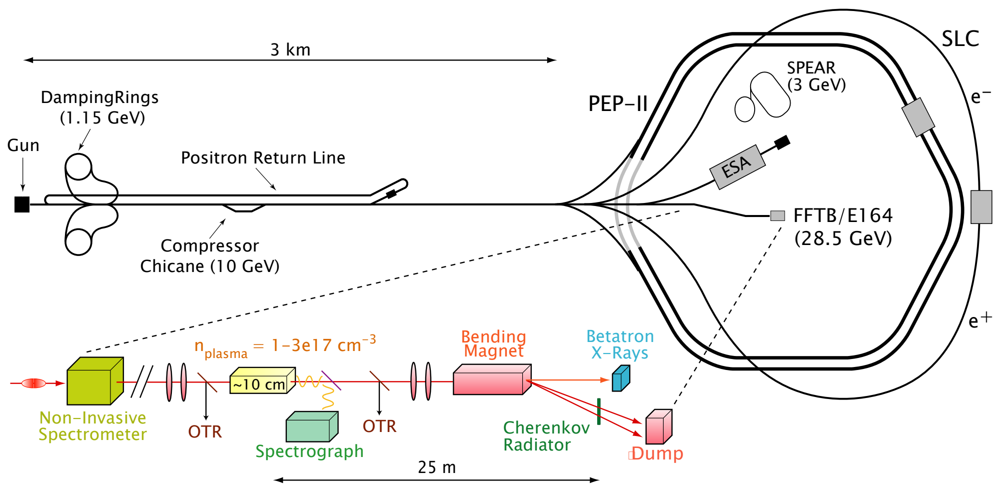

Figure 1.4: The experimental setup for E164

of the accelerator, the bunches can have an RMS length as short as $12\,\mu$m, with peak currents approaching 30 kA. Such beams carry a peak power of over 850 TW. In the E164 experiment, the beams are focused to an RMS radius of about $15\,\mu$m, and so have an impressive maximum intensity of 6×10^{19} W/cm^2. The associated bunch electric fields are greater than 50 GV/m, providing a blowout force comparable to a Ti:Sapphire laser with intensity 3×10^{18} W/cm^2. Thus, like powerful lasers, electron beams from the SLAC accelerator make good drivers for plasma wakes. Being naturally guided for long distances through the plasma, these electron drivers can readily sustain acceleration over meter scale distances.

The original experiment at SLAC, E157, used a 1.4 m plasma oven containing Lithium vapor which was ionized by an ultraviolet laser. As the electron bunches have been shortened in the more recent experiments, the plasma density has been increased so that greater energy gain is achieved in less distance. The length of the plasma is only 10 to 15 cm in the E164 experiment, but this is still more than an order of magnitude longer than in laser wakefield experiments to date.

The guiding of electron beams in a plasma can be understood by a detailed view of what happens as the beam traverses the plasma. When the electron density of the beam exceeds that of the plasma – the underdense regime – all plasma electrons are quickly expelled to a radius greater than that of the beam. This “channel radius,”

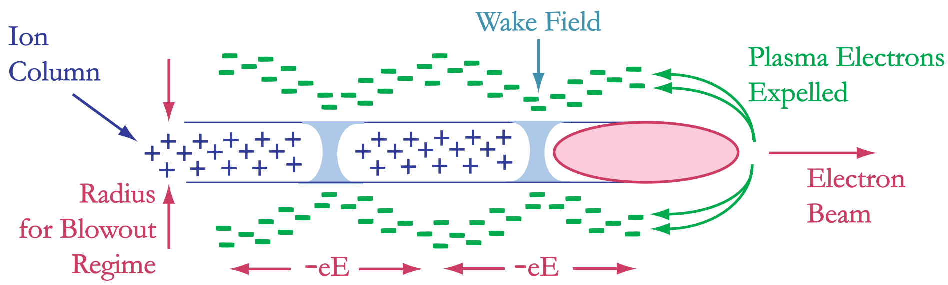

Figure 1.5: Schematic view of beam driven plasma wakefield accelerator.

beyond which the beam's field is shielded, is given by:

$$r_c = \alpha \sqrt{\frac{N}{(2\pi)^{3/2}\sigma_z n_0}} \tag{1.1}$$

where N is the number of particles in our electron beam, σ_z is the RMS length of the bunch and n_0 is the density of plasma electrons with no beam present. In the case where the bunch is of the order of a plasma wavelength long, $\alpha = 2$. For all plasma densities investigated, the channel radius is given by $r_c \geq 27\,\mu$m, which is greater than that of the electron beam in E164 at the location of the plasma.

The positively charged ion column, which remains after electron expulsion, has a radial electric field that grows linearly with radius until one reaches the shell of expelled electrons. This creates an ideal focusing element in both x and y with no geometrical aberrations. The beam-envelope equation describes the behavior of an electron beam traversing such a plasma lens [22]:

$$\frac{d^2\sigma_r(z)}{dz^2} + \left[K^2 - \frac{\epsilon_N^2}{\gamma^2\sigma_r^4(z)}\right]\sigma_r(z) = 0 \tag{1.2}$$

where ϵ_N is the normalized emittance, γ is the usual Lorentz factor of our beam, the plasma restoring constant is $K = \omega_p/c\sqrt{2\gamma}$ with ω_p the "plasma frequency."

This $\omega_p = \sqrt{4\pi n_0 e^2/m}$ is the characteristic frequency of small disturbances in the plasma electrons and can be thought of by direct analogy with classical spring systems where $\omega = \sqrt{k/m}$. In plasmas, the restoring force comes from the other charges. The term K gives the action of the ion column on the *beam* electrons, where the effect is reduced because the beam is relativistic.

Inspection of equation (1.2) shows that there is a σ_r for which the bracketed term is 0 and the beam envelope will thus propagate with no change in size. This “matched condition” is where the so-called emittance pressure exactly balances the focusing force of the ion column, also known as the pinch-force.

In general, the electron beam will come into the plasma with an unmatched size and will therefore undergo oscillation of the beam envelope at half the betatron wavelength of individual particles: $\lambda_\beta = \pi c\sqrt{2\gamma}/\omega_p$. If the beam comes in with twice the matched size, it will pinch down to one half the matched size before expanding back to its original diameter and repeating the process. Similarly if it comes in with ten times the matched size, it will compress briefly to one tenth the matched size.

For both the matched and unmatched cases, the beam envelope is stable or periodic, and this allows transport of the electron beam through long distances. The E157 experiment demonstrated stable transport through approximately four such oscillations in 1.4 meters of plasma.

The focusing forces in plasmas are much stronger than for traditional magnets, and the transverse deflections of beam particles can produce substantial synchrotron radiation of very high energy photons, an important subject in its own right [23]. For acceleration, one seeks to minimize such energy robbing effects. Thus it is optimal to have a beam with the smallest possible emittance and to match the beam as well as possible into the appropriate density plasma. Properly matched, an electron beam driver allows the enormous gradients of plasma wakefields to be sustained over long enough distances to give large absolute energy gains to particles.

The initial E157 experiment at SLAC demonstrated nearly 250 MeV of energy gain to electrons, and the gain is over 1 GeV in the subsequent E164 experiment discussed in this thesis. These are both records for plasma wakefield acceleration due to the advantages of electron beam drivers.

In the following chapter, we outline the basic theory of plasma accelerators, as well as the theory for several effects important in electron beam driven accelerators such as beam ionization of the plasma and the Electron Hose Instability.

Chapter 2

Theory for the E164 Experiment

In this chapter, we discuss theoretical aspects of the E164 program. First, we follow the derivation of plasma wake strength as a function of beam parameters, applying this to predictions for E164. We then discuss the theory of beam ionization of the plasma, a central effect enabling E164.

2.1 Linear Wakefield Theory

The following discussion is adapted from the lecture by Tom Katsouleas of USC given at the Joint US-CERN-Japan-Russia Accelerator School in November of 2002 [24]. It discusses the basic theory of plasma wakefield acceleration by electron beams and gives useful scalings. Although we use MKS units for presentation of most experimental results, theoretical calculations in this thesis are generally performed in CGS units, and we convert results to MKS for comparison with data.

A negative test charge in a plasma will repel the neighboring plasma electrons. This creates a small region where the plasma electron density is at a minimum at the position of the test charge and returns to the original density further away once the plasma has shielded the test charge's electric field. The characteristic scale length of this shielding is known as the "Debye Length" [25, 26] and is given by $\lambda_D = v_{th}/\omega_p$ where v_{th} is the average thermal velocity of the plasma's electrons and $\omega_p = \sqrt{4\pi n_0 e^2/m}$ is the plasma frequency, as described in Chapter 1.

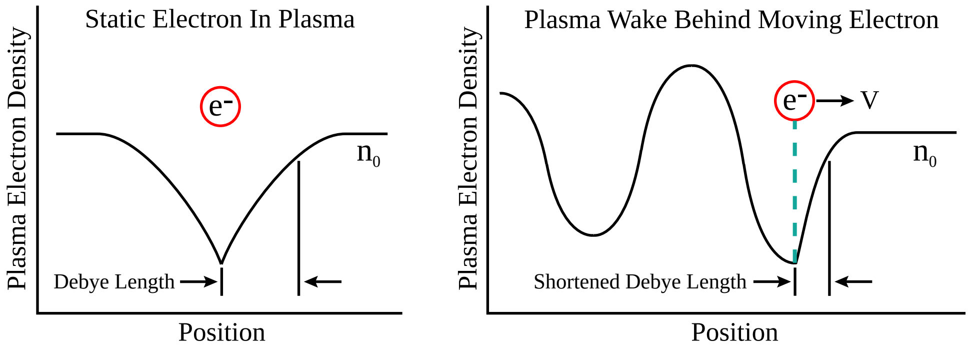

Figure 2.1: Electron Wake Schematic

When a test charge moves through the plasma, the shielding length will contract in front of the charge as the particle's velocity approaches v_{th} of the plasma's own electrons. The shielding length increases behind, and when the particle velocity is high enough, a plasma oscillation is created after the driving particle as the displaced electrons rush back in due to the restoring force of the largely unmoved ions. Figure 2.1 shows a simplified one-dimensional view of such a process. The inrushing electrons can temporarily create a region of even higher electron density than before the disturbance, which therefore has very strong accelerating fields in front of it. The excess electron density on axis leads to a re-expansion and potentially many cycles of the plasma oscillation long after the test charge has passed. This is the basic mechanism for any beam driven accelerator.

2.1.1 Field Strength in Plasma Wakes

Plasmas are of interest because they can support enormous electric fields, and we seek to know the maximum possible field for a given plasma [27]. We start by writing Gauss' Law:

$$\nabla \cdot \mathbf{E} = 4\pi\rho = 4\pi e(n_0 - n_e) = 4\pi e\,\delta n_e \qquad (2.1)$$

where the relevant charge density is given by the difference at any given point between the prevailing plasma electron density n_0 and the local density, n_e. The largest wakes are when all plasma electrons have been expelled, so that $n_e = 0$ and $\delta n_e = n_0$.

The magnitude of the left hand side is proportional to the wavenumber of the plasma: $|\nabla \cdot \mathbf{E}| \sim |ik_p\mathbf{E}| \sim \frac{\omega_p}{c}E$. Note that ω_p is solely a function of plasma density. Intuitively, the denser a plasma becomes, the stronger an electric field it can support when its electrons are displaced. However, the dependence of the field strength is only as the square root of plasma density, because ω_p has that dependence.

For full displacement, we combine the equations and obtain $\frac{2\pi}{\lambda_p}E_{max} \sim 4\pi e n_0$. Substituting and rearranging, one finds that $E_{max} \sim \omega_p mc/e$. This is the non-relativistic wave breaking field, and has a convenient engineering formula. The peak achievable field is approximately $\sqrt{n_0}$ V/cm when the density is given per cubic centimeter.

For the plasma densities of $3 \times 10^{17}\mathrm{cm}^{-3}$ in the E164 experiment, we expect to be able to produce fields of order 50 GeV/m, offering more than a thousandfold increase over the available fields in traditional accelerators such as the SLAC linac.

This analysis ignores relativity for the plasma electrons, which can certainly come into play with strong fields. Still, it is a useful touchstone for understanding the strength of plasma wakes. One immediately sees why the fields possible in plasmas have led to great interest among the accelerator and high energy physics communities.

2.1.2 Linear Cold Fluid Theory

In general, one does not have a solitary charge moving through a plasma, but an electron beam which has an electron density as a function of position within the bunch. For a highly relativistic bunch, $n_b = n_b(z - ct, r)$. In the related case where a laser drives the plasma wake, one has that the laser intensity is given by $I_0 = I_0(z - v_g t, r)$, with v_g the laser's group velocity in the plasma.

The time scale of wakefield generation is short compared to that for the ions to move, so in this analysis, they are treated as remaining fixed in position. In deriving the Cold Fluid Equation for plasma electrons, we start with the Continuity Equation:

$$\frac{\partial n}{\partial t} + \nabla \cdot n\mathbf{v} = 0 \tag{2.2}$$

We linearize (2.2) with an expansion where $n = n_0 + n_1 + \cdots$ and $\mathbf{v} = \mathbf{v_0} + \mathbf{v_1} + \cdots$.

Combining with the equation of motion for $\mathbf{v_1}$, we obtain the pair of equations:

$$\frac{\partial n_1}{\partial t} + n_0 \nabla \cdot \mathbf{v_1} = 0 \tag{2.3}$$

$$m \frac{d\mathbf{v_1}}{dt} = -e\mathbf{E_1} + \mathbf{F_p} \tag{2.4}$$

Usually, one has only one of the terms on the right hand side of Equation (2.4). For the case of driving a plasma wake with an intense laser pulse, the ponderomotive force is given by $\mathbf{F_p} = -m_e c^2 \, \nabla(a^2/2)$ where the normalized vector potential of the laser field is given by $\mathbf{a} = e\mathbf{A}/m_e c^2$. This force can also be described as coming from the gradient of the radiation pressure of the laser pulse [28].

Beam driven experiments such as E164 have only space charge forces from the beam, so we drop the term referring to laser fields and keep only the electric field term $-e\mathbf{E_1}$ in the following analysis.

Taking the time derivative of (2.3) and then substituting (2.4) for $\partial \mathbf{v_1}/\partial t$:

$$\frac{\partial^2 n_1}{\partial t^2} + n_0 \nabla \cdot \left(-\frac{e\mathbf{E_1}}{m} \right) = 0 \tag{2.5}$$

We apply Gauss' Law to the divergence of $\mathbf{E_1}$ term, where $\nabla \cdot \mathbf{E_1} = -4\pi e(n_1 + n_b)$. Substituting into (2.5), we obtain:

$$\frac{\partial^2 n_1}{\partial t^2} + \frac{4\pi n_0 e^2}{m} n_1 + \frac{4\pi n_0 e^2}{m} n_b = 0 \tag{2.6}$$

Note that the terms before n_1 and n_b are just the square of the plasma frequency, thus the Cold Fluid Equation can be written:

$$\frac{\partial^2 n_1}{\partial t^2} + \omega_p^2 n_1 = -\omega_p^2 n_b \tag{2.7}$$

2.1.3 2-D Linear Wakefield Theory

This theory is simplified from a full three dimensional theory by assuming azimuthal symmetry. Equation (2.7) tells the density response of the plasma to a disturbance from an electron beam. We use this to determine the strength of the plasma wake. In the laboratory frame, $n_b = n_b(z - ct, r)$, so we define the comoving coordinate $\xi = z - ct$. Thus $\frac{\partial}{\partial z} = \frac{\partial}{\partial \xi}$ and $\frac{\partial}{\partial t} = -c\frac{\partial}{\partial \xi}$. We can substitute into the Cold Fluid Equation (2.7) and divide by c^2 to obtain:

$$\left(\frac{\partial^2}{\partial \xi^2} + k_p^2\right) n_1 = -k_p^2\, n_b(\xi, r) \tag{2.8}$$

This represents an oscillation, so we use a Simple Harmonic Oscillator Green's function solution to this in integral form:

$$n_1(\xi, r) = k_p \int_{\xi}^{\infty} d\xi'\, n_b(\xi', r)\, \cos k_p(\xi - \xi') \tag{2.9}$$

This integral is taken only over the charge ahead of position ξ because, by causality, nothing behind ξ can affect that position. We recall that n_1 has been assumed to be a small linear perturbation on the prevailing plasma density, and the disturbance i

We now seek an expression for the wakefields associated with our change in electron density, $n_1(\xi, r)$. We use Faraday's Law, taking the cross product of both sides:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\frac{1}{c}\frac{\partial}{\partial t}\mathbf{B}\right) \tag{2.10}$$

For the curl of the magnetic field on the right hand side, we use Ampère's law with Maxwell's Correction:

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \tag{2.11}$$

On the left, we apply the standard vector calculus identity. With both substitutions, we obtain:

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{1}{c}\frac{\partial}{\partial t}\left(\frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\right) \tag{2.12}$$

We can rearrange the two sides of (2.12) such that all currents and charges will ultimately be on the right hand side. Using shorthand for the partial derivatives:

$$\frac{1}{c^2}\,\partial_t^2\,\mathbf{E} - \nabla^2 \mathbf{E} = -\frac{4\pi}{c^2}\,\partial_t\,\mathbf{J} - \nabla(\nabla \cdot \mathbf{E}) \tag{2.13}$$

It is useful to separate the transverse and longitudinal directions in the derivatives. Doing so allows the elimination of several terms because, for a highly relativistic particle, the time and position derivatives cancel. On the left hand side we see that:

$$\text{LHS} \;=\; \frac{1}{c^2}\,\partial_t^2\,\mathbf{E} - \nabla^2\mathbf{E} \;=\; \left(\frac{1}{c^2}\,\partial_t^2 - \nabla_\xi^2\right)\mathbf{E} - \nabla_\perp^2\mathbf{E} \;=\; -\nabla_\perp^2\mathbf{E} \tag{2.14}$$

On the right hand side, we can substitute in for $\mathbf{J}$ and use Gauss' Law for $\nabla \cdot \mathbf{E}$:

$$\text{RHS} \;=\; -\frac{4\pi}{c^2}\,\partial_t\,\left(-\mathbf{v_1}en_0 - cen_b\hat{\xi}\right) - \nabla\left(-4\pi en_1 - 4\pi en_b\right) \tag{2.15}$$

Next we use the equation of motion to substitute $-e\mathbf{E}/m$ for $\partial_t\mathbf{v_1}$. We collect all other terms and divide the gradient of n_b into longitudinal and transverse portions, again cancelling time and position derivatives. Finally, we combine the constants in front of the leading term and find that they equal ω_p/c^2 or k_p^2:

$$\begin{aligned}
\text{RHS} \;&=\; -\frac{4\pi}{c^2}\,\partial_t\,\left(-\mathbf{v_1}en_0 - cen_b\hat{\xi}\right) + 4\pi e\left(\nabla n_1 + \nabla n_b\right) \\
&=\; \frac{4\pi}{c^2}\left(-\frac{e\mathbf{E}}{m}\right)en_0 + 4\pi e\left(\frac{1}{c}\partial_t n_b\hat{\xi} + \nabla n_1 + \nabla_\xi n_b + \nabla_\perp n_b\right) \\
&=\; k_p^2\mathbf{E} + 4\pi e\left(\nabla n_1 + \nabla_\perp n_b\right)
\end{aligned} \tag{2.16}$$

Combining the $\mathbf{E}$ terms, we obtain a differential equation for the total field in terms of the beam density and the plasma density displacement due to the beam:

$$\left(\nabla_\perp^2 - k_p^2\right)\mathbf{E} = -4\pi e\left(\nabla n_1 + \nabla_\perp n_b\right) \tag{2.17}$$

We wish to know the transverse and longitudinal components of the wakefields, which are defined as the force per unit charge:

$$W_r \equiv \frac{F_\perp}{q} = \left(E + \frac{\mathbf{v}}{c}\times\mathbf{B}\right) = E_r + \hat{\xi}\times B_\theta\hat{\theta} \tag{2.18}$$

$$W_\xi \equiv \frac{F_\parallel}{q} = E_\xi \tag{2.19}$$

The Panofsky-Wenzel Theorem [29] relates these two wakefields:

$$\frac{\partial W_r}{\partial \xi} = -\frac{\partial W_\xi}{\partial r} \tag{2.20}$$

With this equation and assuming the separability of the longitudinal and transverse variables, it is useful to define functions $Z(\xi)$ and $R(r)$ such that $W_r = -Z\,\frac{\partial R}{\partial r} = Z\,R'$ and $W_\xi = \frac{\partial Z}{\partial \xi}\,R = Z'\,R$.

Because the longitudinal wakefield is simply given by E_ξ, we obtain that:

$$E_\xi = Z'(\xi)\,R(r) \tag{2.21}$$

We want to solve (2.17) for the longitudinal direction, so insert the separated variable expression for E_ξ. Note that to calculate the strength of the *wakefield*, rather than the total field, we do not include the term involving the beam. Thus only n_1 appears:

$$\hat{\xi}: \quad \left(\nabla_\perp^2 - k_p^2\right)\, Z'\,R = -4\pi e\,\partial_\xi n_1 \tag{2.22}$$

On the left hand side, Z' can clearly be pulled out of the transverse derivative. On the right hand side, we use Equation (2.9) to substitute for n_1. We assume that the description of the electron beam can be separated so that $n_b(\xi, r) = n_b(\xi) f(r)$ in the expression for n_1. Substituting into the Wake Equation:

$$Z'(\xi)\,(\nabla_\perp^2 - k_p^2)\,R(r) = -4\pi e\,\partial_\xi \left[k_p \int_\xi^\infty d\xi'\,n_b(\xi')f(r)\,\cos k_p(\xi - \xi')\right] \tag{2.23}$$

We can pull $f(r)$ out from the integral, and move ∂_ξ in. After reorganizing, we obtain:

$$Z'(\xi)\,(\nabla_\perp^2 - k_p^2)R(r) = -4\pi e k_p \left[\int_\xi^\infty d\xi'\,n_b(\xi')\,\partial_\xi \cos k_p(\xi - \xi')\right]\,f(r) \tag{2.24}$$

We perform the derivative and pull out the term $-k_p^2$ to combine with $f(r)$ for later convenience:

$$Z'(\xi)\,\left[(\nabla_\perp^2 - k_p^2)R(r)\right] = \left[-4\pi e\int_\xi^\infty d\xi'\,n_b(\xi')\,\sin k_p(\xi - \xi')\right]\left[-k_p^2 f(r)\right] \tag{2.25}$$

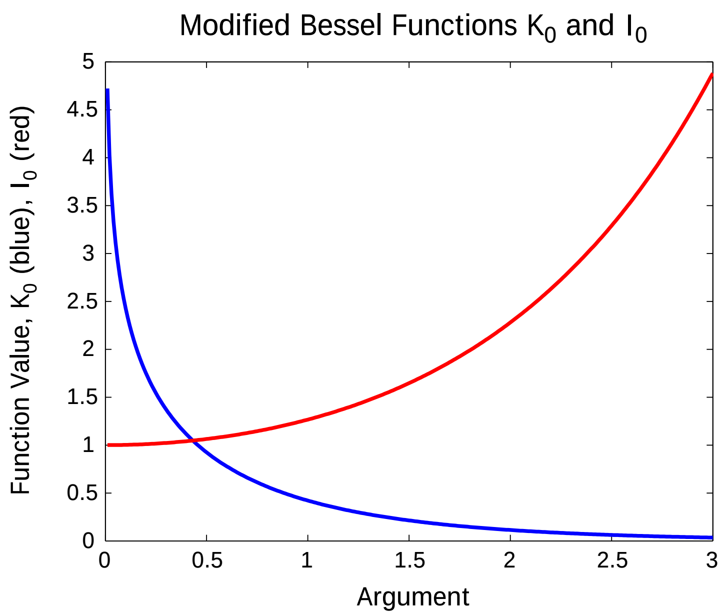

Figure 2.2: Plot of two modified Bessel Functions for arguments 0 to 3.

We use brackets to make explicit the separation of transverse and longitudinal coordinates on both sides of the equation. The only way for (2.25) to be true in general is for the portions dealing with the ξ and r components each to be equal separately. For Z', the relation is immediate, and takes similar form to the expression for n_1. In the transverse direction, we have to solve:

$$(\nabla_\perp^2 - k_p^2)R(r) = -k_p^2 f(r) \tag{2.26}$$

For this, the solution to the Kelvin-Helmholtz Equation serves as a Green's Function, and it involves the modified Bessel Function of the second kind, K_0. For reference, we plot this, as well as its sibling I_0, the modified Bessel Function of the first kind, in Figure 2.2. The Green's Function solution to (2.26) is:

$$(\nabla_\perp^2 - k_p^2)\, G = \delta^2(r) \quad \longrightarrow \quad G = -\frac{1}{2\pi} K_0(k_p|\mathbf{r} - \mathbf{r}'|) \tag{2.27}$$

where the solution gives the effect on a particle at (r, θ) from a particle at (r', θ'). To find R, we must integrate over the distribution $f(r)$. Thus our two functions Z' and

R are given by the expressions:

$$Z'(z) = -4\pi e \int_{\xi}^{\infty} d\xi' \; n_b(\xi') \, \sin k_p(\xi - \xi') \tag{2.28}$$

$$R(r) = k_p^2 \int_0^{2\pi}\int_0^{\infty} r' dr' d\theta' \; f(r') \, \frac{1}{2\pi} \, K_0(k_p|\mathbf{r} - \mathbf{r}'|) \tag{2.29}$$

These provide a general solution to the case of a beam driven plasma accelerator, but obviously the solution needs to be adapted to a given specific case. We now focus on the E164 experiment.

2.1.4 Application of Theory to Narrow Beams

It is useful to rewrite the Bessel function part of the integrand in (2.29) as:

$$\frac{1}{2\pi} K_0(k_p|\mathbf{r} - \mathbf{r}'|) = \frac{1}{2\pi} K_0\big(k_p(r^2 + r'^{\,2} - 2rr' \cos\theta)^{\frac{1}{2}}\big) \tag{2.30}$$

This can be expressed in a different form if we change our notation. Let $r_>$ be the greater of the two transverse positions r and r'. Similarly, let $r_<$ be the lesser of the two. Following [30], this substitution gives :

$$\frac{1}{2\pi} K_0(k_p|\mathbf{r} - \mathbf{r}'|) = K_0(k_p r_>) I_0(k_p r_<) + 2 \sum_{1}^{\infty} \cos(m\theta) \, I_m(k_p r_<) K_m(k_p r_>) \tag{2.31}$$

We note that we are assuming an azimuthally symmetric beam. Thus, we will be integrating θ from 0 to 2π because we are only interested in r here. In this integral, the sum over m goes to zero, and we can immediately drop the last term. With no θ dependent term remaining from the Bessel function, and assuming that the beam is azimuthally symmetric, the first integral in our equation for $R(r)$ is simply 2π, and we obtain that

$$R(r) = k_p^2 \int_0^{\infty} r' dr' \, f(r') K_0(k_p r_>) I_0(k_p r_<) \tag{2.32}$$

An analytically tractable description can provide useful guidance. We assume that the transverse beam density is given by a normalized step function: $f(r) = \frac{1}{\pi a^2} H(a - r)$

such that the beam has radius a. Because a given electron beam can change its diameter through focusing effects while maintaining constant charge, we normalize the step function so that the integral over the beam diameter gives constant charge regardless of beam size. In this case, the solution for $R(r)$ is given by

$$R(r) = \frac{1}{\pi a^2}\big(1 - k_p a\, K_1(k_p a) I_0(k_p r)\big) \qquad r < a \tag{2.33}$$

$$= \frac{1}{\pi a^2}\big(k_p a\, I_1(k_p a) K_0(k_p r)\big) \qquad r > a \tag{2.34}$$

Naturally, we are mostly concerned with the case where $r < a$, because that is where the beam's particles lie, by construction. We consider only the first solution.

Within this solution, two regimes have useful results. For "wide beams," where the diameter is large compared to the plasma wavelength, $k_p a \gg 1$ and R is constant inside the beam: $R(r) = R(0) = 1/\pi a^2$, shrinking rapidly outside of the beam.

In the more common case where the beam is narrower than the plasma wavelength, we have that $k_p a < 1$. For $r < a$, clearly $k_p r < 1$, as well. With this, we can simplify the result (2.33), because for $k_p r$ up to unity, $I_0 \approx 1$ and we ignore it henceforth.

If we take a very narrow beam, such that $k_p a \ll 1$, we obtain a simple result for the accelerating field:

$$R(0) = \frac{1}{\pi a^2}\big(1 - k_p a\, K_1(k_p a)\big) \approx \frac{1}{\pi a^2}\big(\pi (k_p a)^2\big) = k_p^2 \tag{2.35}$$

An approximation for $R(0)$ which is valid for beams where $k_p a \leq 1$ has been found by colleagues on the E164 experiment from USC and UCLA [31]. In this approximation, we have both a constant and a logarithmic term:

$$R(0) = \frac{1}{\pi a^2}\big(1 - k_p a K_1(k_p a)\big) \approx \frac{k_p^2}{2\pi}(0.6159 - \ln k_p a) \equiv \frac{k_p^2}{2\pi} f(a) \tag{2.36}$$

This approximation gives 25% accuracy at $k_p a = 1$, and better than 5% accuracy for $k_p a < 0.5$ with a simple functional form. The function involving a plays no role in the following derivation, so for simplicity we will refer to $(0.6159 - \ln k_p a)$ as $f(a)$.

Very narrow beams have stronger wakes than wider beams. Physically, this is because as the beam shrinks, few plasma electrons start inside the radius of the beam, which is a place where the fields are reduced. Thus all particles are driven similarly, producing a stronger wake.

With this result for R, we combine (2.28) and (2.36) to get:

$$W_z = Z'\,R \approx -4\pi e\frac{k_p^2}{2\pi}f(a)\int_\xi^\infty d\xi'\; n_b(\xi')\sin k_p(\xi - \xi') \tag{2.37}$$

Since many electron beams in accelerators are more or less Gaussian all dimensions, we can solve for the case of a narrow beam with longitudinally Gaussian distribution, where the transverse size does not matter:

$$W_z = -2ek_p^2 f(a)\int_\xi^\infty d\xi'\left[\frac{N}{\sqrt{2\pi}\sigma_z}\; e^{-\frac{\xi'^2}{2\sigma_z^2}}\right]\sin k_p(\xi - \xi') \tag{2.38}$$

The wake does not develop until after the peak of the bunch passes, and the strongest wakefields will be about one quarter of a plasma wavelength behind that point, so that $k_p\xi \approx -\pi/2$. Thus we solve for the wake behind the electron bunch, where $\xi < -\sigma_z$. We note that the dominant term in the integral is the Gaussian envelope, and that having the integral run from ξ to ∞ is very similar to integrating from $-\infty$ to ∞, once ξ is more negative than about one sigma. This approximation gives a simple analytic solution to the integral. Lastly, we substitute $u = \xi/\sqrt{2}\,\sigma_z$ and reorganize:

$$W_z = -\frac{2Nek_p^2 f(a)}{\sqrt{2\pi}\sigma_z}\int_{-\infty}^{\infty}\sqrt{2}\,\sigma_z du'\; e^{-u'^2}\cos\left(\sqrt{2}k_p\sigma_z(u - u')\right) \tag{2.39}$$

The integral portion of this equation is given by:

$$I = \sqrt{2}\sigma_z\left[\sqrt{\pi}\,\exp\left(-\frac{k_p^2\sigma_z^2}{2}\right)\sin k_p\xi\right] \tag{2.40}$$

We combine and cancel several terms to find that:

$$W_z = -2Nek_p^2 f(a)\;\exp\left(-\frac{k_p^2\sigma_z^2}{2}\right)\sin k_p\xi \tag{2.41}$$

Using (2.41), we seek to find the maximum wake available. First, we choose the distance behind the bunch ξ such that $k_p \xi = -\pi/2$ and the sine term gives maximum accelerating field for a negatively charged particle. Assuming that we have properly chosen the distance behind the bunch for a given plasma density, we then are concerned with finding the optimal bunch length, σ_z, to match to our plasma density. To do this, we reorganize the constant terms and perform another variable substitution, letting $v = \frac{k_p^2 \sigma_z^2}{2}$:

$$W_z = -\frac{4Nef(a)}{\sigma_z^2} \left(\frac{k_p^2 \sigma_z^2}{2} \right) \exp\left(-\frac{k_p^2 \sigma_z^2}{2} \right) = \frac{4Nef(a)}{\sigma_z^2} \, v e^{-v} \tag{2.42}$$

We can find the maximum of the function of v: $\partial_v (v e^{-v}) = (1-v)e^{-v} = 0$ has its maximum at $v = 1$. Using this, we see that the wake is maximized for $k_p \sigma_z = \sqrt{2}$.

When we have the correct plasma density to match a given bunch length, another way to express the location of peak field is to say that $\xi = -\frac{\pi}{2\sqrt{2}} \sigma_z$. With this condition satisfied, the magnitude of the maximum possible wakefield is given by:

$$\widehat{W}_z = \frac{4Nef(a)}{\sigma_z^2} \exp(-1) \approx 1.47 f(a) \, e \, \frac{N}{\sigma_z^2} \tag{2.43}$$

While purely a result of linear theory, this result, commonly invoked as "The N/σ_z^2 Scaling Law," [32, 33] provides a decent ballpark estimate for many circumstances.

2.1.5 Numerical Estimates for E164

We wish to include the effects stemming from the transverse beam size, which in E164 is not small compared to the plasma wavelength. To calculate a prediction for this experiment, we rewrite (2.43) with the beam diameter explicitly included:

$$\widehat{W}_z = 1.47e \, \frac{N}{\sigma_z^2} \left(0.6159 - \ln k_p a \right) \tag{2.44}$$

For the experiments at SLAC, the beam diameter is generally smaller than the plasma wavelength, and the logarithm term should dominate. For a rough estimate of what we expect the dependence to be, we can modify (2.44) by dropping the constant term:

$$\widehat{W}_z = 1.47e \frac{N}{\sigma_z^2} \ln \frac{1}{k_p a} \tag{2.45}$$

This gives the expected functional form for narrow beams. The precursor experiments E157 and E162, where there were 4×10^{10} particles in a bunch $600\,\mu$m long of radius $50\,\mu$m radius, found acceleration of about 236 MV/m in a plasma density of 10^{16} cm^{-3}. So we can simply scale the results with the useful engineering formula for how much acceleration is expected in a single bunch plasma accelerator [34, 35]:

$$E_z\ [MV/m] \approx 236 \left(\frac{N}{4 \times 10^{10}}\right) \left(\frac{600\,\mu\text{m}}{\sigma_z}\right)^2 \ln\left(\sqrt{\frac{10^{16}\ \text{cm}^{-3}}{n}}\ \frac{50\,\mu\text{m}}{\langle\sigma_r\rangle}\right) \tag{2.46}$$

We note that $\langle\sigma_r\rangle$ is the average RMS radius of the electron beam. As discussed at the end of Chapter 1, c.f. (1.2), there is a matched size for the electron beam in a plasma. In E164, the matched size for our plasma is (depending on the plasma density) of the order of 1 micron, much smaller than the typical $15\,\mu$m incoming beam spot size, so the beam immediately pinches dramatically. An order of magnitude estimate is that $\langle\sigma_r\rangle$ is about half of the incoming spot size, since the beam repeatedly alternates between full size and tightly pinched.

Realizing that this is only an estimate extrapolated from linear theory, we can plug in the typical values for the beam size and plasma density in E164. As discussed subsequently, $\sigma_z \approx 20\,\mu$m and $\sigma_r \approx 15\,\mu$m, with roughly 1.8×10^{10} particles in each bunch. To match the plasma wavelength to this bunch length, we would use use a plasma density of approximately 1.5×10^{17} cm^{-3}. For this set of parameters, (2.46) predicts an accelerating gradient of about 55 GeV/m. This is in reasonable accord with the 37 GeV/m predicted by the simplistic wave breaking formula. In Chapter 5, we will later see that the actual acceleration is lower, but it is within the same order of magnitude.

2.2 Detailed Bunch Considerations

Linear theory indicates that we want the shortest beams possible. Having generated a strong wake, one ideally wants a second bunch of particles trailing behind that can be accelerated monoenergetically. In a useful accelerator for High Energy Physics experiments, one thus seeks a high charge drive bunch and a lower charge bunch following behind by less than a plasma wavelength.

Using a conventional linac to produce two distinct electron beams with a separation substantially smaller than these plasma wavelengths is being investigated for a future experiment, but has so far proven extremely challenging. For the series of experiments culminating in E164, efforts concentrated on using single beams both to produce a useful wake and to provide trailing particles to witness the wake.

In addition to considerations of multiple bunches, there is theoretical reason to consider tailored bunch profiles to increase the wake strength.

2.2.1 Production of Strong Wakes

Linear theory tells us that short electron bunches are needed for strong wakes. The analysis assumes that the bunch has a Gaussian current profile, but the details of the current profile can make the wake substantially stronger in the linear regime [36]. We define the "Transformer Ratio" R_t to be the ratio of the energy gained by a particle in the wake to the energy loss of the drive beam. It has been shown in general that for beams with symmetric rising and falling current profiles, $R_t \leq 2$ [37].

Asymmetric beams, however, have the possibility to enhance the wake. Of particular interest have been "doorstop" beams where the current profile rises gradually over several plasma periods and then drops rapidly. Figure 2.3 shows the enhanced wake that gets created behind such a beam [38]. In this case, the simulation shows that the Transformer Ratio increases to approximately five.

For E164, long ramped beams have several problems. First, these are not easy to create in an accelerator. Also, with a fixed number of electrons, a long beam has low electron density. We then must use low density plasmas so that the beam will still have greater electron density than the plasma.

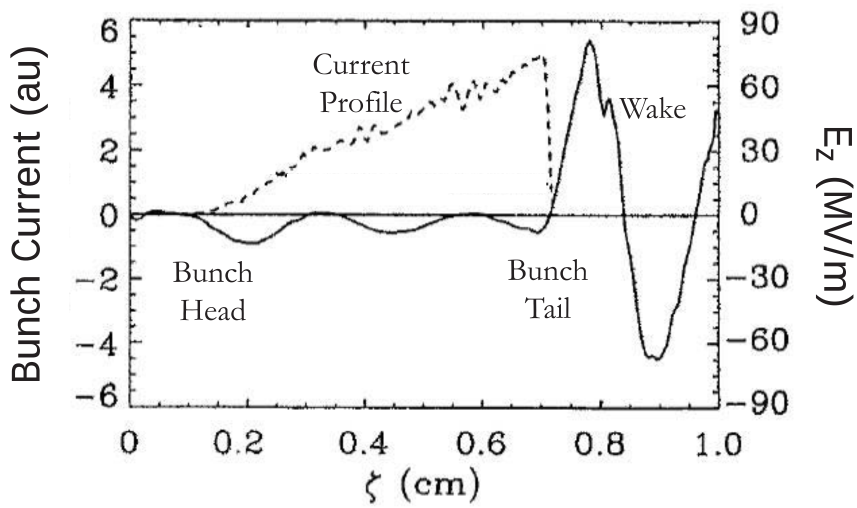

Figure 2.3: Ramped beam shows a strong wake behind. The dashed line is the current profile of the bunch, and the solid line shows the axial wakefield, which develops rapidly once the beam passes. From the simulation presented in [38].

Additionally, in E164, the beam ionizes neutral Lithium vapor by its own intense space charge fields. For a beam of given transverse size, this places a hard upper limit on how long the beam can be and still fully ionize the Lithium so that a plasma wake can be driven. As discussed below, for E164, that limit is less than about 100 μm.

Lastly, for a beam which is long compared to a plasma wavelength, one enters the regime of the "Electron Hose Instability" which is of possible concern for plasma accelerators [39], though current theory does not fully describe the relatively short bunches in E164. So although potentially interesting in some applications, the E164 experiment has avoided long ramped beams and focused simply on creating the shortest beams possible as the best means to increase accelerating gradient.

2.3 Field Ionization Requirements

The plasma source for the experiment is really a source of neutral vapor which can then be ionized just as the electron beam enters the vapor. For the majority of the experiments at SLAC, starting with E157, the plasma source has been Lithium vapor in a heated tube [40]. Originally, ionization was achieved through the use of an

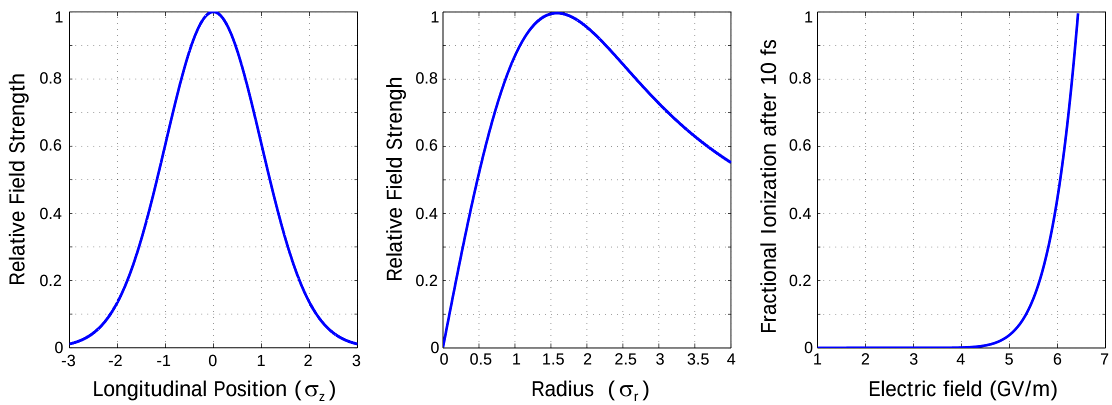

Figure 2.4: Relative field strength in a Gaussian bunch versus z and r. On the right is the fractional ionization of Lithium after 10 fs as a function of the electric field.

ultraviolet laser which was brought into the Lithium vapor collinear with the beam. Absorption of the photons singly ionized the Lithium along the laser's path.

This suffers from obvious difficulties associated with making the small diameter laser and electron beams collinear over long distances, as well as with having the necessary UV optics directly in the path of an intense ultrarelativistic electron beam.

As we seek to use long columns of dense Lithium vapor, the laser pulse energy required to ionize grows rapidly. Our laser system would not be powerful enough to create the desired plasmas. More importantly, creation of a uniform plasma density in a long, high density vapor column is very difficult because of the exponential attenuation of the UV as it propagates through the vapor. Having a consistent plasma density is crucial to understanding the acceleration we create.

A solution presents itself in E164 when we focus our very short electron beam down to very small spot sizes. The space charge fields for our beams are intense enough to drive many GeV/m wakes in plasma and are so strong that they can simply rip electrons from the Lithium atoms to ionize directly. Quantum mechanically, we can view this process as occurring because the barrier for an electron to tunnel out of its atom is depressed strongly by the beam's large field. The electron has a high probability of escaping in the time it takes for the beam to pass by. We summarize the theory briefly, but for detailed discussion of tunnelling ionization in the context of the E164 experiment, as well as results, please see Caolionn O'Connell's thesis [41].

We use atomic physics calculations to estimate the rate of ionization from strong fields. The Ammosov-Delone-Krainov (ADK) model [42, 43] gives an engineering formula for the rate of ionization of Lithium for a given electric field:

$$W[s^{-1}] \approx 1.54 \times 10^{15} \, \frac{4^n \, \zeta \, [eV]}{n \, \Gamma(2n)} \left(20.5 \, \frac{\zeta^{3/2} \, [eV]}{E \left[\frac{GV}{m}\right]} \right)^{2n-1} \exp \left(-6.83 \, \frac{\zeta^{3/2} \, [eV]}{E \left[\frac{GV}{m}\right]} \right) \quad (2.47)$$

where ζ is the ionization potential of the desired atomic species, and E is the electric field. The effective principal quantum number is given by $n \approx 3.69 \, Z/(\sqrt{\zeta} \, [eV])$ where Z is the ionization state being calculated. We are concerned with singly ionized Lithium, so $Z = 1$. Gamma is the usual generalization of the factorial function. Knowing the 5.4 eV ionization potential of Lithium, we can simplify (2.47) to:

$$W_{Li} \approx \frac{3.60 \times 10^{21}}{E^{2.18} \left[\frac{GV}{m}\right]} \exp \left(\frac{-85.5}{E \left[\frac{GV}{m}\right]} \right) \quad (2.48)$$

When we focus beams from the SLC down to spot sizes of $15 \, \mu$m and compress to bunch lengths of about $20 \, \mu$m, the space charge fields consistently are incredibly strong. For a three dimensionally Gaussian bunch travelling relativistically, it is straightforward to calculate the peak field, which occurs at the center of the beam along the $\hat{z}$ direction, and at a radius of about $1.6 \, \sigma_r$, as shown in figure 2.4. The peak field can be given by the useful engineering formula adapted from [41]:

$$E_{max} \, [GV/m] \approx 40 \left(\frac{N}{2 \times 10^{10}} \right) \left(\frac{15 \, \mu\text{m}}{\sigma_r} \right) \left(\frac{20 \, \mu\text{m}}{\sigma_z} \right) \quad (2.49)$$

The peak fields in our plasma are enormous. Although the fields are weak in the very center of the electron beam, from about one half sigma out to many sigma, the fields are at least half as strong as the peak field. This region of good ionization encompasses the vast majority of the plasma electrons, so wake production is not substantially changed from a fully ionized case.

In $\hat{z}$, short bunches are desirable. The greater bunch electron density means that the fields are strong enough to ionize the Lithium an earlier point in the beam. Thus

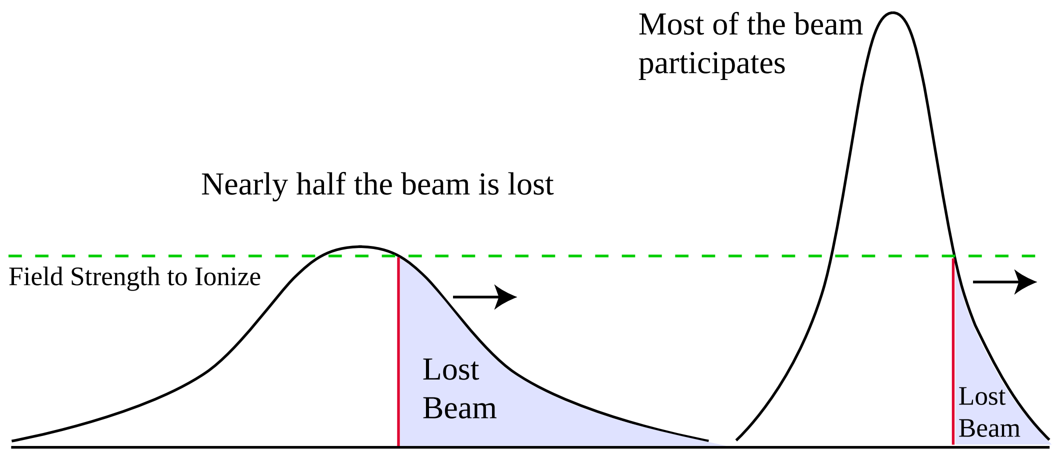

Figure 2.5: Dual Advantages of Longitudinal Beam Compression.

a larger fraction of the beam electrons participate in driving the wake. Additionally, we can increase the plasma density to match the shorter bunch, leading to stronger wakes. See figure 2.5 for a schematic view of the dual advantages of short bunches.

A practical benefit for E164 of using the electron bunch both to ionize and to produce a wake is that alignment issues such as between the laser and the electron beams are removed. Also, there are no UV optics in the electron beam. Thus the experiment is simplified in this regard.

To enhance ionization, we also seek to make the beam smaller in diameter. Currently, the focusing magnets in the FFTB do not allow us to squeeze the beam to be narrower than about 15 μm, but in future wakefield accelerators small diameter beams will be preferred. First is that narrow beams produce stronger wakes as discussed above. Also, we wish to come closer to matching the incoming beam size to that which propagates with unchanged envelope in the plasma, and that requires beams with diameter in the few micron range for dense plasmas; it can be shown that a matched beam produces the least synchrotron radiation over multiple oscillations.

We prefer narrower beams for several reasons, and creating them automatically improves the situation with ionization, which is fixed by the properties of atoms, and can be overcome with better matched electron bunches.

Now that we have discussed the theory of the beam-plasma interactions, we will describe the method of creating the short bunches which are so necessary to E164.

Chapter 3

Producing Short Beams

3.1 Beams in the Stanford Linear Collider

Stanford's main accelerator, the three kilometer linac, has been used for a wide variety of experiments over four decades. Although the machine can produce electron bunches of widely different current profiles and energy spreads, it does impose some constraints on the ability to produce bunches ideal for plasma wakefield accelerators.

E164 is performed in the Final Focus Test Beam Facility, which is directly in line with the main accelerator. This has advantages for shaping beams in well controlled ways. The high quality beams coming straight from the main accelerator have been useful for numerous experiments, and are a reason that the Linac Coherent Light Source will be built on the site of the current FFTB. Figure 3.1 shows the overall layout of the various beamlines at SLAC.

As mentioned in Chapter 1, we seek to use the shortest bunches available. This is because the achievable gradients grow as $1/\sigma_z^2$, so long as the plasma density grows as σ_z^2 to shorten the plasma wavelength appropriately for matching to the beam's length. The $12\,\mu$m bunches are important for this reason.

As we use longer, denser plasmas compared to earlier experiments, the previously used technique of ionizing the Lithium with an ultraviolet laser becomes infeasible because reasonable lasers do not have enough pulse energy to ionize all of the Lithium we require. Additionally, the exponential absorption of the laser energy leads to a

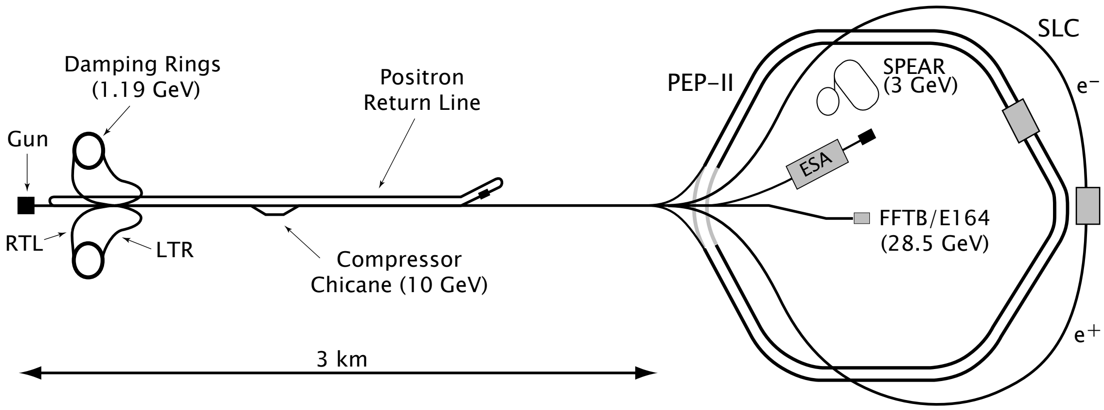

Figure 3.1: Diagram of SLAC with the FFTB facility in which E164 takes place

non-uniform plasma density, even if we had a sufficiently powerful laser. In practice, with our ability to focus the beam to 15 μm spot sizes, we get complete ionization once the bunches from the accelerator are shorter than about 50 μm. With bunches even shorter than this value, the Lithium ionizes at an earlier point within the bunch, allowing more of the bunch to participate in driving the wake. So both for gradient and for ionization, we sought to use the shortest beams available.

To make the short beams, the accelerator has three stages of compression in accelerating and transporting the electrons to the E164 experimental area. Before these manipulations can be performed, the overall phase space of the beams must be reduced through synchrotron radiation cooling in the north damping ring (NDR). The beams are brought there from the gun by the first sector of the linac followed by the linac-to-ring beamline (LTR). After tens of thousands of revolutions in the rings, the electron bunches have suitable emittances for injection into the accelerator. We discuss in detail the linac systems which allow us to create the bunches needed for E164.

3.2 The Damping Ring

Electron bunches from the gun are boosted to 1.19 GeV in the first sector of the linac, and are then extracted for transportation to the north damping ring, where the cooling reduces the phase space substantially. The following discussion summarizes results found in [44] and [45] and uses theoretical treatments in [46].

3.2.1 Single Particle Dynamics

The energy lost to synchrotron radiation in the NDR must be made up by continued boosts from two accelerator cavities located approximately opposite one another and powered by the same klystron. The design of the ring requires that incoming bunches be injected at a phase of the rf which is not far from zero crossing. The bunches will always remain centered at that phase as they circulate in the ring.

The effect on individual particles is more involved, however. The optics of the ring are such that high energy particles take a slightly longer path in completing a revolution than do low energy particles. This is referred to as having positive momentum compaction, α, and this is generally true of storage rings run above the transition energy. A ring operated below its transition energy will have that low energy particles arrive behind the synchronous particle because their lab frame velocities are appreciably lower than c.

In the SLAC damping ring, if an electron lags behind in the bunch, it will receive a weaker boost, and if it gets ahead of the center of the bunch, it will see a stronger boost. This differential acceleration and positive α interact over large numbers of revolutions to constrain the longitudinal phase space in a parabolic potential well. Thus, the electrons oscillate back and forth in both energy and longitudinal position about the center of the beam.

The potential well depth is directly related to the "gap voltage" of the accelerating cavities. The maximum accelerating voltage is 1 MV, and the lower limit is about 600 kV because below that value electron bunches are not reliably captured. R. Holtzapple [45] measured that, at low beam current, the resulting bunch length scales as one over the square root of the gap voltage. This matched the theoretical expectation [46]. A typical operating point is 800 kV, with a corresponding (low current) bunch length given by $\sigma_z \approx 5.0$ mm.

Synchrotron radiation continually takes energy from the particles, and the reference particle will lose 79.2 keV per turn. Differences in energy, however, have several effects on the amount of radiation emitted each orbit. First, for a given bend radius, the synchrotron radiation goes as the fourth power of the electron's energy. Second, the momentum compaction causes higher energy electrons to take a longer orbit, with more total radiation. Third, the optics of the ring are such that off-energy particles go through quadrupoles off-axis, and the net contribution again is that higher energy particles radiate more energy per turn. These three together combine to give an energy dependent radiative loss, and this difference acts like a friction force as particles oscillate back and forth in energy. Based on this largely classical analysis, one might naively expect that eventually the electrons would all settle at the bottom of the potential well, coalescing into a delta function in both energy and z.

The bunches have significant length and energy spread because the synchrotron radiation is quantized. The photon spectrum is very broad, with a critical energy of 1.84 keV, and the stochastic nature of the photon emission acts like a heat bath which continually excites the particles relative to the central energy of the potential well. The competition between this random excitation and the damping properties of differential power emission leads to the equilibrium bunch length described above.

3.2.2 Collective Effects

For highly relativistic bunches, Coulomb interactions are very minor. The primary interaction is indirect through the wakefields left behind by each particle in the vacuum chamber. These result when the beam passes by any change in the shape or cross sectional area of the chamber. At these locations, fields can be excited which act back on the beam.

Within resonant cavities, there is a narrow band of frequencies which get excited by passage of the beam. In most of the damping ring, however, there are no real cavities, just various changes in the beam pipe where fields of many frequencies are deposited. Another way of viewing the situation is that at each of these locations there is a low Q resonator which presents a broad-band impedance to the beam.

For the low frequency components of the spectrum, the impedance has an inductive character, and for the middle frequencies with wavelengths typically of the order of the vacuum chamber discontinuities, the impedance is resistive.

The original damping ring vacuum chamber was constructed of many parts with a variety of locations where the beam could leave resonant fields, such as abrupt changes in the pipe geometry or in the many bellows and flexible joints. The impedance of the ring was largely inductive, with a total inductance calculated to be 33 nH [47]. This led to longitudinal instabilities in the early days of SLC. In 1994, the chamber was replaced with one which minimized sudden transitions in the beam pipe and which removed most bellows and all flex joints. Having lowered the calculated inductance of the ring to 6 nH, the primary component of the overall impedance became resistive.

Inductive wakes have the property that they symmetrically broaden the rf potential well seen by the electrons in going around the ring. Resistive wakes alter the potential well from the rf to be asymmetric without substantially changing its width. The particles therefore move toward the head such that the current passing any given location rises rapidly and then falls off more slowly.

For both resistive and inductive wakes, the effect is stronger as the number of particles in the bunch increases. Interestingly, below the threshold for instabilities, the energy spread does *not* change with the number of particles [44]. Thus, in the NDR, wakes modify the bunch shape, but have minimal effect on the energy spread.

3.2.3 Overall Dynamics

The equations which include the single particle effects of the potential well and the synchrotron radiation loss as well as the collective effects of wakefields can be combined into a Fokker-Planck equation. Solutions to this in the steady state were first investigated by J. Haissinski [48]. Such solutions give the expected longitudinally asymmetric beam with finite energy spread and length.

Figure 3.2 was generated using several Haissinski solutions to our Fokker-Planck equation. The four curves in each graph reflect the changing Intensity Parameter, which depends on the number of particles in the electron bunch. These particular

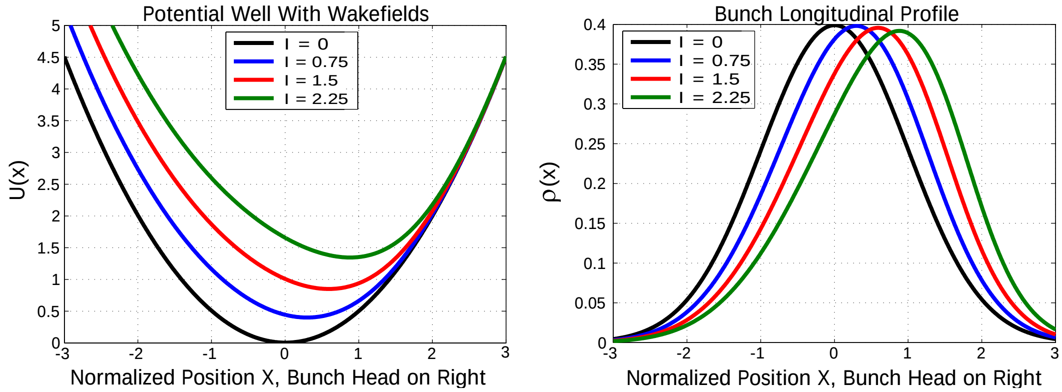

Figure 3.2: The effect of wakefields is to distort the potential well and cause the bunch to tilt toward the head [48].

curves reflect only the effects from resistive wakes, which primarily impart asymmetry, with minimal effect on bunch length. In practice, the bunch is also affected by the inductive wakes which increase bunch length, but have little effect on asymmetry.

We can fit the bunch current profile to an asymmetric Gaussian:

$$I(z) = \frac{N}{\sqrt{2\pi}\sigma_z} \exp\left[\frac{-z^2}{2\,\sigma_z^2\,(1+sgn(z)A)^2}\right] \tag{3.1}$$

where N is the number of particles in our beam and A is the asymmetry factor, with values given by $-1 < A < 1$. In this convention, used throughout this thesis, negative asymmetry means that there are more particles near the *head* of our electron bunch.

We can define σ_L and σ_R for the respective parts of the curve on either side of the peak value at $z = 0$. If σ_z is the average of these two, then we observe that $\sigma_L = \sigma_z(1 - A)$ and $\sigma_R = \sigma_z(1 + A)$. The area under the curve is unchanged by the addition of the asymmetry factor, but the centroid value shifts to $\langle z \rangle = \sqrt{8/\pi} A\sigma_z$.

During SLC running in 1995, R. Holtzapple measured the asymmetry and bunch length as a function of increasing beam charge when the accelerating gap voltage was set to be 820 kV [45]. Additionally, at fixed bunch charge, the bunch length was measured to be a weak function of gap voltage. With a gap voltage for E164 of 790 kV, we expect that at all beam currents, the bunches will be 1% longer than those plotted in Figure 3.3. There are possible systematic errors in these measurements

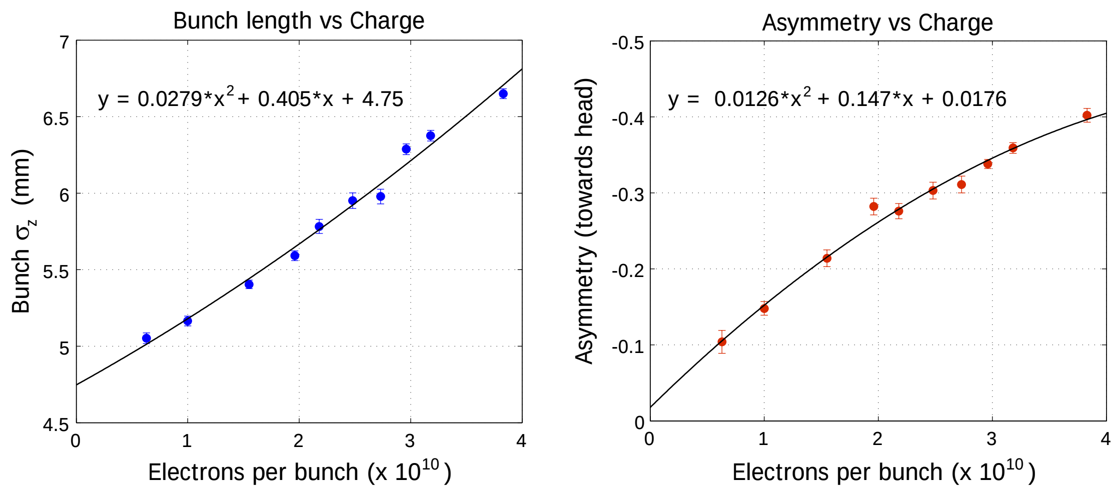

Figure 3.3: Due to wakes, the bunch increases in length and has greater asymmetry toward the head with increasing beam charge. Adapted from [45]

which could be as large as 10 percent, and this dominates the small difference coming from gap voltage. Despite the uncertainties in absolute bunch lengths, we clearly see the expected trend toward slightly longer bunches and increasing asymmetry with more charge as predicted by theory. Given a typical 2×10^{10} electrons, we expect the bunches to be approximately 5.6 mm in RMS length with an asymmetry of -0.27.

3.2.4 Instability Effects

In addition to changing the equilibrium bunch profile of the bunch, wakes can additionally cause potentially destructive collective effects known as microwave instabilities. A general feature of these is that a bunch with high charge will periodically experience a rapid increase in energy spread, and therefore length, before slowly damping back down to the equilibrium values. This has an approximately 1 ms time scale for our ring, and so happens several times during a store cycle. As the timing is random, this instability creates jitter in the bunch length at time of extraction, with attendant fluctuations all the way down the main accelerator.

These instabilities depend on the number of particles in a bunch, having a measured threshold for the NDR near 2×10^{10} particles, and were a source of real concern for the SLC where the desired current was twice that value [44].

Parameter	Symbol	Value
Typical Injected Horizontal Emittance	$\gamma\epsilon_{x\,inj}$	$\sim 150\,\mu$m
Typical Injected Vertical Emittance	$\gamma\epsilon_{y\,inj}$	$\sim 150\,\mu$m
Typical Injected Energy Spread	δ_{inj}	$\sim 1\%$
Typical Injected Bunch Length	σ_{inj}	$\sim 900\,\mu$m = 3 ps
Damped Horizontal Emittance	$\gamma\epsilon_x$	$26\,\mu$m
Damped Vertical Emittance	$\gamma\epsilon_y$	$3\,\mu$m
Damped Energy Spread	δ_0	7.4×10^{-4}
Bunch Length for 2×10^{10}	σ_0	5.6 mm = 18.7 ps
Bunch Asymmetry Toward Head	A	0.28

Table 3.1: E164 beam parameters before and after the damping rings. The incoming values change from day to day, but the outgoing parameters are controlled solely by the properties of the damping ring, and are therefore stable.

Transverse instabilities in the main linac, as discussed below, also grow with bunch charge, so for dual reasons, we were limited to not much more than 2×10^{10} particles. We did observe jitter in the longitudinal properties of the beam from shot to shot, and one possible cause is operation near the microwave instability threshold. In future experimental runs, we can investigate this effect further.

3.2.5 Properties of the Beams After the Damping Ring

The emittances of the beams are all reduced by the cooling action of the damping ring. We summarize incoming and outgoing parameters in Table 3.1.

It is worth saying explicitly that the properties of the beam after the damping ring are completely independent of those coming in. The particles are accelerated by and then radiate away approximately 5 GeV of energy each in going around the damping ring some 70,000 times. The incoming properties are simply overwhelmed, so that the beam's outgoing phase space is solely determined by the properties of the ring as discussed above.

The parameters of the beam as it re-enters the linac matter significantly because we subsequently compress in length by a factor of nearly 500 in several stages. Small changes in these parameters have repercussions down the length of the accelerator.

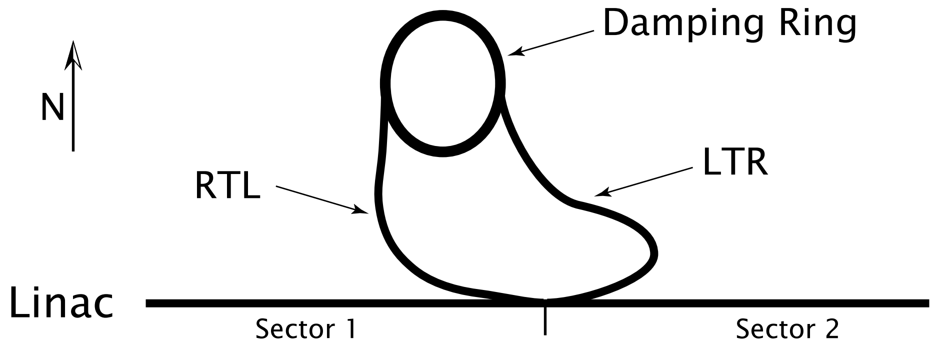

Figure 3.4: Beginning of the linac where the damping ring and RTL are located.

3.3 The Compressor Cavity and RTL Beamline

After the damping rings, the beams are transported back to the main accelerator by the ring-to-linac beamline (RTL). The transverse emittances are greatly reduced in the damping ring, but the bunch length actually *increases* due to the various effects discussed previously. Therefore the beam cannot be injected directly into the linac.

To shorten the bunch, we must induce a longitudinally correlated energy spread, or "chirp," and use the substantial dispersion of the RTL to compress the bunch. The correlated RMS energy spread must be much larger than the intrinsic energy spread from the damping ring to achieve good bunching. After being chirped, the beam traverses a complicated series of optics and bends in returning to the linac, as can be surmised from the shape of the RTL as shown in Figure 3.4. The bends create the necessary dispersion for bunch compression, but great care must be taken with the optics to minimize chromatic effects and preserve the transverse emittances. The specifics of how this is accomplished are discussed in the SLC Design Handbook, Chapter Six [49] and we summarize the main principles.

3.3.1 The Compressor Cavity

To correlate energy and position, the beam passes through an S-band (2856 MHz) compressor cavity. This is located before the RTL, and is designed to run at zero crossing, giving no net energy to the beam. The $2\,\sigma$ central portion of the beam only

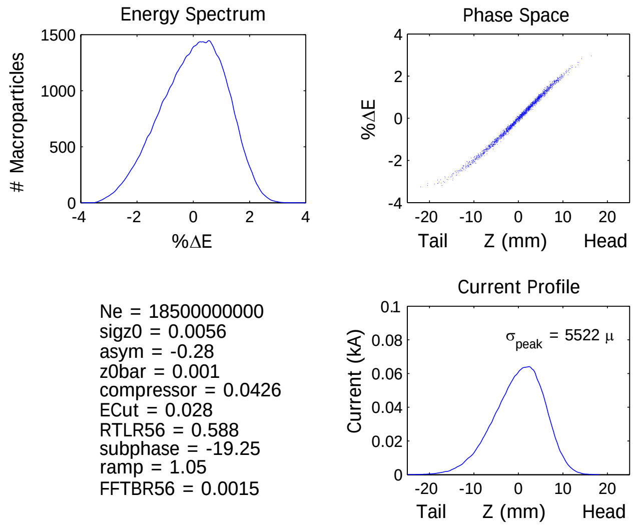

Figure 3.5: Simulation of the beam phase space after the compressor cavity. 100,000 macroparticles represent the beam and they can be histogrammed to provide the energy spectrum and current profile of the bunch.

covers about 38° of rf phase, so the compressor induces a nearly linear energy chirp in the bunches. The magnitude of the chirp is substantial, as the rf amplitude in the cavity is typically between 41 and 43 MV.

In the cavity, the bunch goes from having an uncorrelated RMS energy spread of 0.074% to typically having the particles $1\,\sigma$ toward the head 1.1% higher than the central energy and the particles $1\,\sigma$ behind 1.1% lower. (Recall that in the convention used in this thesis, the head of the bunch is that portion at positive positions, $z > 0$, and the tail is at negative values, where z is measured in the comoving frame.) The induced energy spread totally dominates that coming from the DR, as can be seen in Figure 3.5, which shows a simulation of the beam at this point. We discuss the program for doing such simulations in Chapter 4.

3.3.2 Longitudinal Considerations in the RTL

With the significant bends in the RTL, there is large dispersion throughout the beamline. It is this dispersion that causes the bunch to shorten.

For calculation, it is convenient to use a matrix equation which tells how the properties of the beam are modified by various optical elements. With six dimensions in the full phase space, the general equation will involve a 6×6 "transfer matrix," called **R** by convention. Usually, however, the two horizontal and one longitudinal phase spaces are decoupled, so we can break this into three smaller matrix equations. Our longitudinal phase space is parameterized by position z and fractional energy spread relative to the central energy, written δ. Its modifications through any sequence of optics are calculated with the general equation:

$$\begin{pmatrix} z \\ \delta \end{pmatrix}_f = \begin{pmatrix} R_{55} & R_{56} \\ R_{65} & R_{66} \end{pmatrix} \begin{pmatrix} z \\ \delta \end{pmatrix}_0 \tag{3.2}$$

where the elements of **R** are determined by the specific properties of the magnetic optics and drift spaces in the beamline.

In this terminology, R_{56} has units of distance and is related to the previously defined momentum compaction α in that it is given by αL, where L is the length of the beamline. The matrix equation (3.2) describes the linear effects, but there are also higher order terms which need to be considered for large energy spread beams such as we have in the RTL. The second order matrix is referred to as **T**, and is in fact, three-dimensional. The element of primary interest is that which gives the dependence of final position on the square of initial energy: T_{566}. We can think of these matrix elements as being related to coefficients in the Taylor expansion of the energy dependent part of the transfer function so that $z_f = R_{55}z_0 + R_{56}\delta_0 + T_{566}\delta_0^2 + \cdots$.

Because the energy correlation coming from the compressor cavity is primarily linear, the goal of the RTL is to have a substantial R_{56} with minimal higher order terms, while preserving the transverse phase space.

3.3.3 Transverse Considerations in the RTL

In order to preserve the horizontal and vertical emittances, the RTL has a complicated series of optics for bringing the beams back to the main accelerator. To achieve compression of the beam in the roughly 45 meter long RTL, it is optimal to have the optics present two full betatron oscillations. This means we have two stages, each of which relay images the incoming transverse phase space. Thus, in transport notation, each stage has transverse $\mathbf{R}$ matrices which are just 2×2 identity matrices.

A second order achromat [50, 51] is used because it has symmetries which cause both first and second order geometric terms to vanish. This means that the submatrix of $\mathbf{T}$ dealing with the transverse properties also becomes a (3-dimensional) identity matrix. "Stage 1" fully implements this by using four identical cells of optical elements, the minimum number such that symmetry will cancel all second order geometrical effects. To cancel all second order chromatic effects, e.g. T_{566}, requires only a judicious choice of sextupole strength, as demonstrated by K. L. Brown.

We use the labels "Stage 1" and "Stage 2" for the two halves of the RTL, but these are not actually sequential in the beamline. Due to constraints imposed by the tunnel geometry, Stage 1 is inserted near the beginning of what we call Stage 2. Given this arrangement, the RTL optical setup is referred to as a "nested achromat."

There are a number of features in the Stage 2 which add to its complexity. First is that this section must both extract the beams from the damping ring and also later inject them into the main linac, requiring matching of the optics. Second, the RTL tunnel was unfortunately designed and constructed before the final optics design was finalized, precluding the use in Stage 2 of the full symmetry of Stage 1. Finally, this stage is also responsible for returning the beam to the height of the main accelerator. Allowing for the later possibility of a second ring on top, the damping ring vault was constructed such that the damping ring lies 32 cm below the main beam pipe, and the beams must be deflected vertically to return to the linac.

The second stage of the RTL is itself made from two identical cells which between them have the effect of an identity transfer matrix. Each cell individually, however, has a total $\mathbf{R}$ which is the *negative* of the identity matrix. Each cell requires six dipoles, five quadrupoles, and four sextupoles. All magnets must have the same strength in both cells to cancel second order geometric aberrations and net dispersion.

The chromatic terms are more difficult to cancel, as several of the sextupoles are constrained to be placed at locations with both horizontal and vertical dispersion. These sextupoles must therefore be rotated axially to prevent mixing of the x and y

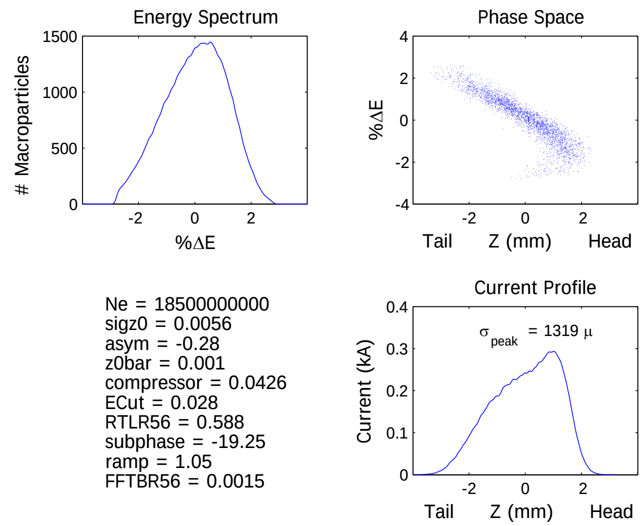

Figure 3.6: Simulated phase space of the beam after compression in the RTL.

plane dispersions with attendant aberrations. Proper *orientation* allows us to zero all of the cross terms between x and y which also depend on energy in the second order transfer matrix. Proper sextupole *strengths* allow us to remove all transverse coordinate dependence on the square of the energy.

3.3.4 Properties of Beams Exiting the RTL

Most of the bunch compression happens in Stage 1 of the RTL, where the design has more freedom to set the optics. Stage 2 is primarily concerned with preserving the beam's transverse properties as they travel back to the main accelerator. With such complicated optics, the energy-position coupling terms are modestly adjustable, and typical values for E164 are that $R_{56} = -0.588\,\text{m}$ and $T_{566} = -1.054\,\text{m}$. The longitudinal position changes for particles with different energy will be dominated by the linear R_{56}, which for even extreme energy particles has 20 times the effect of T_{566}. Any accurate model of the beamline must nonetheless include the latter term for sufficient accuracy.

The negative value for R_{56} is the opposite of what we would expect in a simple dispersive line, and results from the many quadrupoles, whose focusing properties have the effect of altering the dispersion. The R_{56} of -58.8 cm means that a particle having an energy 1% greater than nominal will arrive at the end of the RTL 0.588 cm *behind* an on-energy reference particle.

These path length differences with energy cause all of the particles to collapse to a much shorter bunch with a typical RMS length of about 1.3 mm in traversing the RTL, though the energy spread does not change. Repeated simulation of the whole accelerator to minimize the bunch length at the end of the linac indicates that it is actually optimal to set the compressor cavity and the RTL to overcompress the beam slightly. This means that the bunch is much shorter than at the damping ring exit, but the high energy head overshoots the central position slightly to become the tail as the bunch enters the linac. There is still a correlated energy spread, but it is now reversed in sign. Compare the phase space in Figure 3.6 with that in Figure 3.5.

3.4 To the Sector 10 Chicane

After returning to the main accelerator, the beams are boosted from 1.19 GeV to 9 GeV in Sectors 2 through 6 and then coast until Sector 10. The rf phase is set at $-19.25°$ so that the particles are somewhat ahead of the wave crest. At this position, the average acceleration is still 94% as strong as it would be on crest, but the particles $1\,\sigma$ in front of the beam are accelerated nearly 1.5% less than the central particle, and similarly, trailing particles have greater energy.

3.4.1 Transverse Wakes and BNS Damping

The energy chirp imparted to the beam in the first part of the linac is necessary for bunch compression in Sector 10 as discussed below, but the chirp has unfortunate consequences for the transverse emittances.

The bunches in the accelerator experience both transverse and longitudinal wakes in the linac. For the 1.3 mm bunches at Sector 2, the longitudinal wakes are not a major concern, but the transverse wakes have potential to damage the beam if not carefully controlled. Although nominally outside the present scope, we discuss the mechanism briefly, as it has practical effects on our experiment.

Transverse Wakes

If a bunch passes through an accelerator cavity off-center, the leading particles will excite transverse wakefields which deflect the tail of the bunch. These tail deflections are in the same direction as the original offset, so the overall transverse size and, therefore, the emittance of the beam increases. This instability has some of the same effects on the beam as those from the hose instability in plasmas.

Transverse wakes in the SLAC linac have been investigated in [52]. If we look at Figure 8 from that paper, we see that the wakes in the linac grow approximately linearly for the first 5 ps behind any given charge and then grow more slowly through the first 20 ps until reducing in strength thereafter. Our 1.3 mm RMS bunch has a total length of roughly 4 mm, or 12 ps, so the forces at the tail add up to a substantial kick from all of the particles ahead.

If the beam is off-center at some point in the linac, it will oscillate in transverse position due to the focusing quadrupoles. Although the head of the bunch simply oscillates with fixed amplitude, the oscillations allow the tails to be driven sideways resonantly. With the beam becoming ever larger, the emittance grows.

Perfectly aligning all parts of the accelerator and guaranteeing that the beam is always exactly on axis could theoretically maintain the bunch's emittance, but the tolerances are prohibitive. Instead, a useful technique was used in the SLC to minimize emittance blow up from real-world misalignments.

BNS Damping

The technique to minimize emittance blow-up is called "BNS damping" after the initials of the scientists who first proposed it in 1983 [53]. To achieve best BNS damping, we require that the accelerator be run off-crest to make the tail of the bunch acquire *less* energy than the head, especially at the beginning of the linac where the transverse wakefields are most damaging. Chromatic effects in the quadrupoles mean that the lower energy tail particles are more strongly focused, offsetting some of the effect of the transverse wakes. In the case of the SLC, BNS damping was originally optimized by setting the rf phase to approximately $+20°$ [54].

For E164, due to the requirement that the tail gain energy relative to the head, we run at nearly *minus* 20 degrees (see next section). There is still some damping of the emittance blow-up relative to the worst case where all electrons have the same energy, but it is less than in the ideal case. Great care was needed to center the beam through the accelerating cavities to minimize beam emittance growth. A regular feature of tuning the beam involved tracking down places where the beam was offset relative to the accelerating structures, as well as purposefully inducing orbit deviations to correct other, hard to detect errors.

As the tuning of the linac is never perfect, delivered emittances in the FFTB are substantially larger than those coming from the damping ring. Typical best achievable emittances grow by a factor of 2 or 3 to about $50\,\mu\text{m} \times 5\,\mu\text{m}$ in x and y, respectively. One culprit is probably the lack of BNS damping for this known transverse instability.

3.4.2 Longitudinal Effects

Transverse wakes in the first part of the linac are a problem, but the longitudinal forces between particles are comparatively modest and even useful. These longitudinal wakes have the character that they decrease in strength as the distance behind the source charge increases. In longer bunches, the wakes from the head particles are weaker by the time the tail particles see them than is the case for short bunches. As the longitudinal wakes directly affect the choice of parameters for the accelerator, such as rf phases, we discuss the method of calculating them here.

Wake Calculations

To calculate the wake strengths that we expect, there are a variety of methods one can use. A standard approach is to solve the problem in the frequency domain and then inverse Fourier transform the result to find the fields. For short range forces, we need to include a large number of frequencies to obtain accurate results.

The wakefield is actually calculated for use in LiTrack with a hybrid approach for the frequency domain calculation. We briefly summarize the principles here, but recommend [55] for a more in-depth discussion of the technique. For the low

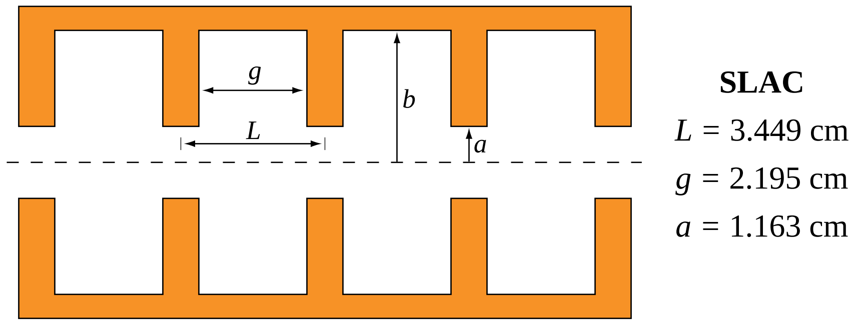

Figure 3.7: Four accelerator cells showing the parameters used to characterize the cell proportions as used in calculating the wake. Our ultrarelativistic electrons never have a chance to see the outer cavity edge, and b does not enter any wake calculations.

frequencies, it uses numerical field matching as described in [52] to find the wave numbers k_n and loss factors κ_n for the first several hundred modes of the structure.

For higher frequencies, an analytic method called the Sessler-Vaynstein Optical Resonator Model gives the dependence [56]. This combines the power spectrum of the beam's field at the iris edges with the diffraction of that power from the edges of an infinite periodic array of thin circular mirrors. In the nomenclature of Figure 3.7, this model assumes that $g \to L$ to make the irises into thin mirrors with holes cut out in the center. Although the irises in the SLAC linac are certainly not infinitessimal in thickness, this model has been observed to agree well with numerical results. The formula for the real part of the impedance at high frequencies is:

$$R_L = \sum_{n=1}^{N} \frac{\pi \kappa_n}{c} \delta(k - k_n) + \frac{2 Z_0 j_{01}^2}{\pi L \psi^2} \frac{\sqrt{\nu} + 1}{(\nu + 2\sqrt{\nu} + 2)^2} \Theta(k - k_N) \qquad (k > 0) \tag{3.3}$$

where $j_{01} = 2.405$ is the first zero of the Bessel Function J_0. With ζ the Riemann Zeta function, the number $\psi = \zeta(1/2)/\sqrt{\pi} = 0.824$. The quantity ν incorporates the specific cell geometry and is given by $4a^2k/(\sqrt{Lg}\,\psi^2)$. Finally, $\Theta(k - k_N)$ is the usual step function, going from 0 to 1 as we move from negative arguments to positive.

After Inverse Fourier Transforming the impedances that are spliced between field matching and the optical resonator model, we obtain a wake plot such as Figure 3.8, which shows the wake behind a charged particle in the SLAC accelerator.

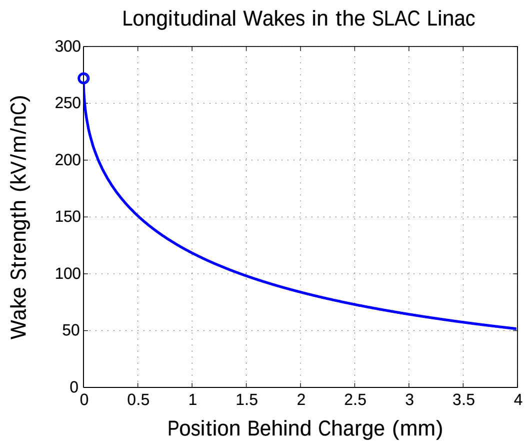

Figure 3.8: Plot of wake strength behind a charge in the SLAC linac. The empty circle shows the maximum field immediately behind the source charge. This calculation [57] was performed in developing the simulation code LiTrack discussed in Chapter 4.

3.4.3 Phase Space After the First Third of the Linac

When we combine the effects of the rf phasing and the wakes, we have a fortunate effect which causes a nearly linear chirp in our beam. There are no wakes at the very head. Then the wakes become strongest in the center of the bunch and have less effect on particles in the tail. This induces a curvature to the energy spread which nicely cancels most of the effects from rf curvature.

Figure 3.9 shows the phase space that is produced at the end of the first third of the linac. We additionally show a second plot of the phase space to afford a direct comparison of what the beam looks like with and without the effect of wakes. We can see that the wakes cause the energy correlation to become much closer to linear.

We recall that after the RTL, the energy spread has an RMS width of about 1.1%, as mentioned above. Being overcompressed in the RTL, there is still an energy-position correlation present in the bunch, though it is reversed from that imparted by the compressor cavity. A particle $1\,\sigma$ in front of the center of the beam returns to the linac with 12 MeV less energy than the central particle. This energy correlation has the same sign as that which develops as the beam propagates down the accelerator.

With the eight-fold energy gain in the first third of the linac, the incoming energy chirp is dwarfed by the energy spread ultimately developed. With the combined

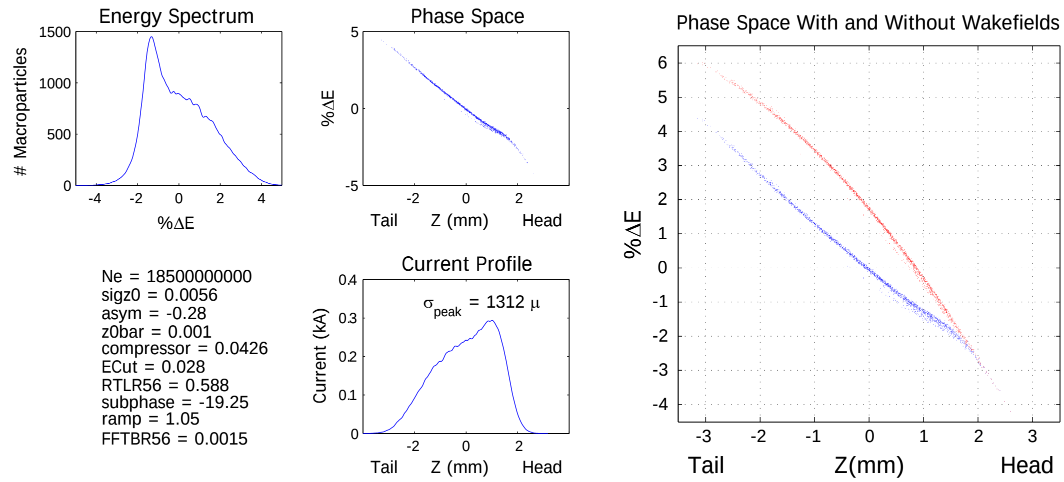

Figure 3.9: Phase space of the beam after traversing the first third of the linac. For comparison, the right hand side shows the beam energy if there were no wakefields in red compared to the actual energy when the wakes are included in blue. The peak loss from wakes is 180 MeV, or about 2%.

effects of the acceleration and wakes, we induce the needed linear chirp for the second stage of compression, and the final correlated energy spread has total width of nearly 8 percent at 9 GeV. At the end, the $1\,\sigma$ leading particle is low by about 1.4%, or 125 MeV, as can be seen in the two phase space plots of Figure 3.9.

3.5 The Sector 10 Chicane

In 2002, a magnetic chicane was added in Sector 10, about one third of the way down the accelerator. It uses the substantial energy correlation of the beam to compress it in a second stage, and can produce beams as short as about $50\,\mu$m. This device was designed with relatively gentle bends such that it has $R_{56} = 7.6$ cm with no significant higher order terms [58]. The positive sign for R_{56} means that the more energetic tail of the bunch takes a shorter path and catches up to the central orbit, while the low energy head lags back toward that same central position. Figure 3.10 shows a schematic of this second stage of compression.

From Table 3.2, we can see that the 9 GeV beams bend with a radius of curvature of 18.52 meters, corresponding to a magnetic field strength of 1.62 Tesla. Our beam is

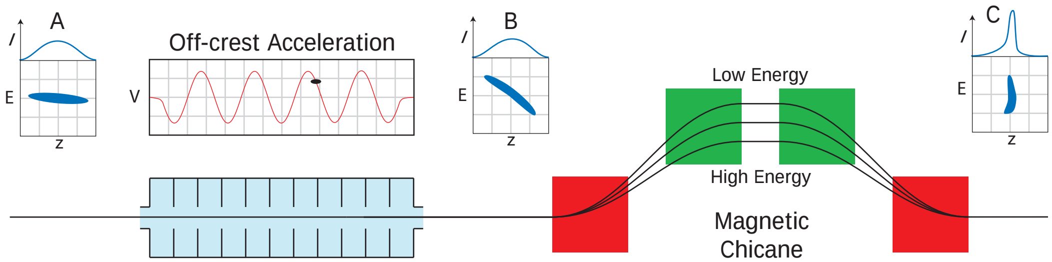

Figure 3.10: Diagram of the Sector 10 Chicane which compresses the bunch down to less than 100 μm using the energy chirp produced in Sectors 2-6.

Parameter	**Symbol**	**Value**
Momentum Compaction	R_{56}	7.60 cm
Peak Dispersion	$\widehat{\eta_x}$	44.80 cm
Dipole Radius of Curvature	ρ	18.52 m
Magnet Length	L_B	1.8 m
Bend Angle	θ	5.57°
First and Last Drift Lengths	ΔL	2.8 m
Central Drift Length	ΔL_c	1.5 m
Overall Chicane Length	L_T	14.3 m

Table 3.2: Sector 10 Chicane Parameters

even more highly relativistic than in the damping ring, and in such a strong magnetic field, it will emit substantial synchrotron radiation in each bend.

As in the damping ring, the stochastic nature of this emission leads to an increase in energy spread and also to an increase in the transverse emittances of the beam. Both coherent and incoherent synchrotron radiation are emitted with such short bunches and the combined effect is to increase the horizontal emittance by about 20% [58, 59].

3.5.1 Beam Properties After the Chicane

After the beam is compressed, it still has a large overall energy spread. Now it is quite short compared to the beams that enter the linac, having a typical overall RMS length of about 70 μm. The bunch is no longer well approximated by a pure Gaussian, so

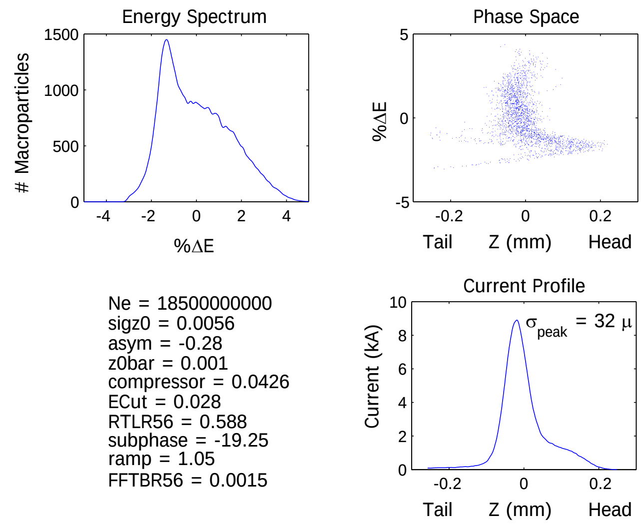

Figure 3.11: Phase space of the beam after the second stage of compression in the Sector 10 chicane.

there is a strong central peak in the current distribution with lower current wings on either side. If we fit a Gaussian just to this central peak, it has a typical $\sigma_z \approx 35\,\mu$m for bunches that are well compressed. As this central portion contains a large fraction of the bunch's overall charge, it is a useful metric and will largely govern the wake behavior in the plasma.

With a large compression ratio in the Sector 10 chicane, the bunch lengths downstream are sensitive to small changes in the accelerator parameters. Just as operation near the microwave instability thresholds in the damping ring is a possible source of jitter on the bunch's longitudinal parameters, variations in the RTL and linac phasing before the Sector 10 chicane can affect the compression significantly.

For example, as is discussed in Chapter 4, the phase of the accelerator regularly varies in a range of a degree or two. In Figure 3.11, we see the phase space after the chicane for the accelerator conditions we have been using to illustrate the evolution of the phase space down the accelerator. The bunch is close to as short as can be produced at this point, with a σ_z of $32\,\mu$m. However, if the rf phase in the linac is changed by $-0.5°$, the RMS width of the central peak grows to $42\,\mu$m, and if the phase changes by $+0.5°$, the RMS width nearly doubles to $56\,\mu$m.

3.6 End of Linac and FFTB

There is still one final stage of compression that takes place after the Sector 10 Chicane. This happens in the FFTB itself and is made possible by effects in the final two thirds of the accelerator. There are a number of considerations which affect our ability to compress one final time, and we discuss them here.

3.6.1 Transverse Wakefield Considerations

Fortunately, as the beam energy increases and the bunch shortens, transverse wakefields are reduced in importance. As discussed in § 3.4.1, the transverse wakes grow approximately linearly in time after a particle passes, so short bunches have few particles available to be deflected by the slowly growing transverse fields.

After the compressor cavity, Figure 3.11 shows that nearly all of the charge is now contained in a length of 0.3 mm $\approx$ 1 ps. The bunch is more than ten times shorter than before so the transverse wake should be ten times smaller. Further, the beam is more relativistic, so transverse kicks have less effect. Lastly, the correlated energy spread developed in the final two thirds of the linac is favorable for BNS damping.

With all of these effects, the transverse wake effects in the end of the linac are minor, as was observed when actually tuning the machine to produce good bunches. The vast majority of the effort was expended at the beginning of the accelerator in reducing transverse wakefields by careful steering of the orbit. After compression in the Sector 10 chicane, our bunches become short enough that the transverse wakefields are weak, but the longitudinal fields are stronger.

3.6.2 Longitudinal Bunch Manipulation

As indicated above, the energy spread after compression in the chicane is such that we have a (now) largely uncorrelated energy spread with RMS of about 125 MeV = 1.4% entering the final two thirds of the linac. The longitudinal wakefields depend on z within the bunch, and are strong enough to induce a correlated energy spread even larger than the total energy spread coming from the first third of the linac.

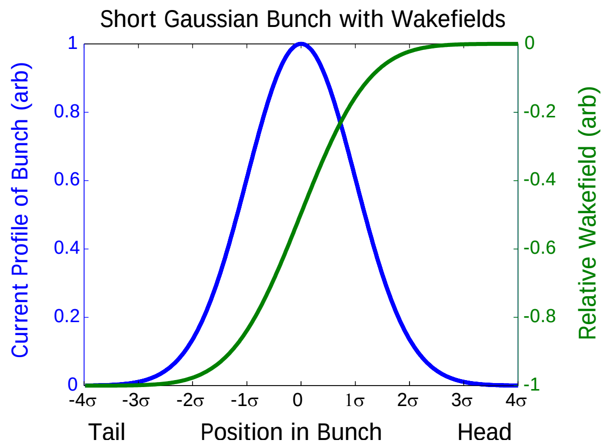

Figure 3.12: The longitudinal wake in a Gaussian bunch which is very short compared to the distance in which the wake decreases appreciably behind a given particle. The wake thus has the shape of the error function, with greatest deceleration in the tail.

As we have seen in Figure 3.8, total longitudinal forces will grow as we compress the bunch to shorter and shorter lengths, the opposite of what happens transversely.

Use of the Wakes

We make these strong wakes into a virtue, however, by using them to imprint a final correlated energy spread on the beams for the last stage of compression. As the bunches have an RMS length about 70 μm, effects from 10.5 cm wavelength rf's curvature are negligible. Thus, in the final two thirds of the main accelerator, the bunches can be run near the rf crest to maximize energy gain. It is solely the wakefields themselves which impart the energy correlation for the final stage of compression.

Our bunch is so short that all of the particles are within a longitudinal region only 0.4 mm long. Thus, all particles in the bunch behind any given electron see a wake which ranges from about 160 to 280 kV/m/nC. With a total bunch charge of about 3.2 nC, we would expect wakes up to about 800 keV/m in the tail, compared to the accelerating gradient, 20 times larger at 17 MeV/m. These wakes acting on a beam should be able to induce total energy spreads of about 5%, or 1.4 GeV. Even so, these strong wakefields do not imprint a perfectly linear chirp and do not completely overwhelm the uncorrelated energy spread from after the Sector 10 chicane.

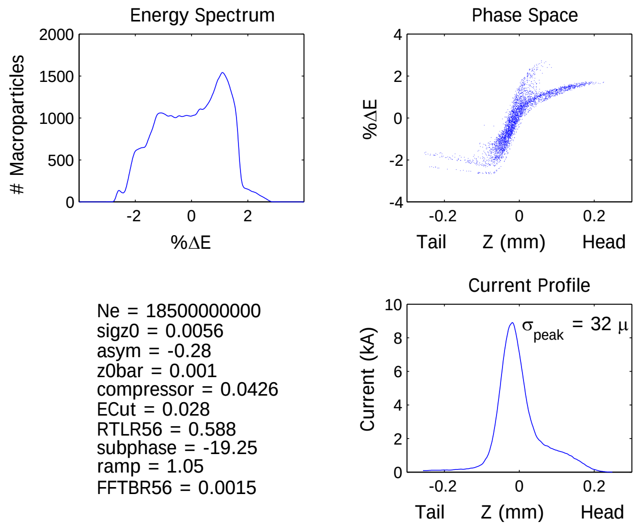

Figure 3.13: Phase space of the beam at the end of the linac.

Most of our bunch particles are in a region only 100 μm long, and the wake field behind any particle drops by only 25% in that distance. Thus, the wake experienced by any given particle is roughly proportional to the integral of the charge ahead of it.

This integral will impart a chirp which looks similar to the error function for a bunch with a Gaussian profile. Figure 3.12 shows the idealized case of the wake progression in a Gaussian bunch where the wake is constant behind any given particle. In the central region, this causes a largely linear chirp.

For our real beam, the wake is not constant over the full length of our bunch, and does decrease noticeably for the tailmost particles due to their comparatively large distance from most of the wake producing charges. They are not decelerated as strongly as the electrons closer to the point of peak current.

A plot of δ versus z thus looks like an "S" sheared to the right such that the central portion still has a positive slope. Figure 3.13 generally shows such a shape, though other features come from the complicated phase space changes before the Sector 10 Chicane: the tailmost particles have slightly higher energies than the particles at the immediate rear of the bunch's central portion. Although not perfectly linear, the 1.1% RMS correlated energy spread at the very end of the main linac allows a final reduction of nearly 3 times in σ_z of the central region and a doubling of peak current.

3.6.3 Compression in the FFTB

This final energy chirp allows a last stage of compression. Clearly, with a substantially nonlinear energy chirp, we cannot compress the beam by as large a ratio as in previous stages. Nonetheless, even the factor of about three that we achieve is important for the plasma experiments we do. Not only does it assure that we can make bunches short enough to ionize Lithium consistently, but it allows us to use dense plasmas in which we can drive the strong wakes needed to achieve well over 1 GeV/m gradients.

The dispersion comes from a horizontal dogleg whose first bend is near the extreme upstream end of the 200 m long FFTB. There are a number of quadrupoles in the dogleg, so its local dispersion is substantially adjustable. As a result, we can modify the R_{56} from -1.0 to -2.0 mm with a straightforward change of quadrupole strengths. This seemingly small R_{56} is all that is required to compress our bunch because the fractional energy spread is substantial, and the required relative position changes of various particles are very small to achieve full compression.

If we bin the above shown beam in z, we see that even infinitessimal slices have a total energy spread approaching 1%. The centroid energy of the slice $1\,\sigma$ ahead of the slice where the peak current is located has an energy about 1.25% higher than the slice on peak. The slice trailing by $1\,\sigma$ has its centroid about 1% lower. A typical R_{56} of -1.5 mm will move these slices toward the center by between 15 and 20 microns. For this particular example, larger R_{56} would compress the bunch slightly more, but this is not always true for our many possible different short bunches.

In this example, the width of the central peak in the electron bunch shrinks to $13\,\mu$m with an associated peak current of 19 kA. This is well above the threshold needed to ionize Lithium with a $15\,\mu$m diameter spot size, and can drive a strong wake in even the densest plasmas so far available to our experiment.

The procedure for creating short bunches is complicated, and our parameter space has many dimensions. We may in future discover a way to make even shorter bunches.

For example, we have implicitly assumed that operating with the highest possible charge in the bunch (up to the point of encountering instabilities in the damping ring) was optimal. Given the many stages of compression, it may prove that reducing the charge could ultimately allow for a shorter bunch. With less charge in the damping

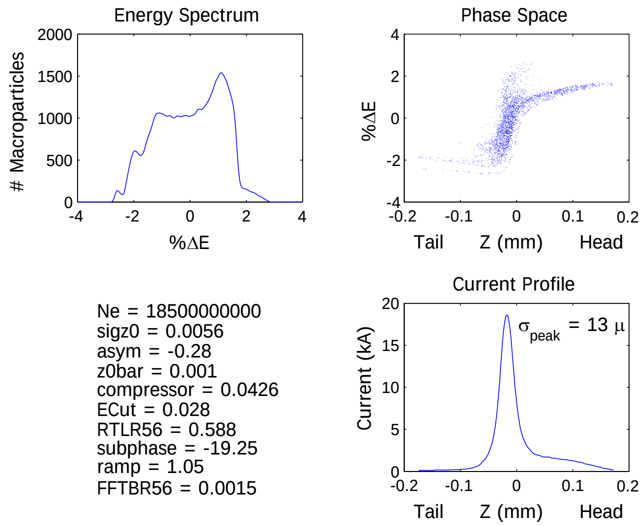

Figure 3.14: Final compressed phase space of beam before the plasma.

ring, the bunch shortens somewhat. With a shorter bunch, the energy chirp from the compressor cavity will be more purely linear for the first stage of compression in the RTL, the R_{56} and T_{566} of which might need to be adjusted to deal with the different beam. The first third of the linac would then need to run at a different phase to balance the different wakefields with the rf curvature to again create a linear chirp for the sector 10 chicane allowing better compression there. With more compression at sector 10, the final energy chirp will be different, possibly requiring a new setup in the FFTB optics.

In this hypothetical case, all of the parameters in the linac that we have discussed would need to be adjusted once we chose a charge for the bunch. It is not easy to visualize, or even calculate analytically, what combination of total charge and linac parameters would lead to bunches with the highest absolute peak current. An exhaustive parameter search to find the shortest possible bunches and highest peak currents as a function of beam charge has not yet been performed.

Nonetheless, the beams we produce are extremely short and intense, rivalling the most powerful laser pulses available and allowing very impressive accelerating gradients to be achieved over 10 cm (see Chapter 5). If we find conditions to produce shorter bunches, then even more impressive gradients will be possible at SLAC.

Chapter 4

Apparatus and Techniques

To perform the measurements in this thesis, a variety of systems installed in the FFTB were important. We first discuss the hardware common to all aspects of E164: the plasma source, the diagnostics for understanding the beam transversely, and the Cherenkov based energy spectrometer which is the central diagnostic for acceleration.

In 2003, we added new hardware at the upstream end of the FFTB. This location was advantageous because there was space available for our 3 ton magnet and associated hardware, but also because there is substantial horizontal dispersion, as discussed in § 3.6.3. Using this magnet, as we discuss below, allows the *non-destructive* measurement of the electron beam's energy spread before it enters the plasma.

We compare these energy spectra, for every data shot, with simulations of the main accelerator. This enables us to determine the phase space for every beam as it enters the plasma and provides greater power in understanding our acceleration and other effects, as discussed in Chapter 5. The specific hardware and spectrum measuring techniques, as well as a discussion of the software we use are the core of this chapter, but first we discuss the existing hardware common to all parts of E164.

4.1 Plasma Source and Diagnostics

The acceleration experiments use a Lithium plasma. The intense electric fields of our electron beams are so strong that they will fully ionize Lithium vapor quickly

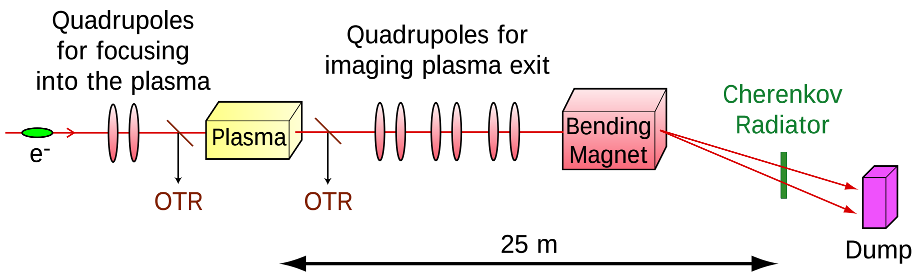

Figure 4.1: Schematic of the plasma cell with diagnostics for understanding the electron beam before and after the plasma. Light from the OTR foils is imaged to view the transverse size of the beam and a series of electron optics image the plasma exit to the Cherenkov radiator while dispersing the beam to measure the spectrum.

compared to the transit time of the beam, as discussed in Chapter 2. We provide a column of Lithium vapor with variable length and density in order to investigate acceleration under a variety of conditions.

After the beam's energy is changed by its interaction with the plasma, we measure the final energy spectrum as a way to understand the plasma's effect on the beam.

To aid in properly focusing the electron beam into the plasma, we also use several beam diagnostics immediately before and after the plasma. We discuss these first.

Figure 4.1 shows a schematic of the beamline at the end of the FFTB where our plasma experiments take place. The major features of relevance are the OTRs, the plasma source and the spectrometer.

4.1.1 OTR Diagnostics

To take an image of the electron beam's transverse components both entering and exiting the plasma, we insert thin metal foils in the beam path to create optical transition radiation (OTR).

A general discussion of transition radiation can be found in [60]. When the electron beam transits the foil, the collapsing dipoles created by the electrons rushing toward their image charges in the foil radiate light across a broad spectrum. As with synchrotron radiation, wavelengths short compared to the bunch length are radiated incoherently, and longer wavelengths will have a coherent nature.

We focus on the incoherent portion for this measurement, as it lies in the visible and is easily viewed with standard cameras. Also, incoherent radiation gives a direct image of the beam, because each electron has an equal probability of radiating, while coherent radiation processes, by definition, are affected by groups of electrons. We want to know where each electron is independent of the others in order to build up a picture of the beam profile.

Radially polarized radiation is emitted from the point where each electron enters the metal and then spreads out from that point. The distribution of the radiation into a given solid angle is only dependent on θ, the angle of the radiation away from the central axis. For ultrarelativistic beams impinging on an infinitely wide metal foil with a large (real portion of the) dielectric constant, the formula for the radiation intensity at a given photon energy and into a given solid angle is, in CGS:

$$\frac{d^2 I}{d\omega d\Omega} = \frac{e^2}{\pi^2 c} \left| \frac{\sqrt{\epsilon}-1}{\sqrt{\epsilon}+1} \right|^2 \frac{\theta^2}{\left(\theta^2 + {}^1/_{\gamma^2}\right)^2} \tag{4.1}$$

The angle where the peak of the radiation happens is, perhaps unsurprisingly, given by $1/\gamma$. In order to have the radiation come out to the side where we can view it, the foil is mounted at 45° to the beam's direction of motion. The radiation thus comes out centered on the direction perpendicular to the electron beam path.

There are several concerns that have been raised about using OTR to produce an image of a beam. The first is that the radiation pattern has the above mentioned peak at $\theta = 1/\gamma$. If this value represents the angular width of the radiation distribution, the resolution limit that we could image would go as $\gamma\lambda/2$, and for visible light with our highly relativistic beam, we could not resolve points on the metal foil less than about 1 cm apart.

Fortunately, this conclusion is actually wrong by a large margin. As discussed in [61], there is still substantial power per solid angle radiated to angles much larger than $1/\gamma$, as the intensity only drops as $1/\theta^2$. Additionally, the total solid angle to be radiated into grows linearly with the angle, so the number of photons coming out at a given angle only drops as $1/\theta$. A more useful metric for the distribution of this light than the naive value of the peak location at $1/\gamma$ is therefore the RMS spread of the photons in angle, $\sqrt{\langle\theta^2\rangle}$. We calculate this quantity for (4.1), and calculate out to some $\theta_{max} \gg 1/\gamma$, where θ_{max} is governed by the actual angle subtended by the optical element being used to capture the light. We find the surprising result that the effective width of the beam at the front of the lens depends primarily on the size of the lens we are using, even though $1/\gamma$ is very small, because there is still so much power out at large angles:

$$\sqrt{\langle\theta^2\rangle} = \frac{\theta_{max}}{\sqrt{2\ln(\gamma\theta_{max})}} \tag{4.2}$$

To capture OTR light, we have a camera lens of roughly 5 cm diameter situated about 25 cm away from the screen. The angle subtended by the camera is therefore $\theta_{max} \approx 0.2$, and for our case, the resolution is about *2500 times* better than expected from naively taking that the characteristic angle is governed by the angle at which the highest flux of photons is emitted, $\theta = 1/\gamma$. We thus expect that for 1:1 imaging in our optical system, we will have diffraction limited resolution of $5\,\mu$m, which is smaller than the pixels on our camera, and we can see the beam's profile clearly. Images of the beam have shown features with roughly this size scale, so the resolution has been as good as expected.

It is worth noting that, unlike the naive case, where resolution gets worse linearly with increasing particle energy, this function depends only very weakly on the relativistic factor of the particles, going as the square root of a logarithm. Therefore, not only is OTR useful in the current SLAC linac, but it should be useful even as electron energies go much higher in the future.

A second concern about using OTR for beam imaging is that there is a distance from the source of radiation, known as the formation length, L_f, within which the Coulomb field of the bunch moving at some v and the radiation field moving at c are not well separated. This is equivalently the distance scale within which we are in the near field, and paraxial optics do not fully apply. This length grows rapidly with the energy of a relativistic particle and is given by $L_f \approx 2\lambda\gamma^2$ [62]. For optical radiation at 500 nm from our 28.5 GeV beam, this analysis predicts that $L_f = 3.1$ km! Our camera is only 25 cm away from the source of radiation, and is thus very much in the near field. There certainly remain theoretical questions about exactly what happens with the radiation impinging on a camera in the near field. However, earlier experiments measured the beam's transverse size with a wire scanner while simultaneously looking at OTR images of the beam nearby. The sizes were consistent and imply that the possible resolution was not compromised significantly from being in the near field.

The beam has very high peak intensity, so an appropriate material for the foil is required. The main damage mechanism for substances placed in the beam comes from the heat load of impact ionization of the material. The rapid deposition of heat can cause small cracks between neighboring crystals from the differential expansion in a kind of shock. The total amount of heat can also simply melt a hole in the foil.

To minimize the ionization, we turn our gaze to metals with low atomic number. Naively obvious choices from the first few members of the Group 1 and 2 metals of the periodic table are ruled out because of toxicity (Beryllium), brittleness (e.g. Calcium) or extreme chemical reactivity (e.g. Lithium). Additionally, the material will heat up rapidly at the point where the beam transits, requiring a high melting point. Lastly, great stress is induced between locations of high and low heating, requiring high mechanical strength. Aluminum has a low melting point and is not particularly strong, so the first good choice is Titanium, which can be made into very thin foils.

Our foils were obtained from Arizona Carbon Foil Company and are grown by vapor deposition on a polished substrate to a thickness of $1\,\mu$m. The substrate is then removed to leave only the foil. Each is tensioned slightly and vacuum epoxied to a stainless steel ring for support.

Our expectation that they would be suitable was borne out, as they have held up well when hit by the beam.

To view the OTR light, we use Photometrics SenSys 12 bit cameras with 768×512 pixels, where each pixel covers an area of 9×9 microns. To capture the light, we use a Nikon f/2.8 AF Micro-Nikkor 105 mm lens. At 25 cm, the lens captures enough light that we expect tens of billions of photons to be collected, for a strong signal. We set the lens to the 1:1 imaging mode such that the pixels on our CCD each see that same area on the Titanium foil. We have plenty of light, and in seeking to image with high resolution, we employ a blue filter to reduce chromatic aberrations. There is also a linear polarizer in front of the lens, as it is expected to improve the resolution in the plane perpendicular to the direction of polarization somewhat [63]. As with other cameras in the experiment, a computer reads out the data at 1 Hz.

We can measure the beam's size both before and after the location of the plasma, where the beam is nominally focused. We have the ability to change the final focusing quadrupole strengths, and can observe the changes in spot size on the upstream and downstream OTRs as the focal point moves. This allows us to verify that we have placed the focal point at the entrance to the plasma as we intend.

Also, any transverse tails that the beam has coming from the main linac will be visible on the upstream OTR camera, and are an indication of input conditions where the beam has a tilt. In § 5.3.2, we see the utility of knowing the transverse properties of the beam before it enters the plasma.

The downstream OTR camera provided the first indication that we appear to be trapping electrons from the plasma itself and accelerating them. We discuss this effect further in § 5.1.1.

These two OTR based diagnostics are an important part of understanding the focusing and other transverse properties of our beams in the vicinity of the plasma oven, and even pointed to unexpected phenomena in the plasma.

4.1.2 Heat Pipe Oven

As we ionize a neutral gas with the beam itself, there are several possible ways to hold gas in a specific region for use in the plasma wakefield experiments. The most obvious simply uses thin metal foils at either end of a section of pipe, into which gas can be pumped in or out to change the density.

When we tried to use such a setup briefly in E164, we found that the downstream metal foils were rapidly punctured, allowing the gas to escape. Damage happened sometimes as fast as after a few minutes.

We therefore used a heat pipe oven, which uses temperature differences and a buffer gas to constrain Lithium vapor to a specific region. The buffer gas is kept from the main vacuum systems by thin Beryllium foils, which are situated a meter or two away from the regions with plasma, and are therefore not damaged.

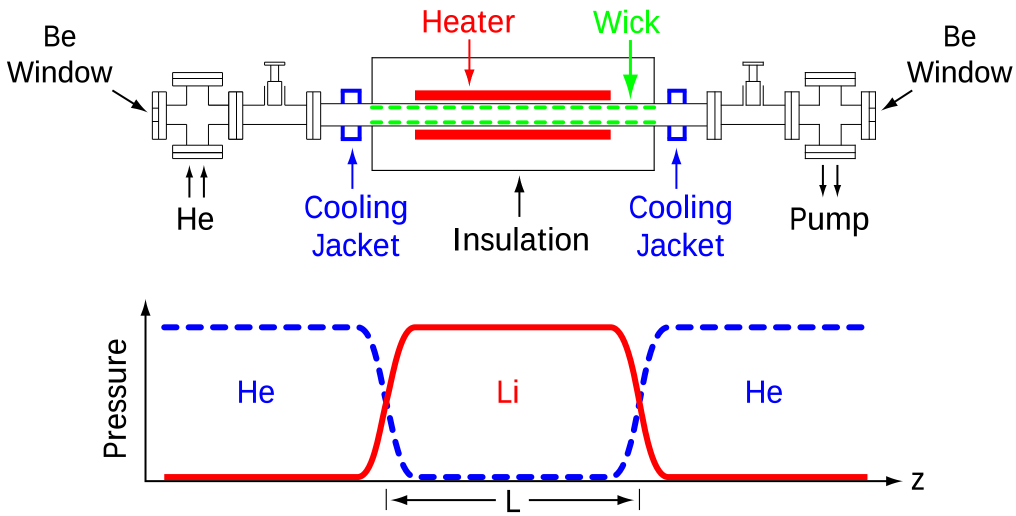

Figure 4.2: Diagram showing the essential elements of our heat pipe oven.

The theory behind the heat pipe oven is described by its original designer C. R. Vidal in [40]. In our implementation, we place a few grams of Lithium inside a pipe which is about 3 cm in diameter and about 50 cm long. We maintain the entire inside of the pipe under an inert atmosphere of Helium gas. The Helium prevents any chemical reactions with the Lithium and is also central to the operation of the heat pipe oven.

With Lithium in place, we heat the central region of the pipe. This heating causes the Lithium to melt and evaporate at a rate dependent on the temperature. Outside this heated region, there are cooling collars wrapped around the pipe at either end to maintain the Helium gas on either side at something close to room temperature.

In this arrangement, a schematic of which can be seen in Figure 4.2, the Lithium evaporates and flows either direction away from the center of the heated region. This flow is at something like the speed of sound, and the rapid Lithium flow simply sweeps all Helium atoms from the central region.

Eventually, the Lithium encounters the cold Helium gas at either end and cools off rapidly. Having cooled, the Lithium condenses on the sides of the pipe where a mesh is installed. The molten Lithium is then drawn back toward the heated central region by the mesh's wicking effect where it again evaporates.

There is a substantial flow of heat from the central regions toward the Helium, but no net transfer of material. We have a central region with essentially a pure Lithium vapor, then wings on either side in which the Lithium density progressively drops as the Helium fraction rises, then the outer regions of pure, cold Helium.

The two parameters of interest concerning the oven are the density of the Lithium vapor inside and the length of the region of vapor. The density is simply controlled by the pressure of the Helium buffer gas. If we evaporate enough Lithium to expel the Helium, the so-called "oven condition," then the Lithium pressure must equal that of the room temperature Helium gas outside.

In order to evaporate enough Lithium for the oven condition, sufficient power must be supplied to the heater. Then by adding more power, the Lithium will flow further before finally condensing, leading to a longer oven. It is worth saying explicitly that once the oven condition is reached, heater power has no effect on the Lithium vapor density, only oven length.

4.1.3 Determining the Oven Length for E164

We seek to know the length of our oven as a function of the vapor density and the heater power that we are using. There are three regions, as mentioned above, all with different properties. In the low temperature regions, there is no Lithium vapor, in the center of the oven, there is only Lithium, and there are the two regions with varying Lithium density, often referred to as the "wings."

When measuring the changing Lithium density profile along the length of the oven, it is not feasible to measure the changing partial pressure of the Lithium vapor directly. However, we can measure the temperature at each point with thermocouples; there is a well defined relationship between the vapor density of Lithium and the temperature. This relationship has been found empirically for a variety of metals in the first column of the periodic table [64]. The equation for Lithium is given by:

$$P = \exp\left(7.530 \times 10^{-4}\, T - 2.053\, \ln T - \frac{1.943 \times 10^{4}}{T} + 32.60\right) \tag{4.3}$$

where the pressure P is in Torr and the temperature T is in Kelvin. Figure 4.3 plots

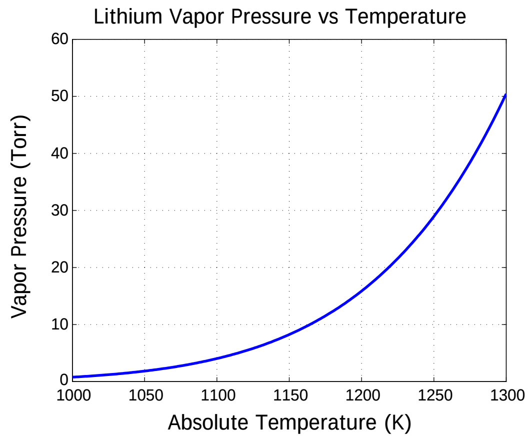

Figure 4.3: Plot of Lithium vapor pressure versus absolute temperature in the region of interest for E164.

this equation for the temperature range of interest in E164. At these temperatures, a vapor density of $3.5 \times 10^{17}\mathrm{cm}^{-3}$ corresponds to a pressure of 46 torr. Even our highest vapor pressure is less than one tenth atmospheric, and the ideal gas law applies very well in this regime. Knowing the temperature and the pressure, we can calculate the density from the standard relation $PV = nRT$.

The temperatures of interest are very high, and limit the ability to increase the Lithium density. We chose to make the pipe from A.I.S.I Type 310 stainless steel, the properties of which are discussed in [65]. This Chromium-Nickel steel has the best tolerance for elevated temperatures of the standard stainlesses. When used at a constant high temperature, it can be used to 1150°C, but when cycled repeatedly from room temperature, the maximum temperature that is recommended is 1040°C. The damage mechanism for repeated heating seems to be that oxide layers form on the surface at high temperatures and then fall off due to differential thermal expansion whenever the material is cooled.

Scaling can eventually weaken the pipe, and constrains our ability to produce high density plasmas. An adequate margin of safety to prevent rupture of our pipe – with molten Lithium inside – meant that we limited the highest temperatures achieved to about 1020°C, giving a vapor pressure of 46 Torr and density of 3.5×10^{17} cm^{-3}. We have observed no scaling of the pipe so far from operation at these temperatures.

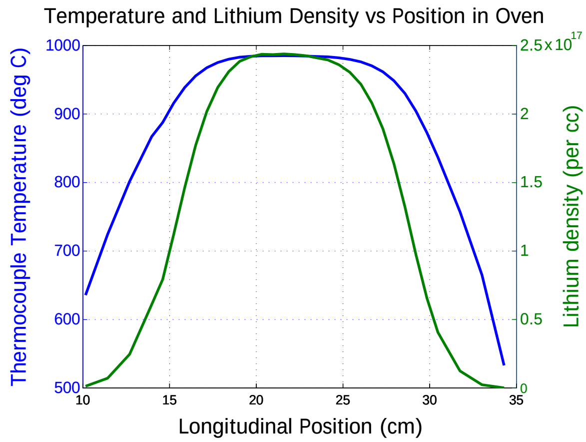

Figure 4.4: Oven profile plot at 2.4×10^{17} cm^{-3}. The FWHM is 13.2 cm.

To measure the oven's vapor density profile, we mount the oven on a stand where we can insert a thermocouple into the middle of the oven and pull it through to record the temperature profile along the longitudinal axis. We can use the above formula to convert to partial pressure of Lithium and determine the oven length from the FWHM of the density profile.

There are wings on either side where the Lithium density drops to zero. Interestingly, the width of these wings appears to be independent of density over the range investigated for E164. The distance from the half density point to where full density is achieved is about 5 cm for all cases. Figure 4.4 shows a sample plot for a peak density of 2.4×10^{17} cm^{-3}. There is a central flat topped region and then we can clearly see the wings on either side. A useful metric is the FWHM, which is 13.2 cm.

There is a slight asymmetry to the curves, and we attribute that to the fact that the thermocouple was at the end of a long metal rod along the central axis of the oven. That rod can conduct heat, causing the thermocouple to read slightly higher temperatures than it should when the rod is extended through the central region. Similarly, the thermocouple reads slightly lower temperatures than expected when the portion of the rod nearby is in cooler regions. This effect appears to be small, but could give a systematic error on the measurement of the FWHM.

For the several densities of particular interest to the investigations of this thesis, we measured the oven's length at several values of the heater power.

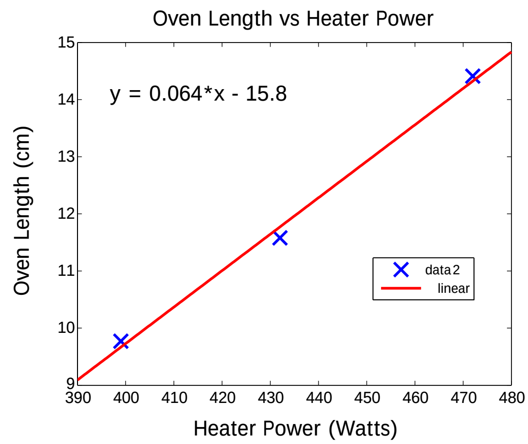

Figure 4.5: Plot of changing oven FWHM with increasing heater power for an oven with plasma density 1.5×10^{17} cm^{-3}.

Figure 4.5 shows the change in oven FWHM as a function of the heater power. The measurements for this were taken at UCLA, where the oven was originally built. The laboratory there is maintained at or below about 20°C. In the FFTB tunnel, the temperature is often closer to 40°C, so we speculate that the same heater power will lead to a somewhat greater oven length than plotted here. This is because the oven is surrounded by a thick layer of insulating bricks, but substantial heat nonetheless flows out through that layer of insulation. In a colder room, it takes somewhat more power to maintain the interior of the oven at a given temperature.

The oven is at 1000°C, so this difference in external temperature should have only a small effect. We will ultimately be primarily concerned with *changes* in the oven length (see § 5.3), and these changes should be similar in the FFTB and at UCLA.

We have such curves for a variety of plasma densities, and Table 4.1 shows the slope of FWHM versus heater power over a ten-fold range in density. Knowing the oven length is certainly necessary to measure the gradient in our plasma.

Given the substantial wings on either side of the oven, if we seek to have a region at the nominal full vapor density, the shortest oven possible is about 10 cm, where the full vapor density is achieved only at the exact center of the oven. By adding more power, we can lengthen the oven, making the peak density region grow, while retaining wings of the same size.

Oven Density (cm^{-3})	**Slope (mm/W)**
0.3×10^{17}	0.68 ± 0.05
0.6×10^{17}	0.70 ± 0.04
1.5×10^{17}	0.64 ± 0.04
2.0×10^{17}	0.58 ± 0.03
2.5×10^{17}	0.56 ± 0.03
3.5×10^{17}	0.58 ± 0.04

Table 4.1: Oven length versus heater power slope for different oven densities. Each value for the slope of length versus heater power is calculated from several measurements of the FWHM at different heater powers. Errors are estimated from the spread in data points. The presence of the thermocouple may also add a systematic error.

Unfortunately, the heater used in E164 is damaged if too much power is put into it. A practical limit was about 500 Watts. At that power level, the oven cannot be longer than about 15 cm, so the range of oven lengths available was limited. Nonetheless, by changing the oven length, the region of full density can be changed in length from about 0 to about 4 or 5 cm long, allowing measurements of the change in acceleration with different oven lengths to be performed, see 5.3.

The other half of the calculation of gradient requires knowledge of the change in particle energies, and that is measured on the Cherenkov based energy spectrometer.

4.1.4 Cherenkov Spectrometer

The electron optics between the plasma cell and the spectrometer involve three quadrupoles and four long permanent magnet dipoles. We can image the exit of the plasma to our spectrometer location and also have substantial vertical dispersion.

At the location some 25 meters downstream from the plasma, with the needed small β_y and large η_y, we send the beam through a piece of aerogel 1 mm thick. When the electrons go through this material, they radiate many optical photons from the Cherenkov effect, as their speed exceeds that of light in the aerogel. The light from each electron comes off in a forward cone with an opening angle of several degrees, and we place a pick-off mirror to the side of the electron beam path to capture a portion of the light from each electron's cone, while allowing the beam to pass by.

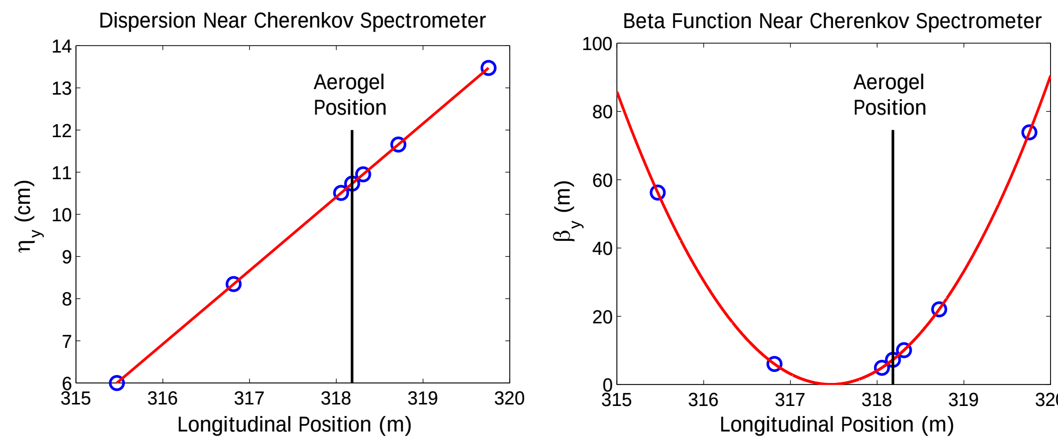

Figure 4.6: Output from ELEGANT showing the vertical dispersion and beta function for the electron beam in the region of our Cherenkov radiation based spectrometer.

This technique of creating Cherenkov radiation to observe the electron positions was originally used in the first plasma wakefield experiments involving long bunches with σ_z of about 650 μm. For such long bunches, a streak camera can be used to measure the energy spectrum of different longitudinal slices of the beam, allowing a direct measure of the acceleration in the tail. Clearly, this requires a prompt method of radiation production, so the Cherenkov screen was used. With the $\sim$40 fsec bunches in E164, we can no longer use the streak camera, so have no specific requirement that the radiation be prompt, but the system had worked well and did not need to be changed.

After picking off some of the Cherenkov light, we then use several mirrors to transport the signal to our 16 bit Princeton Instruments VersArray CCD camera located near the beamline. The camera lies in a region of substantial background radiation, so a lead hutch was constructed to shield the delicate CCD. A Nikon AF Micro-Nikkor 200 mm f/4 lens imaged the Cherenkov light to the camera.

The dispersion at the aerogel location can be measured through surveying, and calculation of its value using beam optics simulations has also been performed in ELEGANT [66]. Both methods give that the dispersion is 10.5 cm with an estimated error of a few millimeters. Figure 4.6 shows the output from ELEGANT with the vertical dispersion and beta function plotted in the region of our spectrometer.

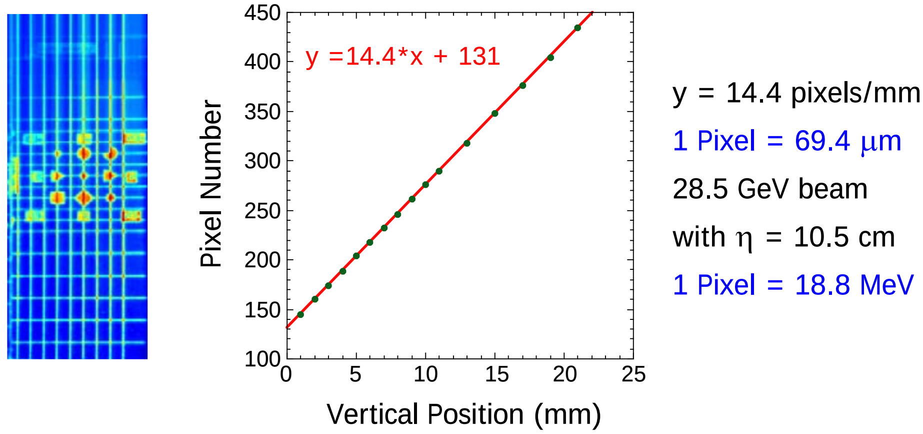

Figure 4.7: Sample image of the Cherenkov Aerogel with the calibration grid visible behind. A plot of the number of pixels per millimeter, combined with the dispersion, allows calibration of the energy per pixel.

With a number of quadrupoles between the plasma and the Cherenkov screen, there is a potential to introduce dispersion from misalignments in the quadrupole positions. To obviate this problem, all quadrupoles were aligned at the beginning of the experimental run to have no effect on the beam orbit with changing magnet strength, so do not contribute to the dispersion.

To understand the energy resolution of our spectrometer, we must know the dispersion and the vertical span viewed by each pixel of our camera. If Δd is the pixel width, or equivalently, the pixel separation, then the corresponding energy range per pixel is given by:

$$\Delta E = \frac{\Delta d}{\eta} E_0 \tag{4.4}$$

To measure the pixel's viewing area, a thin piece of paper is placed behind the aerogel with lines spaced 1 mm apart. We then image the grid of lines to our CCD to determine how many pixels span each millimeter. Figure 4.7 shows that there are 14.4 pixels spanning each millimeter, so the pixels each view an area of 69.4 μm. With a dispersion of 10.5 cm at the aerogel, the energy dispersion is therefore 18.8 MeV/pixel for a beam with central energy of 28.5 GeV.

As this spectrometer lies far from the exit of the plasma, we must verify that it is imaging in both x and y. We verify directly that we are imaging in the horizontal plane by changing the location of beam focus in the vicinity of the plasma and observing that the minimum horizontal spot size at the aerogel occurs when the upstream waist is placed at the location of the plasma exit. The horizontal spot size that we see on the Cherenkov screen is typically of order 400 μm and the incoming beam optics are set up to create a 15 μm spot at the location of the plasma, so the magnification ratio is about 25 times.

With y, we have no way to verify the betatron imaging directly, but because we are able to observe relatively sharp features of the energy spectrum, we estimate that any blurring from the finite β_y amounts to less than one or two pixels on the camera.

This Cherenkov based spectrometer is the central diagnostic for the entire E164 experiment and shows the energy changes resulting from the plasma wake. As there is minimal dispersion in x, and we use optics to focus the beam in that plane, the Cherenkov screen can also be used to investigate deflections in the horizontal plane that may happen to the electron beam in its passage through the plasma.

4.2 Coherent Transition Radiation Measurement

A central concern of the E164 program is to know the length of the electron bunches coming from the linac. The use of Coherent Transition Radiation (CTR) can give information on the bunch length directly, and several different measurements based on CTR are of interest.

The setup for the CTR measurement is shown in the schematic of Figure 4.8. We send the electron beam through a thin titanium foil some 20 meters before it enters the plasma cell. The CTR radiating from the foil is collected by an off-axis paraboloid to be sent through our interferometer. As we are dealing with wavelengths of light in the roughly 5 to 100 micron regime, we use a vacuum window made from high-density polyethene (HDPE) to provide reasonable transmission. This window also has the advantage of being opaque to the incoherent visible light portion, which could confuse the signal. Greater detail on the interferometer and the technique can be found in [67], and we summarize the main results useful for understanding the bunch structure.

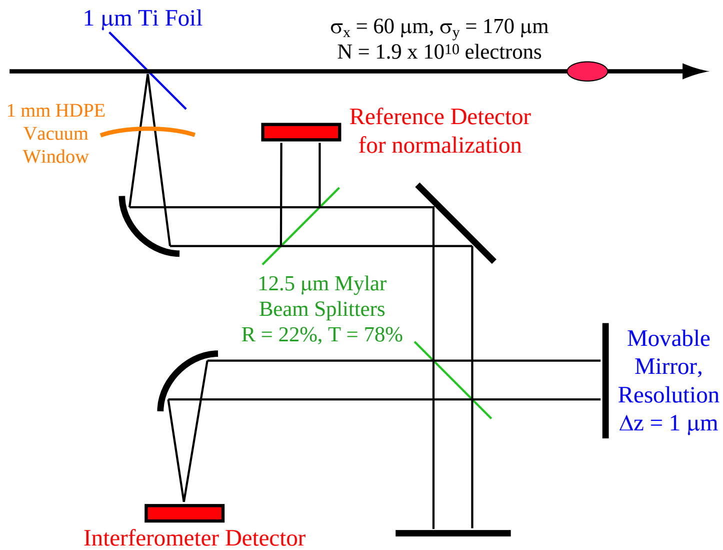

Figure 4.8: Scanning interferometer for determining the electron bunch length.

The beam as a whole emits transition radiation as its electric field encounters the discontinuity of the conducting surface. With a Gaussian bunch, for all wavelengths where $\lambda \gtrsim 0.6\,\sigma_z$, the radiation will be coherent [68]. For E164, this broadband CTR from the foil is emitted at a range of wavelengths. As we expect to produce bunches with a longitudinal RMS as short as 12 microns, the shortest wavelengths that can be produced coherently by this shortest bunch will be $\lambda_{min} \approx 7\,\mu$m. This corresponds to frequencies of over 40 THz.

One way to use CTR to understand bunch length would be to measure the overall spectrum and inverse Fourier transform the result. This requires precise detectors that work over a very broad range of wavelengths. The method chosen for E164 was to split the radiation and send it through an interferometer to obtain the direct autocorrelation of the bunch's electric field.

In both cases, extracting the exact bunch length requires an assumption about the functional form of the bunch's profile, such as Gaussian or hyperbolic secant (as with many fast lasers). Nonetheless, choosing different functional forms has only minor effect on the calculated width. For example, the RMS width of the autocorrelation of a Gaussian pulse is $\sqrt{2}$ longer than the original pulse, whereas the hyperbolic secant pulse has autocorrelator RMS width about 1.5 times longer than that of the original pulse. In general, we can estimate the actual pulse length to be shorter by roughly 1.4 to 1.5 times compared with the measured autocorrelation trace's RMS.

As with OTR, we are interested in the radiation pattern that we expect to observe, because knowing the spectrum helps to interpret results, see the end of this section. Theory again predicts a long formation length. Assuming that we are in the far field, the total power emitted as CTR per unit solid angle and per frequency is given by [69]:

$$\frac{dW}{d\Omega d\nu} = \frac{N_b^2 e^2}{4\pi^3 \epsilon_0 c} \frac{\beta^2 \sin^2\theta}{(1-\beta^2\cos^2\theta)^2} \exp\left[-\frac{4\pi^2\nu^2}{c^2\beta^2}\left(\sigma_z^2 + \beta^2\sigma_r^2 \sin^2\theta\right)\right] \tag{4.5}$$

where N_b is the number of particles, β is the particle speed normalized to c, and θ is the angle at which we look at the radiation.

We note that this function must be convolved with a function that takes into account the diffraction of the CTR from the finite source size in order to give the actual distribution of radiation in the far field [69]. We will integrate this numerically in Chapter 5, when we compare the total power observed against theoretical expectations. Clearly, shorter bunches will produce more total CTR power because the first term in the exponential becomes less negative with shrinking σ_z.

The above equation, and following estimates, all calculate based on assuming far-field optics. The formation length for CTR is longer than the size of our experimental apparatus, as it is for OTR, so these calculations must be considered approximate.

At the location where we measure the CTR, the electron beam is only about 60 by 170 microns in transverse RMS size. Thus, for even the highest energy photons we expect to produce, the source plane is not very many wavelengths across.

If we use estimates based on far field optics, we see that the radiation will diffract away from its point of emission rapidly, as the divergence angle of a beam of radiation is of order $\lambda/d \approx \lambda/2\sigma_r$ [70]. Thus, in the vertical direction, the $7\,\mu$m shortest wavelength produced has a characteristic full angle of about 60 mrad $\approx 3.3°$; all longer wavelengths diffract away more rapidly. In the horizontal direction, diffraction is less severe, with about 20 mrad divergence for the shortest wavelength.

An off-axis paraboloid mirror intercepts the broad cone of radiation to collimate it so that it can be sent through the arms of a Michelson interferometer. The paraboloid

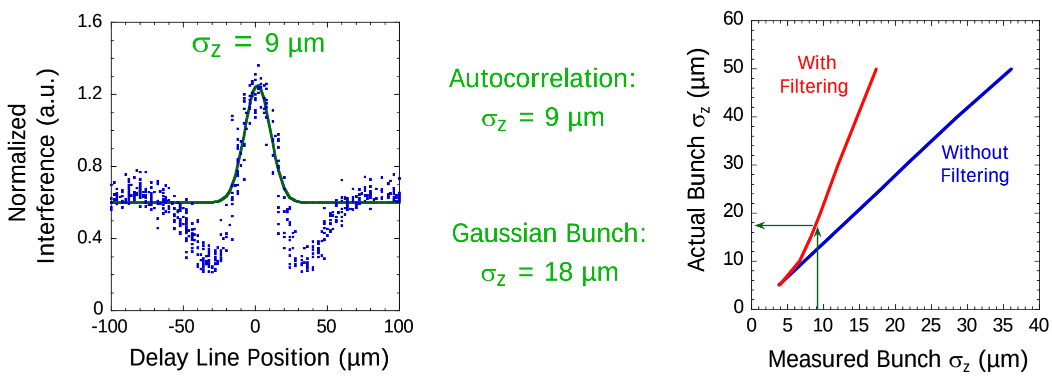

Figure 4.9: Autocorrelation Results. Courtesy Mark Hogan.

is a standard 2 inch optic held at 45°, so the effective width is about 1.4 inches. This optic is situated about 6 inches away from the CTR foil, so any radiation coming from the foil with a characteristic angle of divergence greater than about a third of a radian in the vertical plane or a quarter of a radian in the horizontal will not be collected fully. Given our source size, wavelengths beyond 40 μm are not fully captured vertically, and wavelengths beyond 90 μm are not fully captured horizontally. Again, as we are in the near field, these estimates should be considered approximate.

The detectors that we use for the radiation are pyroelectric elements (pyros) which absorb across a wide spectrum and produce a voltage in proportion to changes in their temperature. The detector elements are quite small, so have a bandwidth well above the 1 Hz repetition rate of our electron beam and measure each shot independently.

For autocorrelation traces, the reference detector is used to select only shots that have total CTR power within some narrow range. We know that the length of the bunches coming from the linac is not stable shot to shot, so it is convenient to select shots having total CTR power that is at the most probable value. We will see in Chapter 5 that that corresponds to bunches of order 20 μm long.

Scanning the interferometer arm distance over many shots, we can build up the bunch's autocorrelation function, as shown in Figure 4.9.

There are several effects which make interpreting the data not straightforward. Any head to tail asymmetries of the electron bunch are washed out in the autocorrelation, which is inherently an even function. Also, the broadband Terahertz radiation being emitted by our source gets absorbed at various wavelengths by the high-density polyethene (HDPE) vacuum window and the air. Moreover, we have seen that longer wavelengths of the radiation are not fully collected. Lastly, there are etalon effects in the beamsplitters. These elements have multiple resonances right in the brightest portions of our emitted radiation. This causes the bunch to appear artificially short and also causes the two dips visible on either side of the peak.

Taking these effects into account, we have calculated what we expect to see in the data for various bunch lengths. With the anticipated correction ratio as indicated in the right graph of Figure 4.9, we expect that the observed data actually correspond to a bunch length of $\leq 18\,\mu$m in the measurement run presented. We will see in Chapter 5, that this is close to the answer provided by simulations. The correlation of these answers provides confidence that each is correct.

4.3 Non-Invasive Energy Spectrum Measurement

Many diagnostics which seek to measure the energy spectrum of an electron beam are destructive in that they present a physical barrier to the beam. This increases the transverse emittances and robs the beam of energy, also causing an increase in energy spread.

We are careful to create tightly focused beams which are short in time, so a physical barrier will clearly not be advantageous. Recalling a technique used in the SLC to measure the energy spread [71], we decided to use a large magnet as the basis for our spectrometer.

By deflecting the electron beam slightly in the vertical direction, each electron makes a vertical stripe of synchrotron radiation. As the electron beam is already dispersed according to energy in the horizontal plane, the sum of all such stripes is imprinted with the energy spectrum of the electrons and can be thought of as a sort of analog “bar code.”

We intercept these X-ray photon stripes on a scintillator to convert the signal to visible light. A scientific-grade CCD camera provides a digitized image of our bar code. Summing this image in the vertical direction, we get a curve which reflects

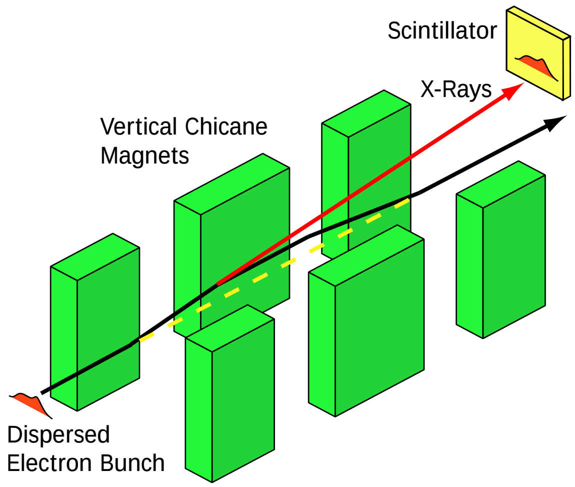

Figure 4.10: Schematic of the synchrotron radiation producing chicane.

the energy spectrum of our electrons. A schematic of the process of translating the electron positions to a visible signal is shown in Figure 4.10.

We need to separate the electron beam and the X-rays, so the magnet was configured as a chicane. The electron orbit returns to what it had been before the magnet, while the synchrotron radiation comes off in an upward fan such that it can be intercepted by a scintillator which is held just above the electron beam path.

4.3.1 Potentially Destructive Effects on the Beam

Although we do not present a physical barrier to the beam, there are still possible deleterious effects from its passage through the magnetic fields, so we calculate them.

Coherent Synchrotron Radiation Longitudinal Effects

The bunch is already very short at the point of our magnetic chicane, so we must verify that the effects of Coherent Synchrotron Radiation (CSR) will not be of concern. This radiation is at much longer wavelengths than the incoherent synchrotron radiation we seek to use. We discuss the properties of the incoherent portion in § 4.3.3.

The wavelength cutoff for CSR is of the order of the bunch length: $\lambda \gtrsim 0.6\,\sigma_z$ (the same as for Coherent Transition Radiation, c.f. Chapter 5). CSR normally affects the energy spread of an electron beam as well as its transverse emittances.

The Linac Coherent Light Source (LCLS) program at SLAC is very concerned with such effects in their proposed bunch compressor chicanes, and several useful formulae for estimating the energy spread due to CSR can be found in the LCLS Design Report [72]. In particular, we seek to know the energy spread induced by CSR, and it is given by:

$$\sigma_\delta \approx 0.22\,\frac{N r_e L_B}{\gamma \rho^{2/3} \sigma_z^{4/3}} \tag{4.6}$$

where N is the number of particles in our bunch, r_e is the classical electron radius, L_B is the length of the bends in our magnet, γ is the relativistic factor, ρ is the bend radius and σ_z is the bunch length.

Our magnet (see next section and Figure 4.12) has four bends of ≤ 50 cm length each. The peak magnetic field available is 1.3 T, with corresponding radius of curvature of roughly 75 m. An upper bound for the CSR effect is given by assuming that the entire bend length is at this full field. We obtain from (4.6) that σ_δ is no greater than 6.5×10^{-5} for the effect from all four magnets.

To find the final energy spectrum of the beam, we simply convolve a Gaussian of width σ_δ with the incoming spectrum, whose energy spread is already 1.1% in RMS. We can immediately see that this convolution will make no discernible change, because the added energy spread changes are negligible compared to the 1.1% RMS.

Coherent Synchrotron Radiation Transverse Effects

For changes in the emittance, we use a complicated approximate formula for the increase due to CSR found on page 7-38 of the LCLS design report:

$$\frac{\epsilon}{\epsilon_0} \approx \sqrt{1 + \left(\frac{0.11}{3}\right)^2 \frac{r_e^2 N^2}{\gamma \epsilon_N \beta} \left(\frac{L_B^6}{\rho^5 \sigma_z^4}\right)^{\frac{2}{3}} \left(L_B^2(1+\alpha^2) + 9\beta^2 + 6\alpha\beta L_B\right)} \tag{4.7}$$

where ϵ_N is the normalized emittance before the bend, and α and β are the Twiss parameters for the beam at the location of the bend.

Given these values at the magnetic chicane, the increases in emittance for x and y are about 0.02 and 0.2 percent, respectively. This will have no effect on our beam.

Wake Fields from the Scintillator Crystal

In passing through an aperture, part of the energy stored in the beam's electric field is scraped off, causing the particle to lose energy. Our crystal sits on one side of the beam path about 2 mm away, but a useful limit on the wakefields comes from the more easily calculated geometry of a circular aperture of radius 2 mm; we follow the treatment in [73]. Any wake loss from the crystal will be something less than half the value calculated. For a bunch of charge q and length ℓ, where $\gamma\ell$ is much greater than the radius of the aperture, r, we find that the energy lost by each particle is given by:

$$W \approx \frac{N_{.,} e^2}{2\pi\epsilon_0 \ell} \ln\left(\frac{\gamma\ell}{2r}\right) \tag{4.8}$$

For our bunches with 1.8×10^{10} particles, a FWHM length of as little as $25\,\mu$m, and a distance to the crystal of 2 mm, we find that the maximum energy lost per particle is about 12 MeV. This is an overestimate not only because the crystal covers less area than an equivalent radius circular aperture, but because the fields are not completely reflected by the crystal, which is very much not a perfect conductor. Based on geometric arguments, every particle loses a bit of energy, with the largest expected energy loss something less than 5 or 6 MeV. The whole beam simply loses a tiny amount of energy, and there is no discernible effect on its structure.

Bunch Compression in the Chicane

Lastly, we can say that there is no measurable bunch compression occurring in our chicane. We know that the bend angles are only 2.5 mrad, and for a chicane with such small angles, we have an approximate formula for the R_{56}, where ΔL is the distance between the first and second bends, which must be the same as between the third and fourth bends [74]:

$$R_{56} \approx -2\frac{L_B^2}{\rho^2}\left(\Delta L + \frac{2}{3}L_B\right) \tag{4.9}$$

Our chicane has $R_{56} \approx -17\,\mu$m and even extreme opposite energy particles will move by less than half a micron in z relative to one another. This tiny change compared to the bunch length is dwarfed by the effects of the FFTB dogleg's R_{56}.

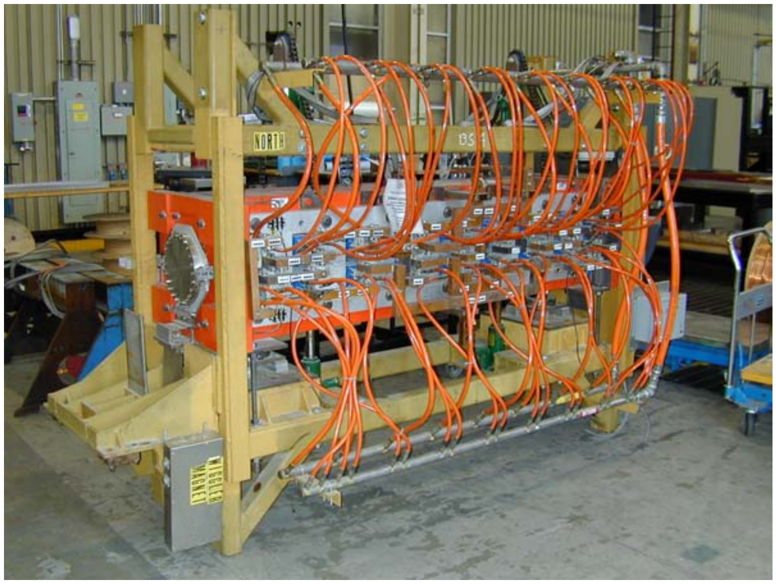

Figure 4.11: Photograph of our magnet rewired as a chicane. The orange hoses are for water cooling. On the front face is one of the two added flux return plates.

Overall, these effects of the chicane on the electron beam are essentially zero, and we treat the beam as returning to its nominal orbit unchanged. Thanks to the bends in the chicane, however, synchrotron photons come off at a variety of small upward angles to be intercepted by a phosphor about 2 m downstream of the magnet.

4.3.2 Magnet Hardware and Details

Due to space constraints, we needed to deflect the electron beam by several milliradians in each bend of the chicane in order to have the X-Rays be at least a few mm above the electron beam by the time they were intercepted by our phosphor about 2 meters downstream.

To bend the 28.5 GeV beam sufficiently, we needed a magnet with high fields sustained over distances of more than a few centimeters. To bend our beam by 2.5 mrad requires that $\int \vec{B} \cdot dl \approx 0.25$ T·m.

The available hardware was a large electromagnet 1.8 meters long and weighing some 3 tons, see Figure 4.11. This had originally been designed ca. 1980 for use as a four period wiggler with the beams in SPEAR of up to approximately 3 GeV. The

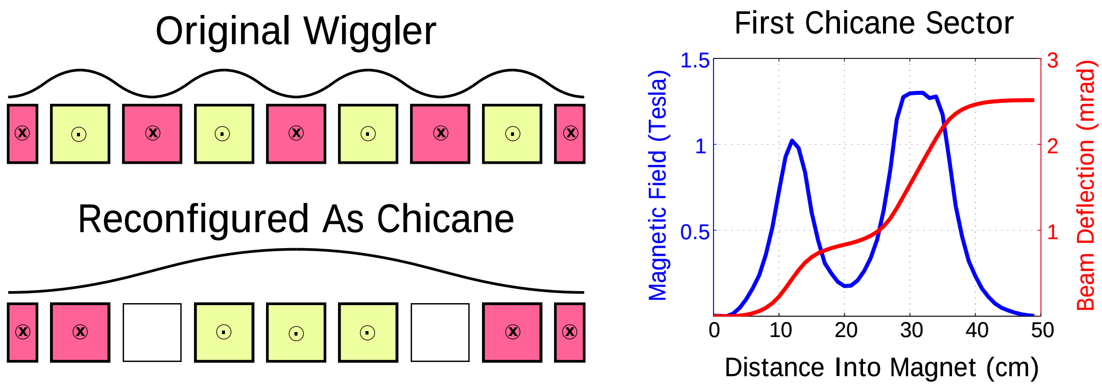

Figure 4.12: On left is a schematic of the rewired magnet. At right is a plot of the magnetic field strength versus longitudinal position going through the first bend.

magnet has nine separate pole pieces, where the central seven are the same and at either end are two poles with half the integrated $\vec{B} \cdot dl$ of these central poles.

With over ten times the beam energy of SPEAR, the deflections from single poles in the wiggler would not have been large enough, so we rewired the magnet as a chicane by combining neighboring poles. The first 1 1/2 poles together constitute the first bend, the central 3 poles make bends two and three and the final 1 1/2 poles are the final bend. We can see a schematic of this change on the left in Figure 4.12.

The magnet was originally an air core magnet where the separation between the two halves could be adjusted to adjust the field. The magnetic fields available were substantially weaker because the magnet had to drive field across not only the air gap between the poles but also across the gap between the two halves of the yoke.

Requiring the strongest fields possible, we closed the yoke gap and held the magnet with two large steel plates. One of them is visible behind the forest of orange cooling hoses in Figure 4.11. These act as flux return for the magnet, increasing the field strength between the poles. With the plates, the magnet achieves about 1.3 Tesla at its maximum rated current of 800 Amps. The right side of Figure 4.12 is a plot of the magnetic field as we progress through the first bend. These measurements were taken with a Hall probe pulled through the magnet at 1 cm intervals. Included on the plot is the total cumulative bend angle for a beam at 28.5 GeV. We see that each of the four bends in our chicane will deflect the beam by 2.5 mrad, as required.

4.3.3 Synchrotron Radiation Properties

The electrons see a varying magnetic field in the course of traversing each bend, and therefore at each of these points, they have a different synchrotron radiation spectrum. We can calculate this instantaneous spectrum by following the treatment in [46]. A useful quantity to know is the total synchrotron power over all wavelengths, P_γ, which is given (in MKS) by:

$$P_\gamma = \frac{e^2 c}{6\pi\epsilon_0}\frac{\gamma^4}{\rho^2} \tag{4.10}$$

where, again, γ is the relativistic factor of the electron and ρ is the bend radius of the beam in the magnet.

To calculate the spectrum, rather than just the average power, we first need the critical frequency, which is the frequency above and below which half the radiated power lies. This ω_c is given by:

$$\omega_c = \frac{3\,c\,\gamma^3}{2\,\rho} \tag{4.11}$$

When we have calculated ω_c, it is natural to normalize all frequencies to it, using $x = \omega/\omega_c$. In these normalized units, all synchrotron radiation spectra look the same. To derive the actual spectrum, we just scale by ω_c and find the appropriate constants to give it unit area. Then we will just need to multiply by the total power, P_γ, to give the actual power at each frequency. Note that this is similar to what can be done with the blackbody radiation from an object at finite temperature, where ω_c is the characteristic parameter of the spectrum, as T is for thermal radiation.

It turns out that the spectrum's shape is given by an integral of a Modified Bessel Function of $5/3$ order. With P_γ and ω_c, we can determine the power spectrum for a single electron. The equation is:

$$P(\omega) = \frac{9\sqrt{3}}{8\pi}\frac{P_\gamma}{\omega_c}\,x\int_x^\infty K_{5/3}(x)dx \tag{4.12}$$

As our magnetic field varies, so will the spectrum's peak value and total power. To find the total spectrum in our bend magnet, we find the instantaneous spectrum associated with the magnetic field strength at each point through the magnet. We

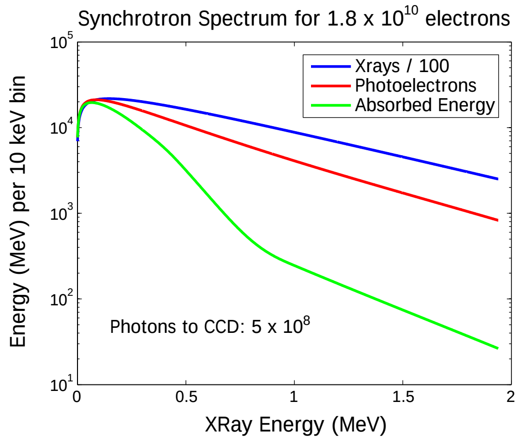

Figure 4.13: Synchrotron radiation spectrum for 28.5 GeV beam with 1.8×10^{10} electrons in the first chicane bend. Blue is the X-ray spectrum divided by 100 for plotting. Red shows the photoelectron spectrum from the (small) fraction of X-rays intercepted by the crystal. Green shows the energy ultimately deposited in the crystal.

use the data as plotted on the right hand side of Figure 4.12 as the basis for the calculation. Adding up the effects from each part of the bend, we get a numerical approximation to the integrated final spectrum of radiation that we expect for each electron and multiply it by the 1.8×10^{10} particles in each bunch. Although the calculation refers to ω, we simply multiply by $\hbar$ to give the more conventional units of photon energy on the horizontal axis.

The first bend's total spectrum is shown in the blue curve of Figure 4.13. There is substantial power in the spectrum, and we divide by 100 for plotting purposes so the curve will lie close to the other curves discussed in the following section.

Note that the total energy lost by each electron in going through the chicane is about 800 keV, composed of photons primarily in the range around 100 keV. The loss of 800 keV is insignificant compared to the beam's average incoming energy. Changes to the energy *spread* are also small, as the average electron emits something like eight photons. Therefore, variation in energy loss between electrons is expected to be of order $\sqrt{8} \cdot 100\,\mathrm{keV} \approx 300\,\mathrm{keV}$, some 10 ppm of the beam energy. The incoherent synchrotron radiation has even less effect on the beam's final spectrum than the insignificant effect of the coherent radiation.

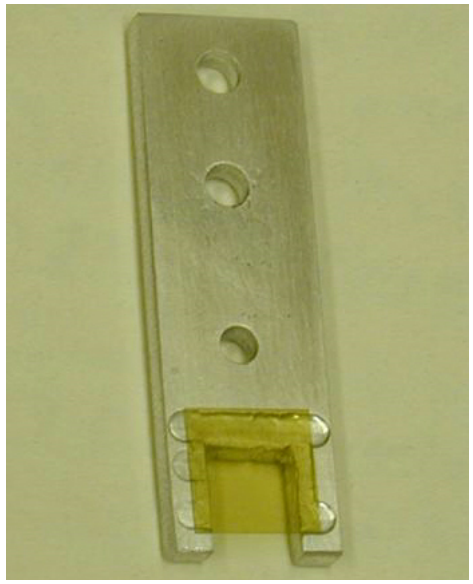

Figure 4.14: Scintillation crystal in its Aluminum holder.

4.3.4 The Scintillator Crystal

To detect the X-rays, we use a $10 \times 10 \times 0.08$ mm crystal made of Cerium doped Yttrium Aluminum Garnet scintillator (YAG:Ce). Figure 4.14 shows our yellowish crystal in its mount which holds it several millimeters above the electron beam path.

We chose this particular crystal because it is mechanically very strong, and has good light output centered in the green portion of the spectrum, where our camera is most sensitive. Due to its strength, we could use a thin crystal without fear of cracking it in the process of mounting to the aluminum holder. The thin crystal reduces blurring from multiple scattering of photoelectrons as they transit the crystal. We discuss this effect in § 4.3.7.

The thickness of crystal presented to the X-ray beam is greater by a factor of $\sqrt{2}$ because the crystal is at a 45° angle to the photon beam. This allows our camera to view it head-on with all parts in focus. (The camera is very close to the crystal with a zoom lens such that the depth of focus is only about one hundred microns.)

We address the issue of spatial resolution below, but first we must know how much energy actually gets deposited in the crystal from our synchrotron X-rays. We needed a crystal thick enough to give an appreciable signal.

X-Ray Absorption

The Atomic Nucleus by Evans [75] gives a variety of useful curves and formulae for the interaction of photons and energetic particles with matter.

We use these curves first to find the expected absorption of X-rays by our thin crystal. Specifically, a graph from p. 686 of Evans for Aluminum is most useful. Because of its greater density, YAG:Ce will stop X-rays 1.74 times as much as Aluminum, so we scale the absorption values accordingly.

The red curve in Figure 4.13 shows the significantly smaller total energy that is actually absorbed by electrons in the crystal, primarily through the photoelectric effect and some Compton scattering. With an effective thickness of $80 \times \sqrt{2} = 113\,\mu$m of material, we expect the absorption to vary from 1.1% at 10 keV to about 0.3% at 2 MeV. At 1 MeV, a photon coming from the magnet has only a 0.45% chance of being absorbed. We note that essentially all of the energy of the photon is transferred to the liberated electron.

Electron Range and Energy Deposition in YAG:Ce

The scattered electrons stop rapidly at low energies, but the faster electrons will fly out the back of the crystal, depositing only a fraction of their energy into it before escaping. On p. 625 of Evans is a useful empirical formula for the range of energetic electrons through matter when the electrons have energies between 10 keV and 3 MeV, which covers our spectrum well. When E is in MeV, and ρ is in g/cc, we find that the range, given in microns, of electrons in material is:

$$R = \frac{4120\,E^n}{\rho} \qquad n = 1.265 - 0.0954 \ln E \tag{4.13}$$

For YAG:Ce, with a density of 4.57 g/cc, we find that 10 keV electrons go less than a micron. At 100 keV, electrons have a range of about $30\,\mu$m and 1 MeV electrons can go about $900\,\mu$m. See the left side of Figure 4.15.

With both the photoelectric effect and, for higher energy X-Rays, Compton scattering, the electrons are not ejected parallel to the direction of the incident photon. There is a spread in angles of emission, but for the lowest energy photons which can

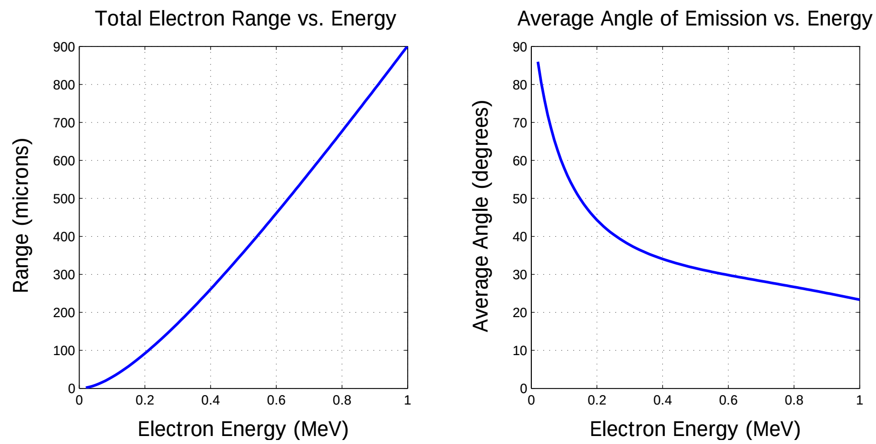

Figure 4.15: Total range and average angle of emission for electrons of varying energies transiting the YAG:Ce crystal.

still liberate electrons from the ground state, the angle of emission tends to be in the direction of the electric field, namely at 90° relative to the photon incident direction.

As the photon energies become higher, the momentum of the photon causes the electrons to be ejected more and more in the direction of the incident X-Ray. At about 0.2 MeV, electrons are ejected, on average, at 45° and by 1 MeV, the average angle is just under 25°, as shown on the right hand plot of Figure 4.15.

The electron loses energy through many small collisions, so its trajectory tends to continue along a fairly straight line. Given this, we can calculate the longitudinal and transverse components of the range for the average emitted electron. We do so in Figure 4.16, and note that these curves will be relevant later for calculations of resolution. For the moment, though, we concern ourselves with the longitudinal range, as it allows estimates of how much light the scintillator will produce.

The electrons liberated by the X-Ray beam are born evenly throughout the thickness of the crystal, so the average of how much energy is deposited by them must take that into account; with a tiny range, any electron born at 10 keV anywhere but the exact back surface of the crystal will deposit all of its energy into the crystal.

After electron energies rise much above about 0.25 MeV, their longitudinal range becomes roughly the same as the crystal thickness. If the electron is born at the exact front of the crystal, it will deposit all of its energy, but if it is born at the

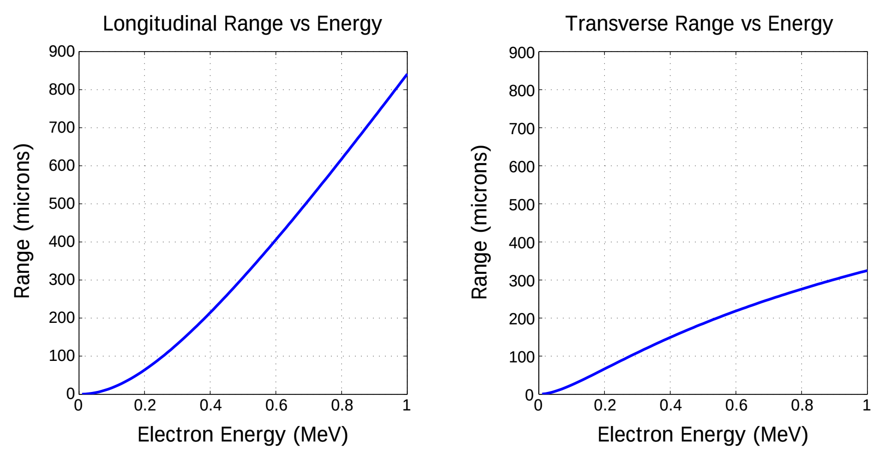

Figure 4.16: Average longitudinal and transverse ranges for electrons of varying energies transiting the YAG:Ce crystal.

back surface, it will deposit none. Averaging over the possible places where the particle could be born, we see that, statistically speaking, 0.25 MeV photons which get absorbed ultimately deposit about half of their energy in the crystal.

At 1 MeV, the longitudinal range is about 850 μm, much longer than the crystal thickness. According to [76], the loss of energy by an electron in matter per unit distance is nearly constant with electron energy, varying only logarithmically. We can thus say that the fraction of the electron's energy which is deposited in our crystal for particles born at the front surface should be approximately $113/850 = 13\%$. For particles at the back surface, again no energy is deposited. Averaging, we see that 1 MeV electrons deposit about 6.5% of their energy before exiting the crystal.

Combining the low chance of X-ray absorption and the average energy loss of the resulting electrons, we can calculate the average energy deposited by each photon. For example, the average 1 MeV photon, which gives birth to a ~ 1 MeV electron, ultimately deposits just $0.45\% \times 6.5\% = 2.9 \times 10^{-4}$ of its energy in the crystal.

The green curve in Figure 4.13 reflects the overall energy that is actually deposited in the crystal as a function of the incoming photon energy. Above about 200 keV, where the electron longitudinal ranges start to become significant relative to the crystal thickness, an ever larger fraction of the electrons exit the back of the crystal, causing a rapid drop in deposited power. Once the electron range is significantly larger than the crystal thickness (above 0.6 to 0.8 MeV), the nearly constant dE/dx of high energy particles in the crystal means that we enter the more shallowly sloped region where the fraction of energy deposited goes roughly as 1/E.

Photon Production in YAG:Ce and Capture in Our Camera

Knowing how much energy is actually deposited in the crystal, we can calculate how much optical light will come out for us to observe. The manufacturer of our YAG:Ce scintillator claims that per MeV of deposited energy in the crystal, about 8,000 photons with a spectrum centered at a wavelength of 550 nm will be created. Knowing the solid angle subtended by our camera lens, we can then calculate how much light the camera will capture from the isotropic emission. In this case, we expect about half a billion photons to reach the CCD. As discussed in § 4.3.6, this is a good signal. A major concern had been whether this thin crystal would have enough signal for the camera, and it worked well when installed.

Reflections back and forth between the two polished faces of our crystal are a possibility which could reduce the resolution. Additionally, if photoelectrons are continually ejected from the crystal, it will be charged to a very high voltage, potentially leading to damage. To ameliorate both effects, we purchased the crystal with a protective coating. The crystal was coated on both sides with a 75 nm Indium Tin Oxide (ITO) transparent conductive layer. The ability to conduct allows the crystal to avoid possible damage from electrostatic fields. The specific ITO thickness was chosen to make a quarter wavelength antireflection coating for the scintillation central wavelength of 550 nm, where ITO has an index of refraction of about 1.6.

Light leaking out from the edges of the crystal is visible, but in software we choose a portion of our camera images which excludes all crystal edges to remove this as a source of possible error in our measurements.

Having the appropriate crystal, we need a way to hold it above the electron beam and actuate it up and down so that it is close to, but not intersecting the beam path. Given our maximum 2.5 mrad divergence between the X-Rays and the electrons, and about 2 meters of propagation distance, the x-rays which are *furthest* from the electron beam at the crystal are only 5 mm above, so adjustability at the sub-millimeter scale

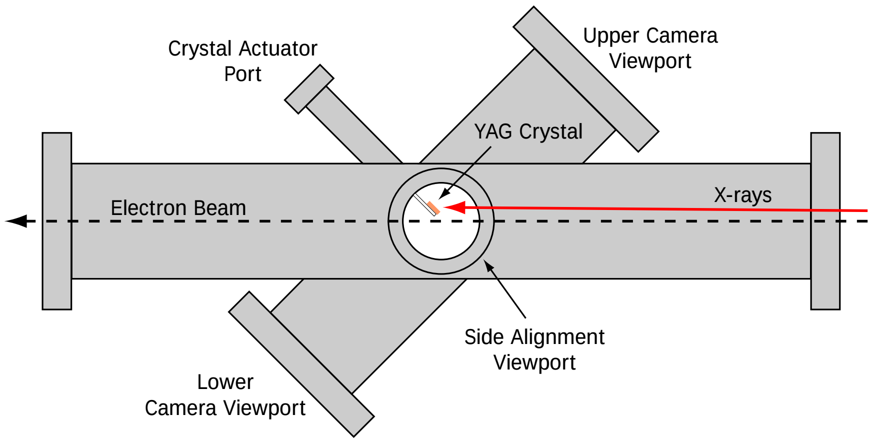

Figure 4.17: Schematic of the stainless steel vacuum chamber for the scintillator crystal. It provides ports for alignment and viewing by two cameras simultaneously.

is needed. The actual mount for the crystal is pictured in Figure 4.14. To move the assembly up and down above the beam, we use a stepper motor driving a linear actuator that feeds through into the vacuum chamber.

4.3.5 The Vacuum Chamber

In order to hold the YAG:Ce crystal above the electron beam at a variable height and also to be able to view the crystal from various angles, we constructed a special purpose vacuum chamber with a variety of ports for alignment and viewing of the crystal. Figure 4.17 shows a schematic of the arrangement used.

Although only one camera is necessary to view the crystal and provide the high quality images useful for later data analysis, it was useful to add the ability to have a second analog camera viewing the back side of the YAG:Ce crystal. The crystal mount was purposely made in the shape of an upside down "U" (see Figure 4.14). Thus the crystal is visible from both sides, making this possible.

The operators of the main accelerator sometimes used the spectrum information coming from this second camera to tune the beam. As the spectra for different running conditions are distinct, this was a useful diagnostic for them, and could be used at all times, unlike the scientific CCD camera for E164.

4.3.6 The Camera

For imaging the light coming from our scintillator crystal, we used a SenSys camera made by Photometrics. This line of air-cooled cameras uses scientific grade CCD imagers which are sensitive to relatively low light conditions.

We imaged onto the CCD with a standard 105 mm AF Micro-Nikkor lens made by Nikon. This has up to an f/2.8 aperture for good light gathering. Due to space constraints, the lens was situated about 20 cm from the YAG:Ce crystal, giving a demagnification of 3 times and a depth of focus of about 100 μm as mentioned above.

Our model, the KAF 1400, has a CCD with 1317 $\times$ 1035 pixels, each 6.8 $\times$ 6.8 microns in size. The quantum efficiency in the region of 550 nm is 42%, and the ADC saturates at between 10,000 and 80,000 electrons, depending on the sensitivity setting. The corresponding noise levels per pixel range between 12 and 20 electrons. We used the High Dynamic Range gain setting with full well capacity of 40,000 electrons and expected noise of 17 electrons. The digitization is with 12 bits, corresponding to about 10 electrons per count. Readout of the camera was performed at up to 1 Hz by a computer located in the FFTB tunnel.

The predicted half billion photons from our scintillator were imaged onto a region of about 300 $\times$ 200 pixels. This gave an expected value of some 3000 photons converted to detectable electrons in each pixel and typical ADC counts of 300. Actual images show a lower average count per pixel of more like 200, with uniform and easily subtracted background counts per pixel of about 60. The actual light seen was roughly half what had been expected. Given the various uncertainties in the calculation of the estimated light yield, this was still reasonably close to expectations, and in practice we had a good signal with no difficulties seeing the spectrum.

4.3.7 Spectrometer Resolution

Computer simulations of the beam properties in the FFTB indicated that the available locations for our spectrometer would provide good spectra. Figure 4.18 shows the Twiss parameters in the available region, and plots the actual location that was finally selected due to space constraints.

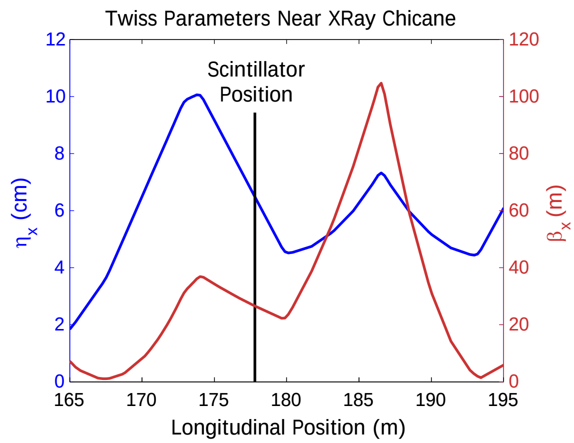

Figure 4.18: Plot of horizontal Twiss parameters at the start of the FFTB.

The fundamental ability of a spectrometer to provide good spectra at any location has to do with the balance between the dispersed and betatron spot sizes there. Our non-zero beta function acts to blur our ability to view the spectrum.

There are also a series of other resolution compromising effects that are unique to our indirect method of observing the electron positions. We discuss each of these sources of blurring in order, starting with the beam properties.

Electron Beam Focusing

The first resolution limiting effect comes about from the imperfect focusing of the electron beam. We recall the usual formula for the spot size in a region of dispersion:

$$\sigma = \sqrt{\beta\epsilon + (\eta\delta)^2} \tag{4.14}$$

where δ is the fractional energy spread as discussed in Chapter 3.

In the circumstance of an ideal cold beam with very small emittance ϵ or a tightly focused beam with small β, the (horizontal, say) position of a particle would be determined solely by its energy, and we map energy to position with perfect fidelity.

For real beams, the $\beta\epsilon$ term will always play a role, and its contribution is to cause uncertainty in particle positions because even particles of the exact same energy will be spread out over the betatron spot size. Mathematically, equation 4.14 can be thought of as the result of a convolution of the energy spectrum with a Gaussian of width $\sigma_\beta = \sqrt{\beta\epsilon}$. This convolution has a smoothing effect which reduces the information content of the spectrum.

Referring to Figure 4.18, we can see that the dispersion is about 6.4 cm. The RMS (correlated) energy spread is 1.1%, giving a spot size σ_δ from horizontal dispersion that has an RMS of about 1 mm. The spot size from the betatron focusing of the beam is not so much smaller than this value. With typical normalized emittance of about $\epsilon_x = 50\,\mu$m and $\beta = 26$ m, we find $\sigma_\beta = \sqrt{\beta\epsilon_x/\gamma} = 153\,\mu$m. This gives the uncertainty in position for an electron of a given energy, and turns out to be the dominant source of blurring in our spectrum.

Three quadrupoles in the vicinity of the X-Ray spectrometer can be used to alter the R_{56} of the FFTB for varying bunch compression at the plasma. For all data discussed in this thesis, these magnets were unchanged, leaving R_{56} at 1.5 mm, but when R_{56} is set to 1 or 2 mm, the dispersion changes to 5.6 cm and 7.2 cm, respectively.

Electron and Photon Trajectories

The spectrometer magnet is a large object almost 2 meters in length, and the vacuum chamber with our scintillator lies almost a meter downstream. This means that the position where the photons are radiated is in a region of significantly larger dispersion than the location where they are intercepted.

In the horizontal plane, each electron radiates an X-ray which propagates in a very narrow cone centered in the same direction as the electron is moving. This will lead to slight blurring, as discussed next, but even though the photons are born some meters upstream, they are still, on average, in the exact same horizontal place as their originating electrons when both reach the scintillator. This is why we must calculate the dispersion and beta function for the electrons at the location of the detector.

X-ray Production

When the electrons radiate X-rays in our chicane bends, the X-rays do not come off in a perfectly collinear fashion. From each electron, the radiation is expected to come off in a cone with a spread of angles in a way analogous to the case of Optical Transition Radiation.

The photon cone is centered on the direction of motion of the source charge, and the distribution of angles about that direction is governed by the frequency of the emitted radiation. The RMS spread in angles σ_θ is given by [77]:

$$\sigma_\theta \approx \frac{1}{\gamma}\left(\frac{\omega_c}{3\,\omega}\right)^{\frac{1}{2}} \tag{4.15}$$

where γ is the relativistic factor of the emitting electron while ω_c is the critical frequency for the synchrotron radiation and ω is the frequency for which we seek to know the opening angle of our cone.

Even for photons with only four percent the critical energy (of which we do not have many), the opening angle of the cone of radiation is only about three times as large as the value $1/\gamma$, and all photons with higher than the critical energy are even more collimated. As our relativistic factor is about 56,000, even the lowest energy photons come out in a cone narrower than 60 μrad and the vast majority of the power comes out within an angle more like 20 μrad.

With some 2 meters of propagation to our scintillator screen, the RMS spread of the X-rays is of order 40 μm and this blurring effect is much smaller than that from the imperfect focusing of the electron beam.

X-ray Absorption

We discussed the thickness of our scintillator crystal above, and one of the reasons we chose to use a thin crystal came from concerns about resolution. We refer again to Figure 4.16 in order to estimate the blurring arising from the fact that electron trajectories are not collinear with the incident X-Rays.

As noted in § 4.3.4, above about one quarter of an MeV, the average electron has a longitudinal range greater than the thickness of our crystal as presented to the photon beam. Understanding the resolution limits for electrons at this threshold in energy is useful, as it represents nearly the *maximum* possible blurring of our signal.

Having a longitudinal range slightly exceeding the 113 μm thickness of the crystal as presented to the photon beam, every electron above 270 keV of energy will eventually exit the crystal. If that electron is born at the front surface of the crystal, its average transverse deflection is about 100 μm before leaving the crystal. That electron can be born with essentially equal probability anywhere through the thickness of the crystal, and particles originating at the back surface of the crystal clearly experience no transverse deflection before exiting. The average transverse deflection within the crystal experienced by each 270 keV electron is thus 50 μm.

Electrons with less energy simply have a total transverse range which is smaller. Competing with that reduction in total range is the fact that a higher fraction of these lower energy electrons will traverse their entire range within the crystal, once their longitudinal range shrinks below the crystal's thickness. Thus, they will also go fully as far to the side as possible. It turns out that the increasing fraction of particles going to their full transverse range is overwhelmed by the rapid diminution of that range. For example, at 100 keV, most of the electrons have a longitudinal range short compared to the crystal's thickness and will be stopped in the crystal. However, those particles have a total transverse range of under 30 μm and will therefore cause less blurring than the average electron at 270 keV.

For electrons with greater than the threshold of 270 keV, we know that they are emitted with angles ever more closely parallel to the direction of the original incoming photon. Although their overall ranges are longer and longer, the transverse deflections achievable in our thin crystal will progressively be reduced because the typical angles shrink and the thickness of our crystal is fixed. For example, at 500 keV, the typical angle of emission is only 32° and an electron born at the front surface only goes 60 μm to the side before shooting out the back of the crystal. Again averaging over the thickness of the crystal, electrons at half an MeV only blur by 30 μm.

The maximum blurring from the electron trajectories is about 50 μm and the average effect will be smaller, perhaps half as much. This is a minor source of blurring.

This estimation was borne out in the course of E164. In 2003, we had used a thicker crystal of 200 μm. Changing in 2004 to the 80 μm crystal made no observable change in the resolution, implying that even the thicker crystal had blurring which was small compared to that from the imperfect electron beam focusing.

Blurring From the Camera

We took several images of the YAG:Ce scintillator and its holder to focus the images and check resolution. We can see the 1 by 1 cm YAG crystal in full frame views from the camera. This crystal spans about 880 pixels as viewed by the camera, so the distance between points on the crystal viewed by neighboring pixels is $11.4\,\mu$m.

When we look at sharp edges in images taken of the entire aluminum holder and crystal, the transition is clearly visible within two pixels, implying that blurring from the lens is at or below that scale, $\sim$25 microns. This is also minor compared to the major resolution-limiting effect in the electron beam itself.

Total Blurring Effects

Nominally, we can resolve light coming from our scintillator that originates from points only $11.4\,\mu$m apart, but the actual resolution is poorer due to blurring.

As mentioned above, we add the various blurring sources in quadrature. When we combine the four blurring effects, of size 153, 40, 25 and 25 microns, respectively, we come out with a total resolution wash-out with an RMS of $160\,\mu$m corresponding to 14 pixels on the camera.

Energy Per Pixel in the Camera

At the scintillator, models of the electron optics indicate that the dispersion for our 28.5 GeV beam is near 6.4 cm for the common case of $R_{56} = 1.5$mm, so we would expect that each pixel will span about 5 MeV.

The dispersion at the location of our X-Ray based spectrometer is not well measured, so we do not have a good estimate of the accuracy of the simulations. In practice, a good way to calibrate this resolution directly is to compare it with the results from the Cherenkov spectrometer downstream from our plasma where the optics are well understood and have been measured many times. We have discussed the Cherenkov spectrometer and its calibration of 18.8 MeV/pixel in § 4.1.4. Because the X-ray based spectrometer is non-invasive, we can compare the energy spectra at both screens for the same shot to determine the relative calibration.

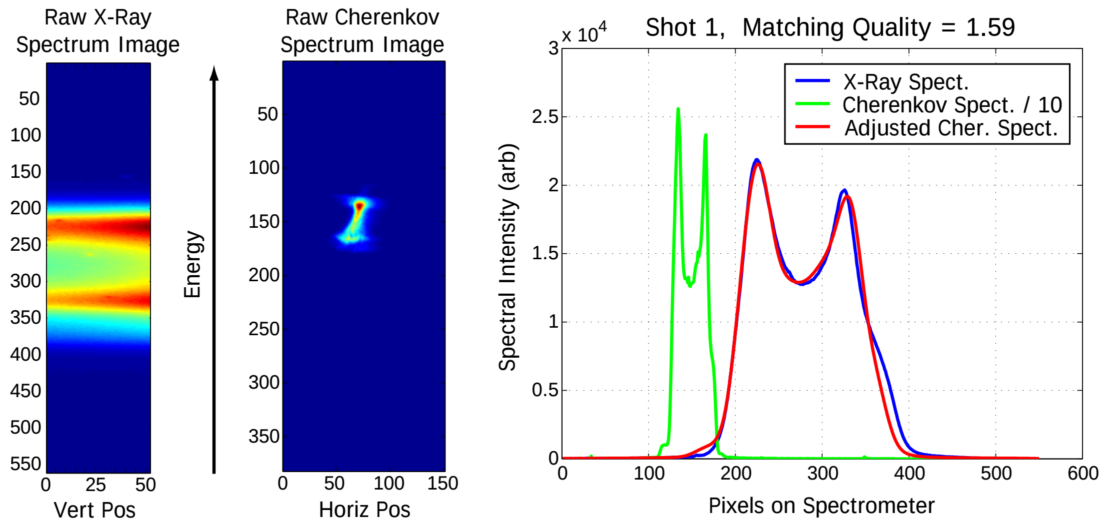

Figure 4.19: Comparison of X-Ray (left) and Cherenkov (center) based electron beam energy spectrum measurements. At right is a graph showing the two raw spectrum curves (blue and green) with the stretched and recentered Cherenkov spectrum (red) matching well. To match the X-Ray's lower resolution, the green curve is convolved with a Gaussian of 10 pixels width after it is stretched horizontally.

Although we take full images of the beam, we are interested in just the energy spectrum curve. This is made by summing the images sideways to find the number of particles at a given energy. We can see example images for the same data event as viewed by the X-Ray and Cherenkov cameras in Figure 4.19. In both images, higher energy particles are at the top. We typically plot the energy spectra horizontally for ease of viewing and note that the raw data from our cameras has the high energy particles at *low* numbers. Both cameras have uniform background light and electronic noise, so we subtract the background from each image by taking the average values at the beginning and end of our curves.

When we plot the energy spectrum curve for these two diagnostics together, the curves can have an arbitrary horizontal offset in the central energy and the overall amplitude of the two spectra will be unrelated. On the right of Figure 4.19, we plot the spectrum from the non-invasive X-Ray spectrometer in blue and that from the Cherenkov camera (divided by 10 in y for ease of plotting) in green. The red curve is the result of our fitting routine as described below.

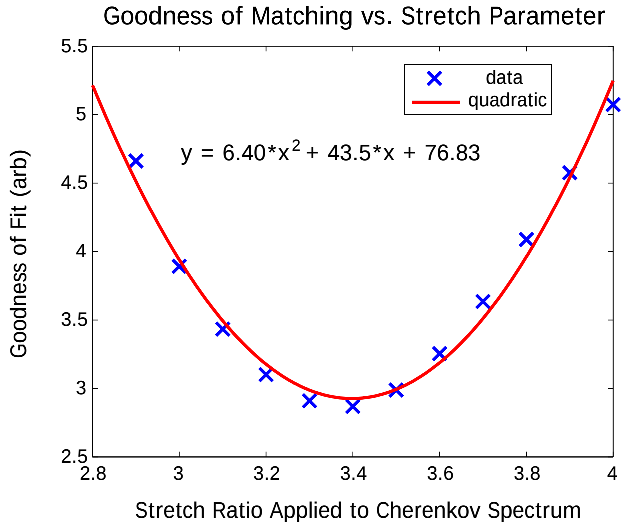

Figure 4.20: Matching Quality between the Cherenkov and X-Ray spectra as a function of the amount of stretching applied to the Cherenkov.

Clearly, the overall width of the Cherenkov curve is smaller than that of the X-Ray, so we are interested in the stretching ratio that will make them overlap most precisely, and try several values to see which gives best agreement.

To do this, we stretch the Cherenkov curve by a fixed ratio, which we scan between 2.8 and 4 times, by using a cubic spline interpolation and resampling. Then we blur the image by a Gaussian of a typical width of 10 pixels, to recreate the reduced resolution images from our X-Ray spectrometer.

Next, we use a fitting routine to align the two spectra horizontally and scale the Cherenkov image vertically to match the X-Ray. The parameter which is minimized by the fitting routine is the sum of the squared differences between the points on each curve, similar to a standard χ^2 metric.

For each of our stretching ratios, we go through a 100 shot run of data and align the Cherenkov and X-Ray images, calculating the "Matching Quality," where smaller numbers reflect better matching. We then average for all 100 shots and plot it versus the stretch ratio that had been used. We see the results in Figure 4.20.

The best overlap happens when the Cherenkov images are stretched by 3.4 times, meaning that the X-Ray spectrum has a dispersion of 5.5 MeV per pixel. This is not

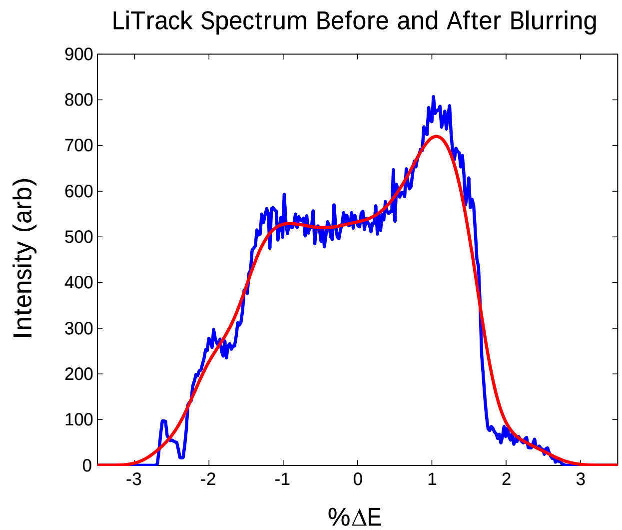

Figure 4.21: Example spectrum from LiTrack before (blue) and after mathematically blurring the image (red) to mimic the physical loss of resolution in the energy spectrometer. The blurring is performed by convolving a Gaussian of width 0.27% with the raw output from simulation, and several features are lost.

in complete agreement with the expectation from ELEGANT, and implies that the dispersion at our X-Ray about 10% smaller than expected. Because it is more closely based on a directly measurable quantity, we choose the value of 5.5 MeV/px, with an estimated error of 0.5 MeV/px.

Overall Energy Blurring Effects

Our estimated 14 pixels of uncertainty correspond to about 75 MeV, or 0.27%. This will cause washing out of some small features, but the major elements of our 4% full width energy spectra will still be visible. Better electron optics could reduce the dominant source of blurriness to the point where the other effects must be taken into account, but they are not of significant concern here.

Figure 4.21 illustrates the loss of resolution with a sample simulated spectrum from LiTrack before and after the expected blurring. The major features remain visible, but we lose information about the small peaks and valleys in the spectrum.

Knowing that there will be limited resolution, we have several options when we seek to compare data with the perfect fidelity spectra from simulation. As was used in the relative calibration of the Cherenkov and X-Ray spectrometers, the technique employed in this thesis is simply to take the "ideal" spectra output by LiTrack and blur them out as shown in Figure 4.21. Then a comparison between simulation and experiment will be on an equal footing.

If the electron optics were better, the blurring could be reduced by several times, allowing finer details of the spectrum to be visible. That greater information would allow even better precision in the comparisons with simulation discussed next.

4.4 LiTrack

We now discuss some of the specifics of the simulation program LiTrack. This software suite for simulation is used to understand the beams from the SLAC main accelerator, and provides the energy spectra against which we can compare our data.

LiTrack was developed by K. Bane and P. Emma of SLAC [78]. Originally written for IBM VM machines, this code was rewritten more recently in MATLAB. LiTrack has been benchmarked against ELEGANT [66] for accuracy. Although it is a general code for linear accelerators, it was originally developed for modelling SLAC.

In many accelerators, longitudinal phase space can be investigated independent of the transverse phase space properties. LiTrack tracks particles only in the 2D longitudinal phase space given by z and $\delta \equiv \Delta E/E_0$, for a large increase in speed.

Acceleration is represented as a pure sinusoid without approximation, and compression from any optical element with bends is represented directly by the path length versus energy coefficients up to third order, e.g. R_{56}, T_{566} and U_{5666}. Wakefields are included as the point charge wake, which is then applied to the evolving time profile of the bunch. The effects of synchrotron radiation are also included.

Unlike slower particle tracking codes such as ELEGANT which explicitly define every optical element in the beamline, LiTrack combines many elements into single units. For example, the first third of the accelerator is rendered as a single object 809.5 meters long which has a final energy of 9 GeV and accelerates using an rf wavelength of 10.4969 cm. Longitudinal wakefields as discussed in Chapter 3 are calculated for the SLAC linac and their effect is included as a single element.

Similarly, the entire RTL is rendered as a single object with a defined R_{56}, T_{566} and central energy for the beam. The U_{5666} is small and ignored. For complex beam lines such as the RTL, these values are normally calculated using another code, such as ELEGANT. Lumping elements like this greatly speeds execution compared to alternative codes such that a typical simulation of the entire SLAC accelerator from the damping ring to the FFTB with 400,000 macroparticles takes about 25 seconds on a PC having a 2.4 GHz Pentium processor and 512 MB of memory.

The number of macroparticles needed for good accuracy depends on the complexity of the system being studied. For conditions similar to those that match our experimental data, experience shows that the results are fairly consistent beyond 100,000 particles and do not change once the number of macroparticles grows beyond about 300,000. To maximize accuracy without unnecessary computational time, we use 400,000 particles for the simulations discussed in this thesis.

With fast execution, it is feasible to simulate the accelerator many times with slight variations in the input parameters. However, there are usually ten parameters which can be adjusted. If each of them were scanned over only three values, we would have 3^{10}, or nearly 60,000 simulations to perform for all possible combinations.

In practice, we focus on several parameters which are likely to vary more than the others to reduce the number of simulations substantially. Even so, performing a set of simulations with several hundred members is commonplace, and would be impractical with a slower code than LiTrack.

Based on measurements in the accelerator, we have an approximate idea of the various input parameters for LiTrack, but the specific combination of parameters for any given data shot is subject to jitter in the machine, so cannot be measured with required accuracy. Thus, we need to simulate many possible states for the machine before trying to match to any given measured energy spectrum.

For a given set of data, we initially perform a rough scan, changing the accelerator parameters by comparatively large values. Once we find simulations that appear to match the data reasonably well, we redo the set of simulations with finer changes in the parameters of interest. This process is usually iterated several times to accurately find the phase spaces corresponding to a particular set of data.

Parameter	Typ. Value	Range	Units
Number of Particles	1.9	$\pm$ 0.2	10^{10}
Bunch Length From Damping Ring	5.5	$\pm$ 0.2	mm
Bunch Asymmetry From D.R.	-0.27	$\pm$ 0.03	-
Phase From D.R.	1.0	$\pm$ 0.5	mm
Compressor Cavity Voltage	42.5	$\pm$ 0.5	MV
Energy Collimation in RTL	2.8	$\pm$ 0.2	%
RTL R_{56}	-58.8	$\pm$ 0.2	cm
Sub-Booster Phase	-19.25	$\pm$ 0.25	degrees
Overall Linac RF Phase	1.0	$\pm$ 2.0	degrees
FFTB R_{56}	-1.5	$\pm$ 0.5	mm
Compressor Phase	90		degrees
RTL T_{566}	105.4		cm
Mean Energy at Sector 10	9		GeV
Sector 10 Chicane R_{56}	7.6		cm
Mean Energy at End of Linac	28.5		GeV
FFTB T_{566}	4		mm

Table 4.2: LiTrack parameters for simulations in this thesis. The top section shows variable values, and the bottom shows fixed parameters.

4.4.1 Parameters in the Linac

For the SLAC main accelerator simulation in LiTrack, ten parameters can be varied to find conditions matching those of the real machine. These parameters, their typical values and typical ranges are given in the first part of Table 4.2. These parameters could be substantially different for other experiments, but the listed values and ranges reflect conditions pertaining to E164 data runs as discussed in this thesis. The parameters which are expected to stay fixed are listed in the second part.

We discuss each accelerator parameter below, and for those which affect timing in the accelerator, we refer to Figure 4.22 to visualize their relationship.

The number of particles in the accelerator tells LiTrack how strong the wakefields will be. Although some investigations were performed with as few as 6×10^9, normally we ran with as many electrons as allowed by instabilities in the damping ring, in the range of 1.9×10^{10}. This particular parameter is measured by various toroids, and is known for every data shot to within a few percent.

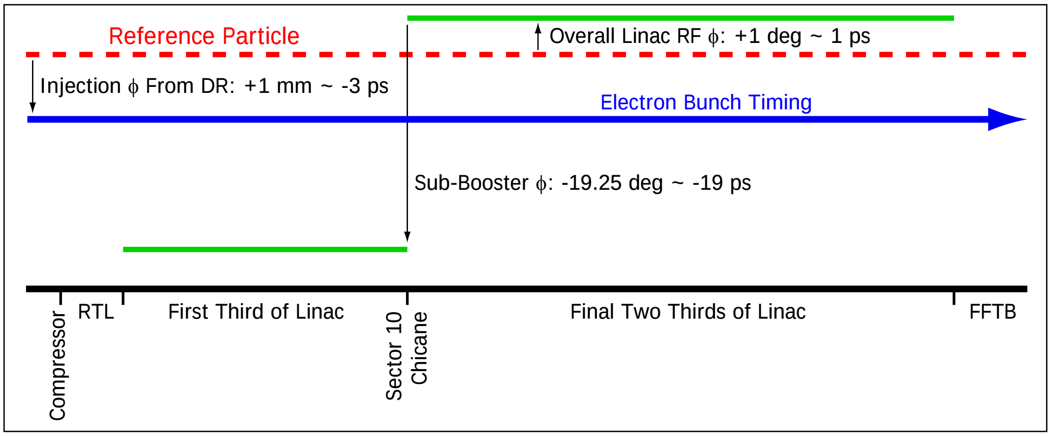

Figure 4.22: Timing schematic for LiTrack's representation of the SLAC main accelerator in E164. Positions above the reference particle line indicate early arrival.

The bunch length and asymmetry of the beam coming from the damping ring are not independent parameters, but are determined by the total bunch charge that we have. We recall Figure 3.3, showing the measured values as a function of current. There were possible systematic errors, so the exact values are not known for a given bunch charge, but our best estimate is that for bunches of 1.9×10^{10} particles, we should use a bunch length of 5.5 mm and asymmetry of -0.27. Simulations with these values have agreed well with the data.

When we specify phases, they can be thought of as being taken relative to the timing of a speed of light particle traversing the main accelerator which had hit the compressor cavity exactly at zero crossing. This speed of light particle in the simulation is equivalent to the Master Oscillator in the real accelerator, which propagates timing information down the full length of the machine.

LiTrack treats the Compressor Cavity as always fixed in phase relative to the reference particle, but we may inject at various times into the cavity. The parameter known as "Phase From Damping Ring" is what we use to change injection timing, and has units of mm. This seemingly odd choice of units may be because LiTrack deals directly with longitudinal coordinates z and δ.

We normally inject the electron beam somewhat after the reference particle. This is the parameter known as the phase from the damping ring, and is a positive quantity expressed in millimeters, with a value typically around 1 mm. Thus, there is a slight net deceleration of the beam in addition to the large energy chirp. This timing offset on injection into the accelerator propagates all the way through the main accelerator.

The voltage of the Compressor Cavity affects the first stage of compression. For different desired levels of compression, we can change this value, but it remained very close to 42.5 MV during the experimental run discussed in this thesis.

Although it was never changed during the experiments relevant to this thesis, the exact value of the R_{56} is not known exactly. It has been calculated to be in the region between -58.6 and -59.2 cm, and we have found best agreement with data using a value of -58.8 cm in LiTrack, so we use that value and do not scan this parameter in our series of simulations. The T_{566} value has been calculated to be -1.054 m and is also not scanned here.

The energy collimation in the RTL does not come from a specific object. Rather, it reflects the loss of some particles as they transit that complicated beamline with various places of reduced aperture. Because we do not have a specific aperture, the energy collimation that occurs throughout the RTL is also not known exactly. Good agreement was observed with an energy acceptance half width of 2.8%, and this parameter was not scanned.

The overall linac phase is phase shifted relative to the Master Oscillator in the real accelerator. In LiTrack, the speed of light particle traversing the compressor cavity at zero-crossing will go through the linac on crest if the overall linac phase is set to zero. When this quantity is positive, the reference particle will go through the linac behind the wave crest, because a positive value means that the phase has advanced ahead of the nominal position.

The first third of the linac is run, as discussed in Chapter 3, at $-19.25°$, placing the beam ahead of the rf crest in order to give the tail particles higher energy than the head for compression in the compressor chicane. This normally fixed value is known as the sub-booster phase, and is not absolute relative to the reference particle, but rather is added to the overall linac phase. Thus, if the overall phase were $+2°$ and the sub-booster phase were set to $-19.25°$, then the actual rf crest in the first third of the linac would be only $-17.25°$ behind the nominal reference particle. We recall that

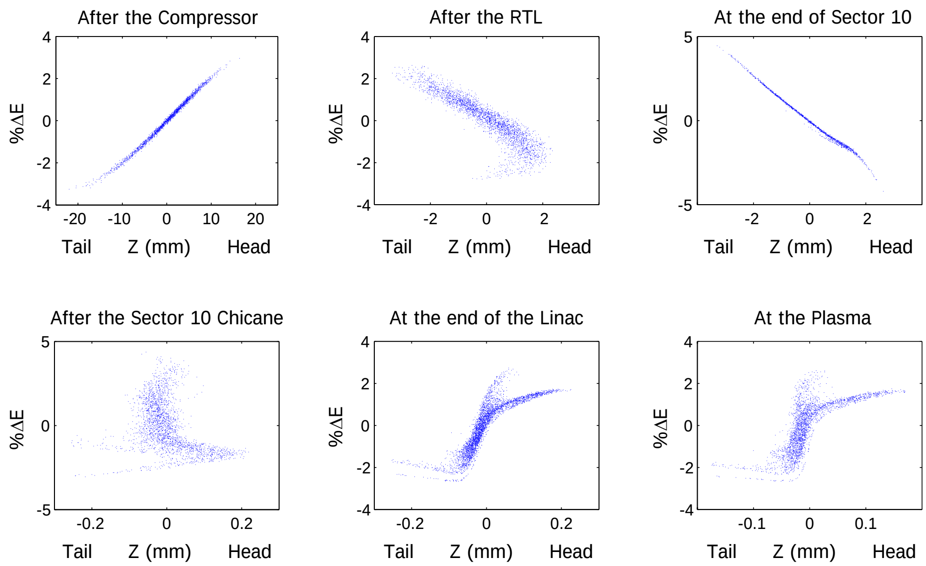

Figure 4.23: Progression of the beam phase space through the linac.

the bunch leaving the damping ring is usually slightly behind that reference particle, so the center of the bunch will be even *less* than 17.25° ahead of the wave crest.

The rf in the last two thirds of the linac is simply at the overall phase, and so bunch is only a few degrees behind the rf crest for near-maximum energy gain. The energy correlation we impart comes almost completely from the longitudinal wakes.

Sample phase spaces from the various elements of the accelerator have been shown in Chapter 3, and we include them again in Figure 4.23 to show the evolution of the phase space clearly. Note that energies are expressed as a fraction of the central energy, which changes by a factor of about 25 through the linac.

We have seen the evolution of a beam responding to the many elements in the main accelerator as illustration of the physics. The plots of the phase space in Chapter 3 come from particle dumps out of LiTrack at various points through the accelerator.

4.5 Finding the Phase Space for Data Events

When we seek to know the phase space of the electron beam for any given shot of the linac, we must compare its energy spectrum with phase spaces from simulations.

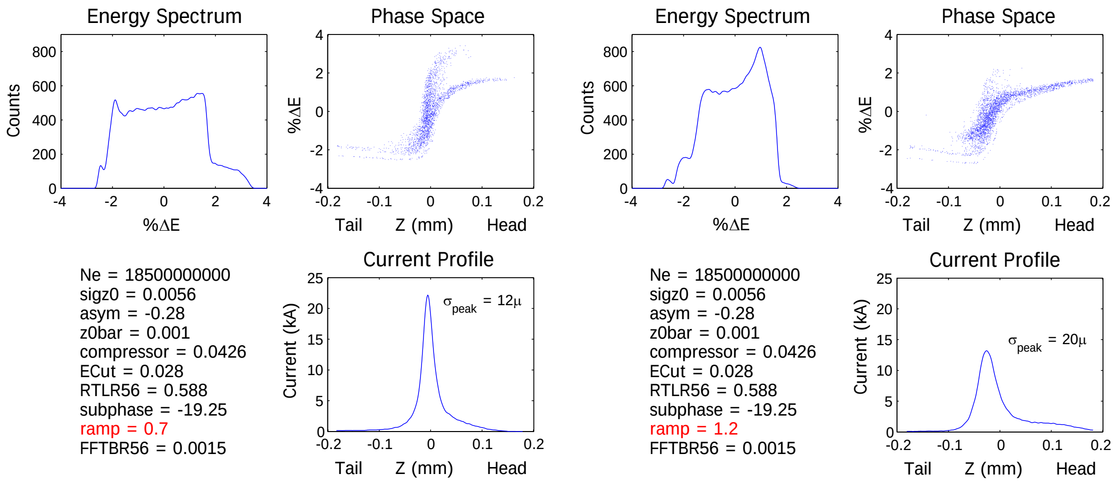

Figure 4.24: Plots of energy spectra and associated current profiles for two different values of overall linac phase. Less than half a degree leads to substantially different beams in the FFTB.

4.5.1 Simulation Parameters for this Thesis

As mentioned above, we do not know the exact state of the main accelerator at any given time. Luckily, we do know the approximate state, and therefore, we allow some of the parameters discussed in the previous section to vary. The overall linac phase, in particular, is known to shift in a range of a degree and sometimes more over time scales as short as a minute. Any given data run of, for example, 100 shots will take 100 seconds. During that span of time, there can be variations in phase of at least half a degree between the various shots. This half of a degree in overall phase can lead to a substantially different beam coming from the accelerator. Figure 4.24 shows the difference that 0.5° can make to the energy spectrum and current profile, even if all other parameters are held constant.

The other parameters above are less likely to vary on scales with such dramatic impact, but several have some effect. For the investigations discussed in Chapter 5, we performed a suite of simulations where three possibly varying parameters were allowed different values. In particular, we simulated conditions where the phase from the damping ring varied between 0.7 and 1.1 mm. The uncertainty in the compressor cavity voltage is larger than that required to change the final phase space, so we allow it to have values of 42.3, 42.5 and 42.7 MV. Lastly, the overall linac phase varies over roughly a one and a half degree range between 0.3° and 1.9°.

In this instance, there are a total of 255 simulations we can compare against each data shot of interest. It is certainly possible to create a vast suite of simulations by varying all of the parameters in the simulation, but these are the three that experience has taught are most likely to vary and have an effect on the final phase space.

4.5.2 Comparison of Simulation and Data

After performing the series of simulations, we save a file which has information on the input conditions to LiTrack for each simulation, as well as each resulting energy spectrum and current profile. With this data file in hand, we can then compare the simulations to data in a separate program.

To perform the comparisons, we have a program written in MATLAB which first loads the incoming beam energy spectrum images which have been stored for later analysis such as this. It then sums the images sideways as discussed above to get spectrum curves that span several hundred bins. There is some background light in the images, so that is removed by taking the average of the first ten plus last ten bins and subtracting that value from each point.

Next, we load the suite of simulations previously performed. We bin them with our 5.5 MeV/pixel calibration to match the binning of the data.

Because we know that the data has some blurring in it, we then convolve each simulation with a unit area Gaussian whose RMS width is 10 pixels, or 0.21%.

At this point, there remain only two unknowns for the comparison, as with Cherenkov vs. X-Ray comparisons discussed in the section on calibration of the energy per pixel on our camera. There are energy feedbacks within the accelerator, and sometimes individual klystrons will go off line. As a result, the centroid energy of our electron beams varies by up to a few hundred MeV out of 28.5 GeV. This manifests itself as the spectrometer images moving up and down modestly. Summed sideways, and plotted, we see the spectrum curves moving sideways by a noticeable amount, though it is still smaller than the overall typical width of the curves.

Small changes in the total energy of the beam in the linac should have no significant effect on the various effects which allow us to produce short beams at the end of the FFTB. The simulations are always, by construction, perfectly centered on $\delta = 0$. Therefore, each simulation is allowed to move back and forth in the matching routine to get the best overlap with the data.

In the vertical direction, the scales are quite different. The spectra come from 12 bit digitized images. Each pixel therefore has a value that can range as high as 4095 and we sum a whole row of pixels, so the points on our spectrum can be over 100,000.

The simulations are typically run with 400,000 macroparticles. When binned across several hundred bins, the spectra have a maximum number of these particles per bin that is typically less than 1000. So in the fitting routine, the simulations are simply multiplied by a constant to account for the mismatch. The seed value for this constant can be calculated by the sum of the data spectrum curve divided by the sum of the simulation spectrum curve.

With approximate initial values, we can then perform the comparison. The heart of the program is a routine which takes the two curves, and calculates the χ^2 between them, seeking to minimize this value by varying the sideways shift of the two images and their overall vertical size. To calculate the χ^2, we simply take the raw difference at each bin, square it and then sum over all bins. The only difference from a standard χ^2 is that we do not normalize by the errors, as they are not known here and are assumed to be constant. For shorthand, we will refer to this value as the χ^2 with the understanding of its modified meaning. The process is nearly identical to that described above for comparison of spectra from the Cherenkov screen and the X-Ray based spectrometer, and output from the fitting routine looks much like Figure 4.19.

To minimize χ^2, we give the fit the ability to vary only the two parameters mentioned above. When the best overlap is found by the fitting routine, we record the χ^2 that was achieved between that particular simulation and the data spectrum. Clearly, if the simulation spectrum has a shape and overall width that is similar to that of the data, we get good overlap, and if the shape or overall width is wrong, then no amount of stretching or translating the curve sideways will allow the χ^2 to be small and the fitting routine will return a large value from the optimization.

Having found the best χ^2 for a given simulation and the data, we then compare the same data against the next simulated spectrum. We record which simulation yielded the smallest χ^2 and then repeat the entire process for the next data shot.

The number of times the fitting routine is called goes as the number of data shots times the number of simulations, so it is beneficial to have as small a number of simulations against which to compare the data of interest as possible, while still spanning the space of accelerator parameters that pertains to any given run. As mentioned in the previous section, 255 simulations was enough to provide a decent match to almost all of the data discussed in this thesis.

Having performed the comparisons, we then save a file with the information about the simulation which matched any given data shot, as well as the quality of that matching. Separately, we calculate properties of the images from the Cherenkov spectrometer and store them.

A suite of MATLAB analysis programs has been complied by our group at SLAC, Accelerator Research Department B, and all data is formatted so that it can be read in and analyzed using this suite. We can thus easily plot calculated parameters of the bunch against measurements from Cherenkov images or from the accelerator, such as the charge or total amount of CTR power produced by any given bunch. This software is the source of many of the graphics in Chapter 5.

With all of these tools in hand, we can now investigate the longitudinal phase space and its effect on ionization, acceleration, and other interactions of the beam and the plasma.

Chapter 5

Data and Results

Knowing the longitudinal phase space of the electron beams in the E164 experiment is important for understanding a variety of effects that occur in the plasma, primarily, the acceleration of the beam itself. To have confidence in the technique, we first seek to find verification that matching with LiTrack provides accurate data. There are several ways to test this method.

5.1 Verifying the Technique

As our technique is indirect, we have sought a variety of sources of confirmation that it is accurate. When we calculate the beam profiles, we can then correlate the properties of each bunch with several other measurements. If these all track one another in the expected ways, that greatly increases confidence in the method.

To evaluate the accuracy of our technique, we use data from several runs taken in 2004. As illustration, we focus on several 100 shot sequences of data taken at 1 Hz on 13 July of 2004, known in our database name convention as runs "07131cw" and "07131dc." In all of the runs from this day, the accelerator was set to create the shortest bunches possible to investigate acceleration at the very highest plasma densities achievable with our oven: 3.5×10^{17} cm^{-3}. The natural variations of the accelerator's various parameters mean that we would expect a variety of beam profiles at the position of the plasma.

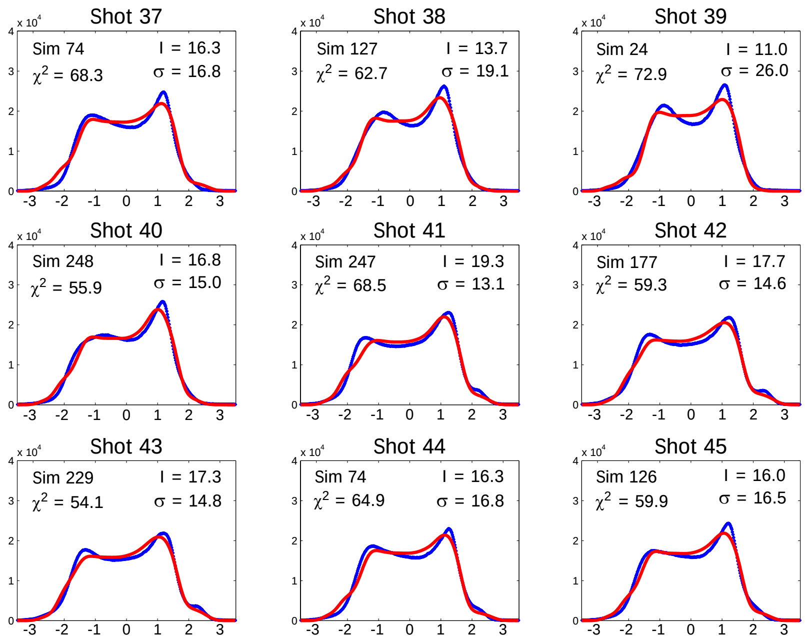

Figure 5.1: Energy spectra for nine sequential shots in Run 07131dc, each matched to one of the 255 simulations with slightly varying parameters. Horizontal axis is in units of δ. Blue is data and red is the best matching simulation spectrum. The peak current in kA and bunch width in μm for the corresponding simulation are indicated.

With all parameters close to optimal, we produce beams as short as 12 μm with corresponding peak currents of about 20 kA. When any of the various linac parameters discussed in Chapter 3 jitter away from optimum, the accelerator produces substantially longer bunches with less intense peak currents. This natural variation allows us to probe several phenomena of interest and is what allows us to see if calculated bunch properties do, in fact, track other measured quantities.

Figure 5.1 shows the matching of our simulations to nine shots from Run 07131dc. Each shot has a slightly different energy spectrum, but still matches one of the simulations, which are different from one another in just a few of the possible variables. Specifically, the compressor cavity voltage is not well known, so is allowed to take the values 42.3, 42.5 and 42.7 MV. Similarly, the phase of our bunch coming from the damping ring is allowed to vary somewhat in the range from 0.7 to about 1.4 mm. Lastly, the overall phase of the linac was scanned over a range of about one and a half degrees, with most shots determined to lie within a range of 0.4°.

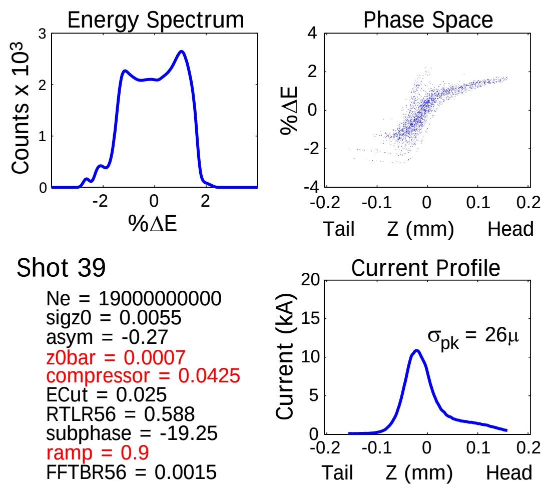

Figure 5.2: Simulation phase spaces, the left shows the long current pulse associated with Shot 39 of Figure 5.1 and the right is the short pulse matching Shot 41.

The overall phase of the accelerator has been measured to vary on short (less than 30 second) timescales, so we always recreate this variation in the suite of simulations for matching. The compressor cavity temperature is also known to vary, sometimes rapidly, when PEP requires changes in the main accelerator. Similar changes are observed throughout the linac, so there are both long term diurnal drifts and much more dramatic jumps of as much as a full degree in just one or two seconds. As discussed in Chapter 4, changes of even 0.1 or 0.2 degree can make appreciable changes to the energy spectrum. This parameter has the largest variability compared to the scale at which changes in it make observable differences in the energy spectrum. Overall linac phase must always be scanned in doing a set of simulations.

For reference, Figure 5.2 shows the phase spaces calculated for shots 39 and 41. They represent occasions where small changes in the linac led to peak currents which were a factor of two different, even though the initial conditions were similar on the scale of our ability to measure the three important quantities highlighted in red.

Determination of the current profile allows us to make a variety of predictions about the behavior of the electron bunches in the plasma and at other diagnostics. We discuss four different ways in which we connect the determined profiles with other measurable quantities, as a means of verifying the technique.

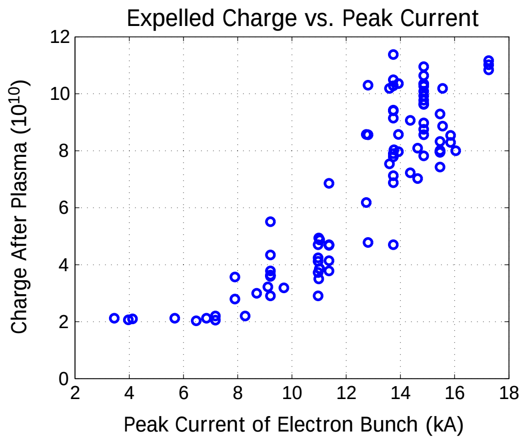

Figure 5.3: The total charge measured downstream of the plasma is related to the peak intensity of the incoming electron beam. Data is from Run 07131cw.

5.1.1 Trapped Charge Measurements

The mechanism is not yet understood, but we have observed that immediately downstream of the plasma, more charge often comes out than was sent in, by a factor of up to five or even more. This is probably some sort of trapping of plasma electrons by the strong wake, analogous to that seen in laser wakefield acceleration experiments.

Trapped particles must start from rest, rather than ~ 28.5 GeV, ending with only a few GeV of energy. They will be lost in the sequence of strong quadrupoles downstream of the plasma and are not seen at the diagnostic Cherenkov screen.

The downstream charge measuring toroid is close to the plasma, so could be influenced by the significant radiation there. We also cannot yet measure the energy of these particles, which could allow us to understand sources for what we see.

Nevertheless, for the above data run, if we plot the peak current of each individual machine shot against the measured charge flowing down the pipe immediately after the plasma, we clearly see that the more intense incoming bunches associate with much larger amounts of charge downstream. See Figure 5.3.

The exact mechanism of trapping is not understood, but it seems reasonable that a more intense electron bunch will drive the plasma wake harder and be more likely to trap particles. As we have substantial variation in the length of the 100 bunches in this run, we have a natural way to see the effect of changing bunch length on

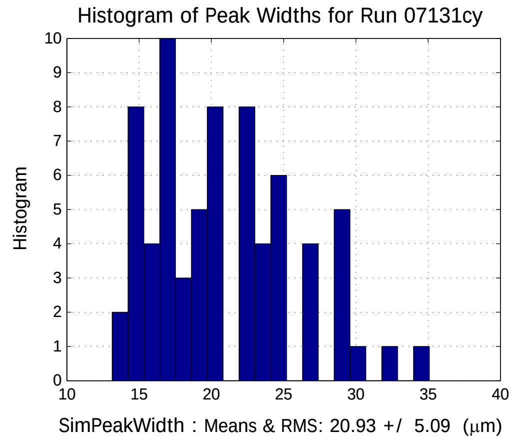

Figure 5.4: Histogram of the calculated bunch lengths in microns for Run 07131cy.

trapping. This provides an interesting result as well as evidence that the simulations do give an understanding of the beam's incoming phase space.

5.1.2 Autocorrelation Measurements

The autocorrelation discussed in § 4.2 is a multi-shot measurement that gives some idea of the bunch length for a given set of accelerator conditions. The measured bunch length is of the order 18 μm long, in reasonable agreement with the shortest bunches that simulation tells us that we can achieve.

As we have seen, the accelerator varies in its input conditions over time scales shorter than the multiple minutes required to build up the autocorrelation trace. Whenever the accelerator strays from ideal conditions, the bunches will be longer.

For example, if we look at a histogram of the peak widths, as determined by comparison with simulations, from Run 07131cy, we obtain Figure 5.4. This shows that we get substantial variation of bunch lengths from roughly our expected minimum of 12 μm to 30 and 40 μm. The most common length is of order 20 μm.

We recall that in choosing events to include in the autocorrelation trace, we had used the most probable total CTR power value as a cut. If total CTR power is inversely related to bunch length, we expect that we had selected the most common bunch length to measure. The autocorrelation value of 18 μm for that bunch length

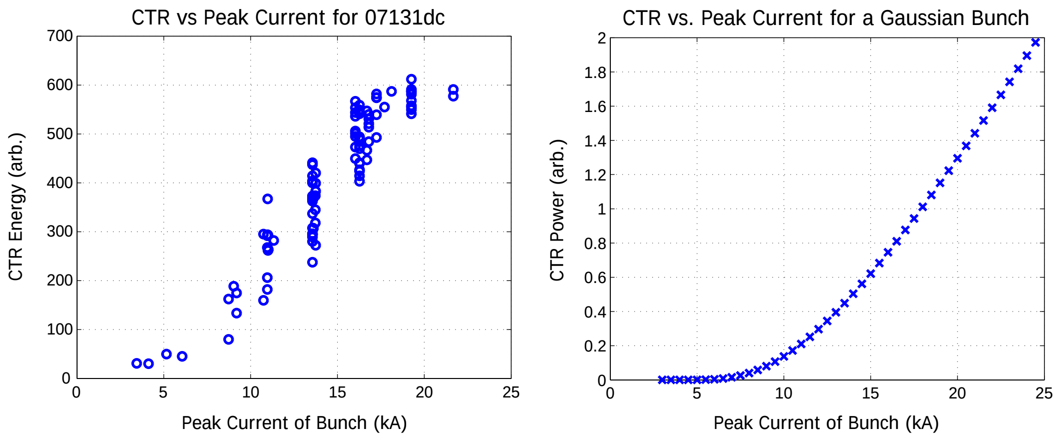

Figure 5.5: Plot of total Coherent Transition Radiation power as a function of the peak intensity of the driving electron bunch. Data is from Run 07131cw.

therefore agrees quite well with the bunch length derived from simulation. This gives confidence in both methods, as they are totally independent of one another.

5.1.3 Pyro Peak Measurements

The autocorrelation measurements give confidence that our short bunches have an overall absolute length close to that estimated from the simulation comparisons. A further piece of useful information comes from the simpler measurement that measures only the total CTR power coming from the bunch. Clearly, we cannot perform the bunch length scan for each shot of the accelerator, but we can measure the total broadband power emitted by each individual bunch on a shot by shot basis.

To find the expected total power for bunches of various lengths, we integrate Equation (4.5) numerically. For each bunch length, the minimum wavelength is taken to be 0.6 times σ_z. As discussed in § 4.2, wavelengths beyond about $100\,\mu$m are progressively attenuated by the apparatus itself, so we choose to stop the integration at $150\,\mu$m.

We can integrate over all angles, and the effect of the acceptance of our system is taken care of by the choice of limit on wavelength above.

The left hand side of Figure 5.5 shows our measured CTR power plotted against the peak current as inferred from simulations for Run 07131dc. Clearly, short high current bunches produce more CTR than the longer low current bunches. The right hand side of the same figure shows the theoretical curve we expect.

That the total CTR power we detect as a function of the bunch's peak power follows the general trend that we expect from theory gives confidence that we do understand the longitudinal profile of our bunches.

5.1.4 Comparison with Ionization Measurements

A less direct verification of our understanding of the beam properties comes from understanding the process of Lithium ionization. Because we seek to use our intense beams to ionize the Lithium, it is important to verify the threshold at which ionization occurs. One of the ways to investigate this is to change the electric field of the bunch in a controlled way and see at what point ionization is initiated.

Caolionn O'Connell, in her thesis on the E164 experiment [41], investigated the onset of Lithium ionization for different beams. Understanding the bunch longitudinal profile was particularly useful for the case where the beam was changed in length to see when it would finally have strong enough fields to dissociate Lithium.

With the high charge bunches that we normally use: 1.8×10^{10} particles, the ionization of Lithium has been observed to happen readily, so in this investigation of the threshold for ionization, the bunch charge was dropped to about 0.9×10^{10}.

The spot size at the entrance to the plasma was held at approximately $15\,\mu$m. Thus the peak fields associated with the different bunches were determined only by their instantaneous currents.

As we have seen in § 2.3, Lithium ionizes rapidly once the space charge fields of the beam rise above roughly 5 to 6 GV/m. Given the instantaneous current of the highly relativistic bunch, then the radial electric field at all points in the associated infinitessimally thin pancake at that z location is given by:

$$E_r(r,z) \;=\; \frac{I(z)}{2\pi\epsilon_0 c\,\sigma_r^2\,r}\int_0^r e^{-\frac{r^2}{2\sigma_r^2}}\,r\,dr \;=\; \frac{I(z)}{2\pi\epsilon_0 c}\,\frac{1}{r}\left(1-e^{-\frac{r^2}{2\sigma_r^2}}\right) \tag{5.1}$$

We recall that the pancake's peak field occurs at $r = 1.6\,\sigma_r$. With a transverse size of 15 μm, this peak electric field of our bunch is given by the engineering formula:

$$\widehat{E}_r(z)\ [GV/m] = 1.81\, I(z)\ [kA] \tag{5.2}$$

Peak fields of 5 to 6 GV/m thus will correspond to about 3 kA of beam current for our beam radius. Away from $r = 1.6\,\sigma_r$, the fields are lower, so full or nearly full ionization of the ion column is not expected to happen until the beam current rises to 4 or 5 kA.

In one run of 200 images from 16 July 2004, the bunch length (and peak current) was varied by changing the overall phase of the accelerator. Comparison of the X-Ray based spectrometer images with simulation shows that the bunch lengths ranged from about about 130 μm to about 20 μm during this sequence of runs. After selecting only those events within a tight range of charge at the plasma and after removing the few events which failed to match well to any of the set of simulations, 115 shots remain.

It is useful to create a condensed view of all 115 Cherenkov spectrometer images. We take each image, an example of which is the right hand one in Figure 4.19, and then sum it sideways to create a single pixel wide lineout. We can then place these lineouts immediately next to one another and order them by the calculated peak current of the incoming electron bunch to get a picture of what is happening to the Cherenkov energy spectrum as the bunch length changes. Figure 5.6 shows such an image, with a plot of the increasing peak current of the bunches below. The overall trend of increasing energy spread at the Cherenkov screen with increasing peak current is very clear.

We see that for the first few images, there is a bit of noise, but there is no ionization. Once the peak current rises a little above 2 kA, the Cherenkov spectra begin to broaden, primarily with particles dropping in energy. The only mechanism available to decelerate these particles in our experiment is that they are driving a plasma wave, so they clearly must have started to ionize the Lithium. With increasing peak current, we see that the deceleration of the particles increases to some maximum, at which point it mostly levels off. There appear to be two interrelated effects causing this behavior. The first is that we go from having no ionization to having full

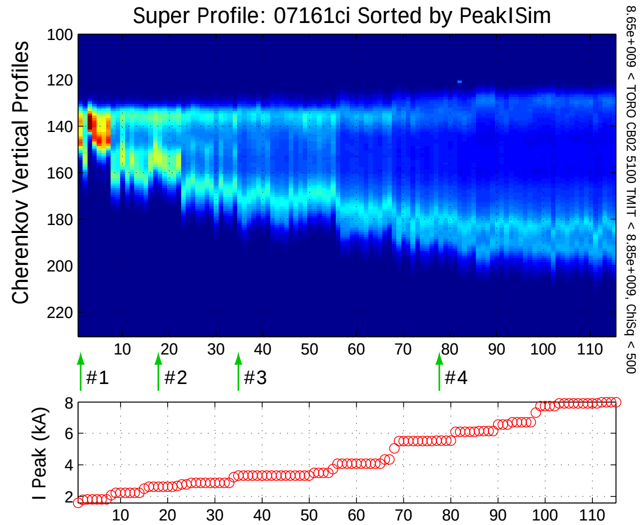

Figure 5.6: Cherenkov spectra summed to single pixel width and ordered by the peak current of the incoming beam as determined through simulations.

ionization, allowing for ever denser plasmas in which to drive wakes. Additionally, the more intense bunches ionize the beam ever earlier within the bunch, so that a larger fraction of the particles can participate in driving a wake. We quantify these statements below.

We know the current profile of the electron beam and its transverse size at the plasma. Thus, we also know the electric field at all points, and can use Eq. (2.47) to find the instantaneous ionization rate. The peak electric fields of our relativistic electron bunch lie at the $r = 1.6\,\sigma_r$ point, so we integrate the total expected ionization in the annular sheath at that optimal radius and plot it for four of the calculated profiles from Figure 5.6, giving Figure 5.7.

The first current profile corresponds to the very first shot. The second profile is that which matched shots 16 through 21, and had a peak current 2.6 kA. The third profile is for shot 34, with a peak current of 3.2 kA, and the last profile is for shots 77 to 80, with peak current of 5.5 kA. For all of the bunches, the ionization at other radii than $1.6\,\sigma_r$ will be lower, which is why, for example, the third profile shows that we expect full ionization, but the corresponding Cherenkov lineout does not yet have the maximal amount of deceleration there. Only part of the plasma sheath

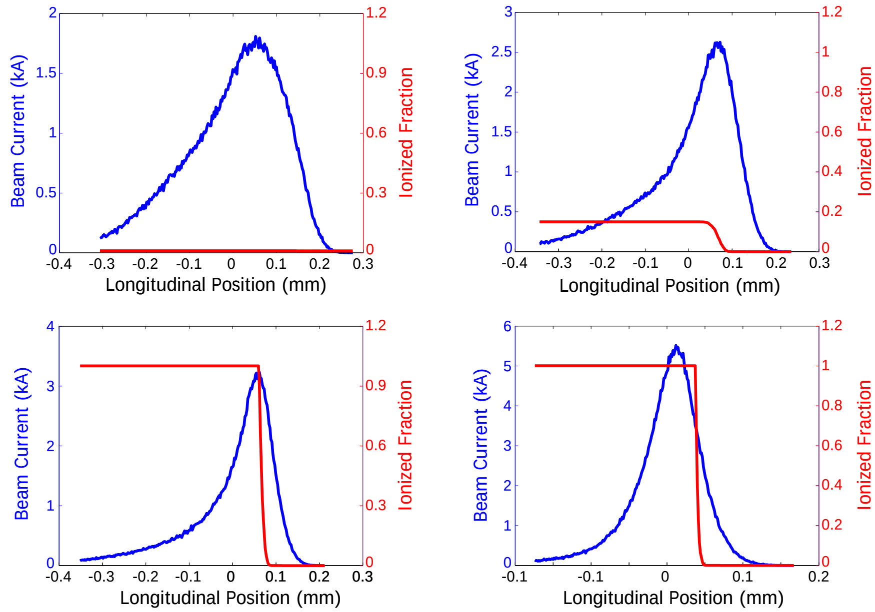

Figure 5.7: Current profiles for the four bunches indicated in Figure 5.6. Ionization is rapid compared to the bunch length above about 3.5 kA.

experiences full ionization and therefore can produce the largest fields available. Once the ionization is complete at some early point in the bunch, the wakes we drive should not change significantly, which is what we see for the final third of the shots.

In looking at these various plots in conjunction with Figure 5.6, it is clear that ionization at $r = 1.6\sigma_r$ turns on rapidly compared to the beam transit time when the current rises above 3 or 3.5 kA, corresponding (for a beam radius of $15\,\mu$m) to peak fields of about 5 GV/m, just as predicted by theory. Additionally, the *entire* ion column seems to be ionized for peak currents above 5 or 6 kA. Our understanding of the bunch profiles and of ionization reinforce each other very well, and these numbers have immediate consequences in the following section.

5.1.5 Post-Plasma Energy Spectrum Features

During the summer 2004 data run, we observed an unexpected feature on the Cherenkov based electron beam spectrometer. Specifically, there was a large amount of charge at or near the highest incoming energies which appeared to be unaffected by passage

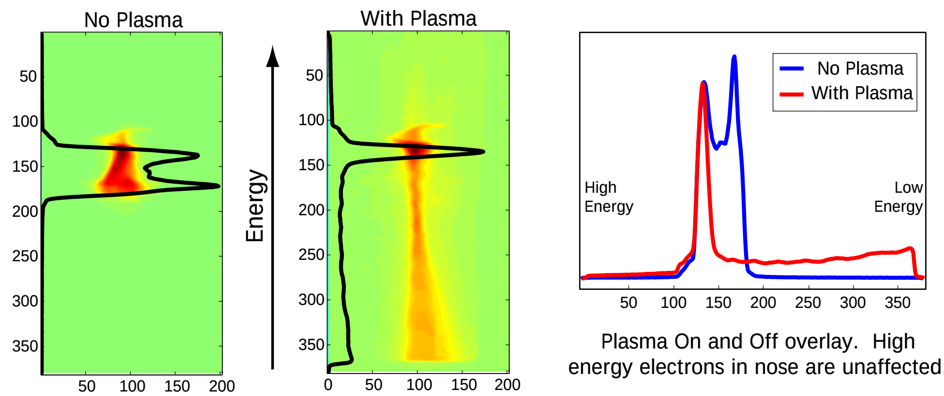

Figure 5.8: Two Cherenkov spectrometer images for beams with nearly identical incoming energy spectra. On left is a shot without, right is with plasma. High energy particles for both shots make it through with no effect as shown on right hand overlay. The axis corresponding to energy has units of pixels on our camera.

through the plasma. With a normal Gaussian bunch that ionizes the Lithium early on, so that almost all but the tailmost electrons are decelerated, this is puzzling. The feature was better explained after the experimental runs, when we simulated the accelerator in detail to match the observed energy spectra.

For data runs from 13 July 2005, we have two shots with nearly identical incoming energy spectra. The first Cherenkov spectrum was with no plasma in the beam path, and the second is of the beam after going through the plasma, see Figure 5.8. With these two shots, we can compare the Cherenkov energy spectra to get a good idea of what changes downstream as we insert the plasma cell into the beam. For reference, Figure 5.9 shows the phase space we have determined for this pair of shots.

If we look closely at the current profile of the bunch in Figure 5.9, we note that there is a long "nose" at the front where the current of the bunch remains low for about 100 μm. As we saw previously, until the current rises above 3 to 5 kA, the Lithium is not ionized. With no plasma for the initial portion of the bunch, nothing should happen to the energies in this population of high energy incoming electrons.

The "Nose" of our bunch, before the current ramps up to over 5 kA, contains roughly a third of the total charge in this instance. We have also observed that in many cases, about one third of the total charge goes through the plasma unaffected.

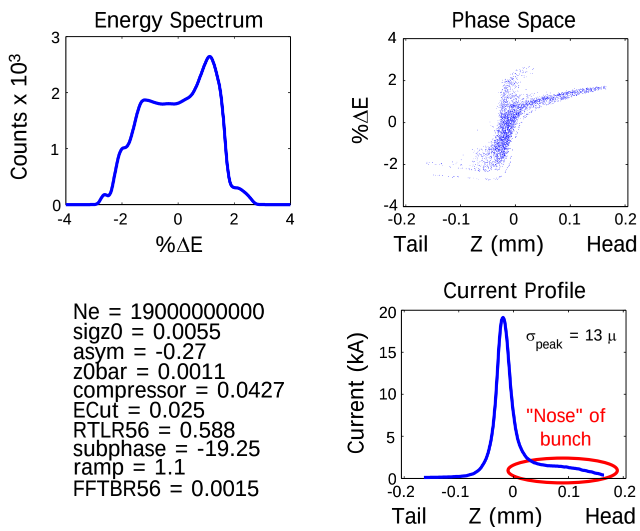

Figure 5.9: Plot of the simulated phase space for the two accelerator shots discussed in this section. One shot had plasma off, and one had the plasma on.

Our understanding of the bunch current profile and of Lithium ionization allow us to explain the initially puzzling collection of charge at high energy that we observed on the Cherenkov spectrometer even with the plasma cell in.

5.1.6 Uses for Knowledge of the Phase Space

Having verified through these various techniques that simulations do, indeed, tell us the phase space, we can move on to understanding various aspects of the beam's interaction with the plasma. We use knowledge of the beam's phase space to inform investigations of hosing instabilities, and of greatest importance, the plasma acceleration effect itself.

5.2 Understanding Acceleration

The main motivation for seeking to know the longitudinal phase space of the beams is naturally so we can understand acceleration more fully. We can use the knowledge of the phase space in several ways.

5.2.1 Direct Acceleration Determination

An important application of the phase space reconstructions is to use them to determine the magnitude of acceleration that we actually have achieved in our accelerator. The technique discussed here is similar to that presented in [79], and the basic idea is that identifying the precise incoming energy of particles in the tail allows a more accurate determination of their overall energy gain.

In E164, we have a short plasma and as discussed previously, our incoming bunch has a substantial energy spread in order to make it short enough for our experiment. As we see above in Figure 5.9, the intrinsic energy spread of the beam coming into the plasma has a full width of 4%, or nearly 1.2 GeV.

With peak gradients expected to be less than 50 GeV/m over a plasma of only 10 cm, the best possible energy gain does not dwarf the beam's own energy spread. As a result, there is potentially substantial uncertainty in determining the actual energy gain by particles when we look at the images of the Cherenkov spectrometer downstream of the plasma.

If we can identify the particles which are being accelerated and know their incoming energy, then we can improve the accuracy of our gradient measurement.

Identifying Particles

We have seen in a variety of simulations of our plasma wakefield accelerator that the wake forms a surprisingly long distance behind the bunch. In linear wakefield theory, we would expect, as shown in Chapter 2, cf. (2.41), that the wake will be at a maximum when $k_p\xi = \frac{\pi}{2} \rightarrow \xi = \frac{\lambda_p}{4}$ behind the bunch. This is because the electrons are assumed to have been blown out gently and immediately recollapse after the peak current portion of the beam passes. From fully blown out to recollapsed back to the axis is one quarter of a plasma wavelength.

Related to this, in linear theory, we see that there is an optimal bunch length for any given plasma density, where we match using the condition that $k_p\sigma_z = \sqrt{2}$ for Gaussian bunches. For too long a bunch, the wake is not strongly driven, and for too short a bunch, there are no particles available in the tail to be accelerated.

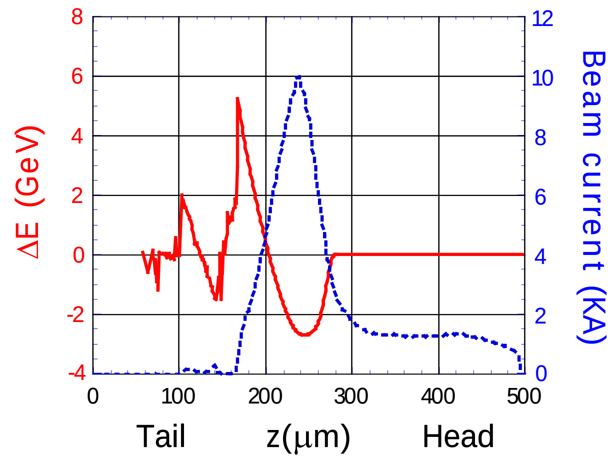

Figure 5.10: Plot of wake developing behind a bunch simulated in QUICKPIC. The plasma density is 2.8×10^{17} cm^{-3} with an associated plasma wavelength of $\lambda_p = 63\,\mu$m. The maximum accelerating field can be seen to lie almost a full λ_p behind the peak current point of the electron bunch. Adapted from [80].

For parameters as in E164, plasma electrons are expelled ballistically by the short electron beam passing through. After the point of peak current, the plasma electrons are still moving *away* from the beam axis before they are ultimately pulled back in by the ion column. Being relativistic, the electrons cannot accelerate arbitrarily as they would in simple harmonic motion. Thus, the plasma electrons oscillate with a phase delay and a substantially longer period than predicted by the naive linear theory, where we assume motion as in a classical harmonic oscillator.

With these two effects, the wake forms not at a distance behind the main current pulse of $\xi = \frac{1}{4}\lambda_p$, but at $\xi \lesssim \lambda_p$. The simulation output shown in Figure 5.10 shows this long distance. The simulation was performed for a 10 kA peak current beam in a plasma of density 2.8×10^{17} cm^{-3} with associated plasma wavelength of $63\,\mu$m. The wake is strongest about $65\,\mu$m behind the peak current point of the bunch, just in front of the point where the blown out plasma electrons come back to the axis.

With this, we know that the particles which will be accelerated are relatively far into the tails of our bunch. With even our shortest available plasma wavelength of 56 microns, the distance behind our bunch for largest acceleration is more than one or two σ_z behind the peak.

The wake evolves from decelerating to accelerating to decelerating again as we move progressively backward behind the bunch. This means that electrons in the tail will see more or less acceleration as a function of their position.

In an ideal plasma accelerator, we would have a very short bunch of electrons trailing the main drive bunch in a position such that all of the witness bunch is accelerated strongly and equally. With the single bunch setup of E164, we have particles at all of the phases of the evolving plasma wake, so some are accelerated more strongly than others. This manifests itself as the jet of accelerated particles that we see on our Cherenkov detector. The particles with highest energy are whichever lie at exactly the best longitudinal position behind the bunch.

Finding the Acceleration Strength

An example data event from 13 July 2004 where we seek to know the gradient showed many particles being accelerated. On the right side of Figure 5.11 we see the energy spectrum as measured on our Cherenkov diagnostic after the plasma with a strong jet of particles above the incoming maximum energy. We can see the substantial number of particles which have gained a continuum of energies at least up to the top of the screen, about 2 GeV above the highest incoming energies.

Improvements to the E164 Cherenkov based spectrometer have given a larger energy acceptance so that all particles will be viewable, but here we are constrained by the available hardware, and some of the particles may have been accelerated by more than we could see.

Using the measured energy spectrum from before the plasma, we have found the phase space for the above event, as shown in the left portion of Figure 5.11. The intrinsic energy spread is quite large.

At a density of $3.5 \times 10^{17}\text{cm}^{-3}$, with even our shortest plasma length of about 10 cm (see Chapter 4), there is enormous acceleration. Recall that when we claim the plasma is 10 cm long, that refers to its FWHM, and the actual flat top portion which is at the full nominal plasma density is only a few cm long.

The maximally accelerated particle must have seen an accelerating gradient of at least 2 GeV in 10 cm, or 20 GeV/m. However, the peak gradient achieved in the

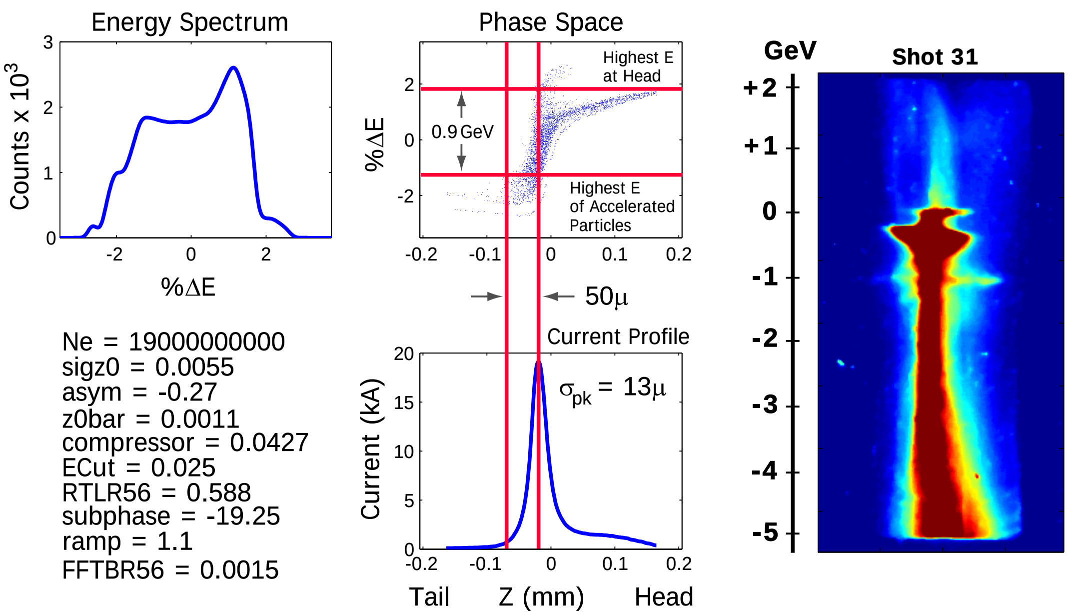

Figure 5.11: Incoming phase space for Shot 31 of Run 07131co and post-plasma spectrometer image showing substantial acceleration.

full density flat top must be higher than this. Further, we can show that the actual energy gain to these particles was even greater than the observed 2 GeV.

As discussed in the previous section, best acceleration happens nearly a full plasma wavelength behind the point of maximum current in our bunch, and at a density of 3.5×10^{17} cm^{-3}, the point of maximum acceleration thus lies about 50 μm behind the position of the bunch's peak current. In Figure 5.11, we draw a line on the simulation phase space corresponding to that position.

As we can see, the particles in a position to be accelerated all lie below about -1.5% in energy relative to the mean particle in the bunch. The head of the bunch, which we have seen will be unaffected by the plasma, has its highest energy particles at nearly +2%. Just for the accelerated particles to become visible, they must have gained more than 3%, or over 0.9 GeV of energy. Thus, when we see 2 GeV of gain, the actual acceleration was greater than we can directly observe, and our peak gradient at 3.5×10^{17} cm^{-3} is now calculated to be nearly 30 GeV/m. Further discussion of this technique for direct gradient determination can be found in [79].

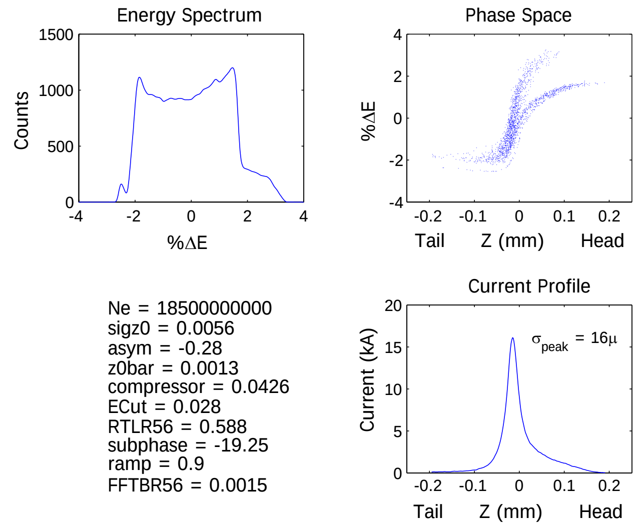

Figure 5.12: The phase space to which we match the three shots shown in this section.

5.2.2 Acceleration Properties Under Varying Conditions

Naturally, we seek to understand the acceleration for more than just one plasma density. Toward this end, we performed a series of data runs in July of 2004 with substantially different plasma densities and therefore, different plasma wavelengths. Although linear theory does not apply, there is still an optimal bunch length for each plasma density. Long bunches drive gentle wakes, but have many particles in the tail to be accelerated. Short bunches create very strong wakes, but have few particles far enough behind the main portion of the beam to see any acceleration.

During one week, we used plasma densities from 1.5×10^{17} to 3.5×10^{17} per cubic centimeter, so that the associated plasma wavelength shrank from 85 to 56 microns. Thus, we would expect that if we sent the same bunch into these different plasma densities, we would see differing amounts of acceleration.

We found the phase space for each event in the various data runs, and so could say what the approximate bunch length was for every shot. To do a direct comparison of one plasma density versus the others, we choose only a subset of the data events from various days which all matched to the same phase space from simulation.

A bunch profile that matched to data from each plasma density has a σ_z for the main peak of about 16 μm. Figure 5.12 shows the output from LiTrack. Linear theory

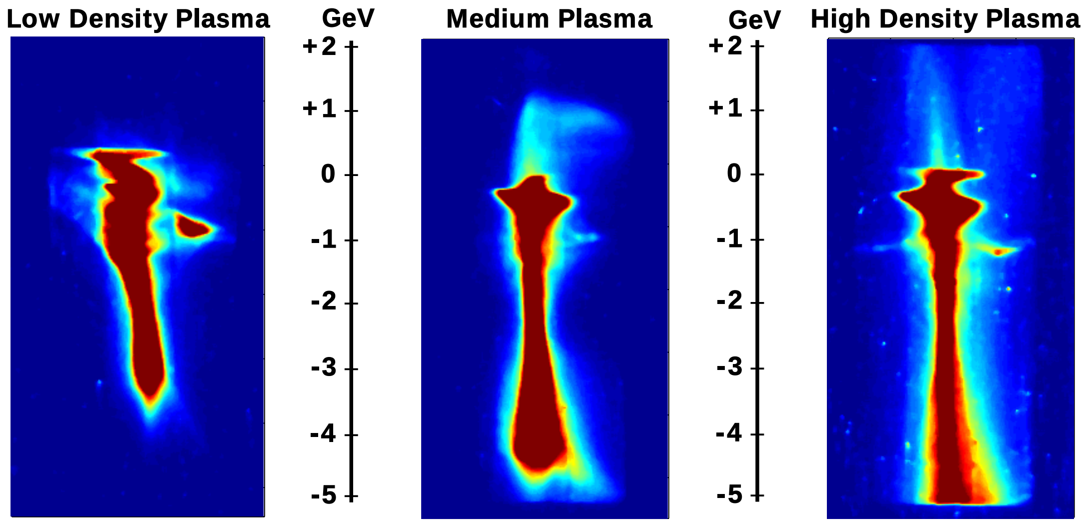

Figure 5.13: We compare the effect of different plasma densities on similar incoming beams. At left is 1.5×10^{17} cm^{-3} with no clearly accelerated particles. In the center, the plasma density is 2.5×10^{17} cm^{-3} and we clearly see an accelerated tail. On the right is the densest plasma: 3.5×10^{17} cm^{-3}, with many strongly accelerated particles. The approximate energy scale shown is relative to the highest incoming energies.

would predict the plasma density with maximum energy gain to be 2.2×10^{17} cm^{-3}, but we have non-Gaussian bunches with uneven wings. Also, the previously described phase delays in the formation of the wake make it less easy to predict the ideal plasma wavelength. We expect that we need a higher density for maximum acceleration than predicted by linear theory.

Having chosen events with the same incoming conditions, we then quantify the acceleration as a function of plasma density.

Figure 5.13 shows representative shots from the three densities of plasma. (Note that the first image had a slightly higher centroid incoming energy. The accelerator varies somewhat from day to day.) The first image represents the effect of the 1.5×10^{17} cm^{-3} plasma, where there are no accelerated particles visible. The second image shows the acceleration after the higher density plasma of 2.5×10^{17} cm^{-3}, where we now clearly see a wisp of electrons which have been accelerated beyond the highest incoming energies. The third image, with plasma density of 3.5×10^{17} cm^{-3}, shows particles with energies all the way to the top of the image. Beyond that point, we cannot see the electrons accelerated by more than about 2 GeV, though it is clear that there must be some such electrons.

Using the same technique as in the previous section, we can calculate the gradient available for acceleration in each of these cases.

The distance behind the peak current of the bunch that the wake lies changes from 75 μm to 50 μm between these cases, but for each of them, the particles that experience that maximum wake start out lower than the head particles by almost 4% = 1.1 GeV. For the case with 1.5×10^{17} per cc, we do not see any accelerated electrons, so the gradient is ≤ 11 GeV/m.

With the higher plasma density of 2.5×10^{17} cm^{-3}, we see acceleration of about 1.2 GeV above the highest incoming energy for a maximum gradient experienced by a particle in this bunch of the order of 23 GeV/m.

At the highest density of 3.5×10^{17} cm^{-3}, we now see acceleration of a full 2 GeV, so as with the case in the previous section, we see a gradient ≥ 31 GeV, limited by our spectrometer range. It appears that the maximum acceleration for our 16 μm bunch would actually be achieved in a yet higher density plasma, but 3.5×10^{17} cm^{-3} is the maximum density in our oven.

This analysis is direct and gives information about the absolute maximum acceleration, but it suffers from the significant problem that we do not know the exact plasma length very well, and thus cannot quote a gradient to very high accuracy. With multi cm "wings" on either side, it is difficult to separate out the effect just from the highest density region in the center. The next section discusses a more involved technique which allows better investigation of the gradient as well as information about transverse deflections such as hosing.

5.3 Acceleration Analysis - Statistical

A more accurate way to understand acceleration is to change the plasma length and see how the acceleration is affected. Plotting these versus each other gives another way to determine the gradient. Because the wings where the Lithium density drops to zero have the same length whatever the central plasma length, any questions about

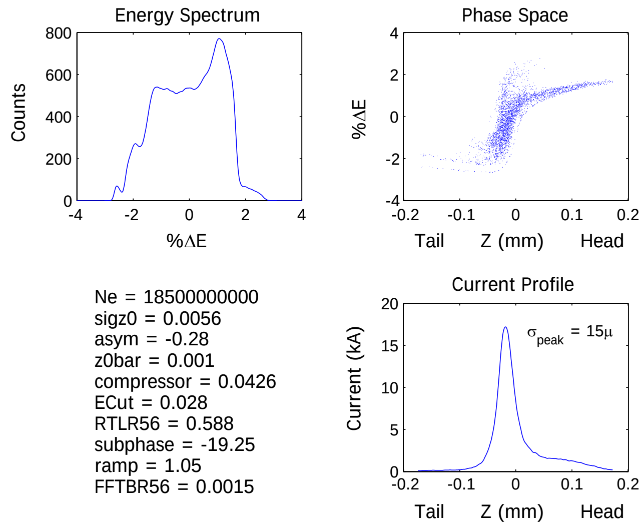

Figure 5.14: The phase space for the beams analyzed in this section.

the effect of the wings can be removed as common mode noise, and we only see the differential effect of adding or subtracting plasma length at the full desired density.

As discussed in Chapter 4, changing the oven heater power changes the plasma's length, but not its density. When reducing the power, changes in length require about 15 minutes to stabilize. As discussed in Chapter 4, for different plasma densities, the changes in length as a function of power are similar.

At each density, 1.5, 2.5 and 3.5×10^{17} cm^{-3}, we changed the oven length by several centimeters in the course of about two hours. The most rapid changes come when cooling down, so we started with the highest power. In reducing the oven length by about 4 cm, we typically took 8 to 12 evenly spaced data runs of 100 shots each.

To make direct comparisons of the acceleration, we need to have similar incoming electron beams from each data run and for each of the plasma densities explored. We found that electron beams with the phase space as shown in Figure 5.14 were present in all of our data runs to be discussed below, and therefore provided a good basis for statistical analysis. Nonetheless, due to the natural variations of the main accelerator, some runs have only 1 or 2 matching shots, some have more than 40.

With the various data events all having the same incoming conditions, we seek some metric for the amount of acceleration experienced by the beam in each event.

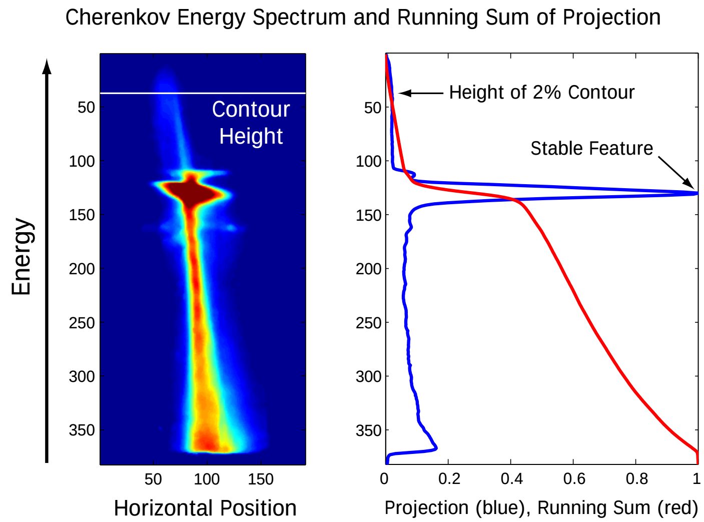

Figure 5.15: On left is a Cherenkov spectrometer image from 13 July 2004. On right is the projected energy spectrum in blue scaled so that the maximum is unity for plotting. In red is the running sum of the projection which is used to calculate the height of features such as the 2% contour labelled at top.

One way to quantify the acceleration that we see is simply to add up the total charge which has higher energy after the plasma than any particle had coming in. A more useful measure is to calculate contours of the beam in energy.

When dispersed at the Cherenkov screen, the vertical position of the electrons is dominated by their energy. If we calculate a running sum of the beam starting at the highest energy, we can calculate contours above which 5% of the beam's charge lies, above which 10% or 50% lies, and so on. Because a relatively small fraction of the beam is accelerated under even the best conditions, we focus on the contours in the range of about a half to several percent. Figure 5.15 shows a sample image from the Cherenkov screen and the calculation of the running sum which gives these contours.

As the incoming beam energy fluctuates slightly from shot to shot and between days, the relevant metric for the height of our 2% contour is not its absolute height, but its height above some stable feature in the beam itself. Thanks to the "nose" on our various beams before ionization takes place, we have such an unchanging feature. We calculate the height of the contour relative to this high energy spike, which is unaffected by the plasma, such as we have previously seen in Figure 5.8.

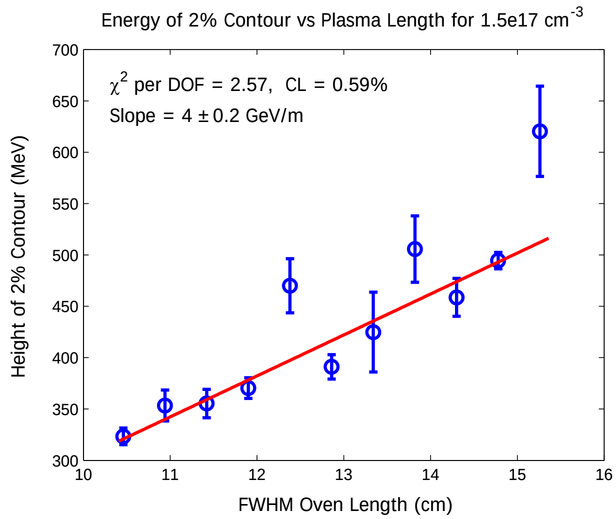

Figure 5.16: Plot of change in acceleration for the 2% contour vs oven length for plasma density 1.5×10^{17} cm^{-3} with linear fit.

5.3.1 Lowest Plasma Density

We use the above technique to analyze data taken on 11 July 2004 with the plasma set to a density of 1.5×10^{17} cm^{-3} to see what gradient had been achieved. By varying the oven power between the 390 and 460 Watts, we were able to change the oven length by about 5 cm in length. We took one set of 100 shots at each setting of the power, where the settings were 7.5 Watts different from one run to the next.

To have a consistent basis for comparing various events, we use only the shots from each run that match to the phase space shown in Figure 5.14. The number of successful matches varies, with between 2 and 17 matches for each of the runs here.

Having selected the subset of shots from each run, we calculate the height of various contours. As relatively few particles are accelerated, we calculate the gradient for only the 0.5% through 4% contours. For each shot, we calculate the height of the various contours above the stable feature afforded by the beam's "nose." We then plot the height of each contour versus the length of the plasma to see the gradient. An example of such a plot is shown in Figure 5.16 for the 2% contour.

The linear fit does not match the data all that well; the confidence level is quite low indicating that the linear fit is not capturing all of the physics or that the error

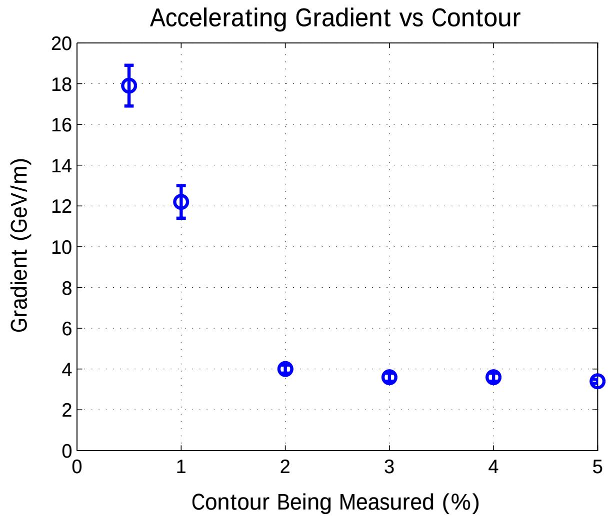

Figure 5.17: Plot of our accelerating gradient vs. the contour at which we choose to measure it for plasma density of 1.5×10^{17} cm^{-3}. Maximum achieved gradient is extrapolated to be about 20 GeV/m.

bars are incorrectly small. A discussion of the statistical techniques used to provide confidence levels can be found in Appendix A.

Nonetheless, by eye one can see that the gradient is something of the order of 4 GeV/m as given by the linear fit. We can create such plots for the various contours and it is then useful to plot the measured gradient versus the percent contour being studied. That way, we can extrapolate toward 0 to find the maximum gradient produced in our plasma. Such a plot is shown in Figure 5.17.

% Contour	**Gradient [GeV/m]**	**C.L.**
0.5	17.9 ± 1	0%
1	12.2 ± 0.8	0%
2	4.0 ± 0.2	0.6%
3	3.6 ± 0.2	0%
4	3.6 ± 0.2	0%
5	3.4 ± 0.1	0%

Table 5.1: Various parameters of the gradient for the lowest density plasma of 1.5×10^{17} cm^{-3}

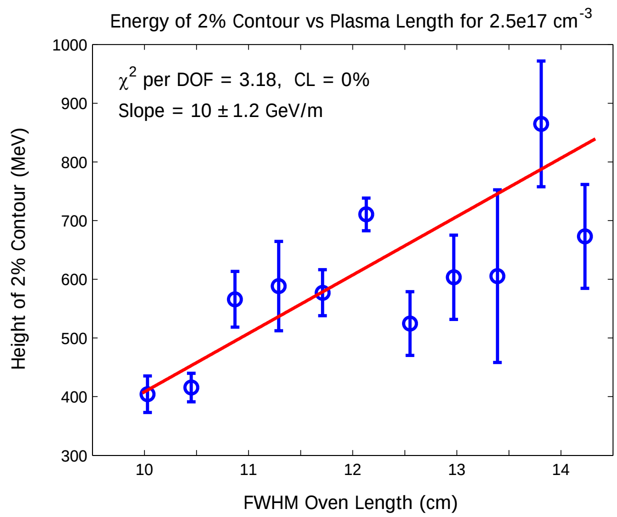

Figure 5.18: Plot of change in acceleration of the 2% contour vs oven length for the intermediate plasma density 2.5×10^{17} cm^{-3} with fit to a line.

We only include plasma lengths greater than 9 or 10 cm, where there are always visibly accelerated particles. Because the gradient is calculated for *changes* in plasma length, we remove questions about the size of the "wings" in the Lithium distribution, and no longer care exactly what the tail particle incoming energies had been.

We reiterate that the errors quoted on the gradients are unrealistically small, as the confidence levels for each linear fit are very low. Still, we can see that the gradient grows as we look at smaller and smaller groups of particles near the peak of the observed acceleration. Extrapolating to 0, we can estimate that the peak achieved gradient for our plasma of 1.5×10^{17} cm^{-3} was of the order of 20 GeV/m.

5.3.2 Intermediate Plasma Density

We perform a similar analysis with the 2.5×10^{17} cm^{-3} plasma and get interesting results. If we measure the acceleration of our 2% contour with beams that have the same longitudinal profile as mentioned above, we see the results in Figure 5.18.

The gradient determined by the fit is stronger than for the lower density plasma, as we would expect. Again, the line fit is definitely not capturing all of the physics. Unlike the previous case, however, the positions of the points are somewhat suggestive.

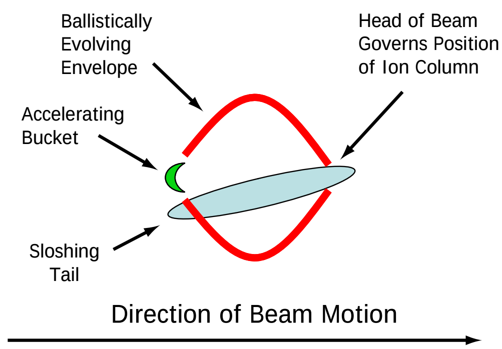

Figure 5.19: Diagram of the transverse size of the accelerating bucket. Particles that start off to one side will oscillate transversely at the betatron frequency, passing through the region of strong acceleration *twice* per oscillation.

We recall that we have seen transverse asymmetries in the beams coming into our plasma. If those offsets are in the tailmost particles, they will oscillate back and forth in the focusing fields of the ion column created by the beam blow-out. The strongest accelerating fields are located right on axis behind the head of the bunch. Although offset tail particles lie at the right longitudinal position, they will be in the right transverse position to experience the peak acceleration only part of the time, seeing an acceleration that oscillates in strength. The schematic of this effect is presented in Figure 5.19.

In each betatron oscillation, they will enter and then leave the accelerating region twice. Thus, not only will the average acceleration be lower than it would otherwise be, but we will see modulation of the effect at twice the betatron frequency, providing a useful hallmark to distinguish this particular effect from other possible explanations.

We can see on the upstream OTR screen that our beams do, in fact, come into the plasma with transverse asymmetries. Although we cannot distinguish the tail from the head in these images, offset tails are a reasonable explanation. We consistently notice that there are asymmetries to the same side. Figure 5.20 shows an example of the beam with a clear tail to one side.

Interestingly, the transverse profile of the bunch clearly does not match to a Gaussian. Rather, to quantify the asymmetry, we fit the bunch to an asymmetric Lorentzian, which is a very good match. Lorentzians are normally associated with

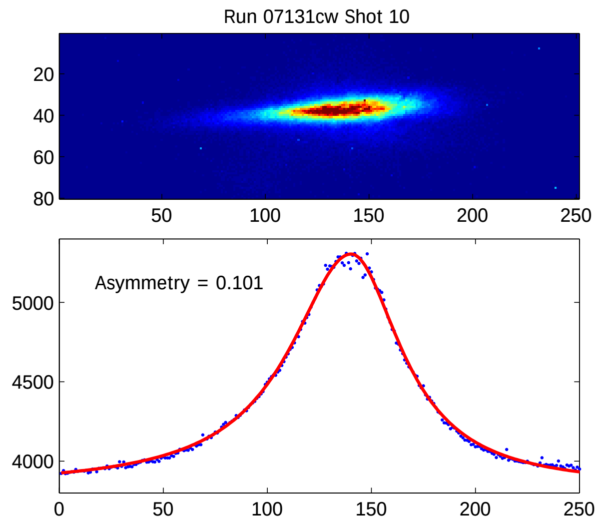

Figure 5.20: Sample image from the upstream OTR foil showing the small tail off to the left. The projection of the image onto the horizontal axis gives a curve which is well fit by an asymmetric Lorentzian.

resonance phenomena, so perhaps the periodic transverse focusing in the main accelerator optics participates in creating this shape. Our equation is:

$$L = C + \frac{W}{1 + \frac{2(1+sgn(x)A)^2}{x^2\sigma_L^2}} \tag{5.3}$$

where C is the pedestal on which our Lorentzian sits, W is the amplitude, A is the asymmetry factor and σ_L is the width parameter of the curve. This is by direct analogy with our definition of the asymmetric Gaussian (3.1).

With a transversely asymmetric beam coming in, it is natural to investigate what happens transversely to the tails of the beam *after* the plasma. As the Cherenkov spectrometer is imaging in x, we can see horizontal deflections of the beam, just as we saw them in the upstream OTR images.

The technique is to take a full Cherenkov image and find the height of whichever contour is of interest, as usual. We then take a horizontal lineout of the Cherenkov image at that height and fit the curve to a Gaussian. The transverse position that we quote is simply the mean value from the Gaussian fit. This is illustrated in Figure

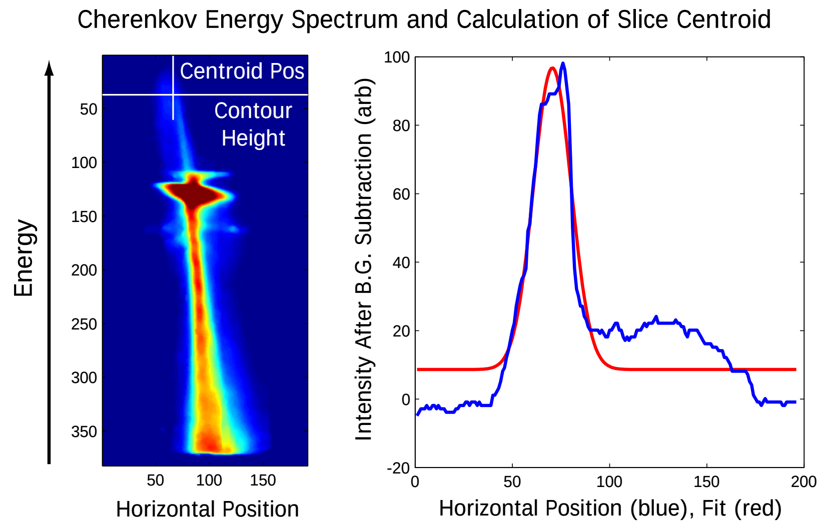

Figure 5.21: By taking a slice from the full image, we can then fit it to a Gaussian to get an estimate of where the centroid lies. In this case, the mean position in the Gaussian fit is not in the same place as the raw centroid would be, but this technique is more robust against noise such as X-ray hits.

5.21. In order to convert the observed transverse positions to something meaningful, we use the magnification of our electron optics to calculate the actual size of the transverse deflections at the plasma exit.

Using this technique, we can find the average horizontal position of the slices at the height of various contours for each plasma length. If tails are oscillating in the focusing fields of the plasma, we should see sinusoidal oscillations at the betatron frequency. We can then plot our tail transverse positions vs. plasma length, as in Figure 5.22.

The amplitude of these oscillations is of the order of the transverse size of the beam, σ_r. The wavelength for the oscillation returned by the fitting routine is a bit larger than the expected 2.23 cm, but we already know that our understanding of plasma length as a function of heater power is only accurate to about 10%, which could account for some of the discrepancy between these two values. Combined with the uncertainties in the fit parameters, our measured value is close enough to the expected value to believe that this really does represent betatron oscillations of an offset tail.

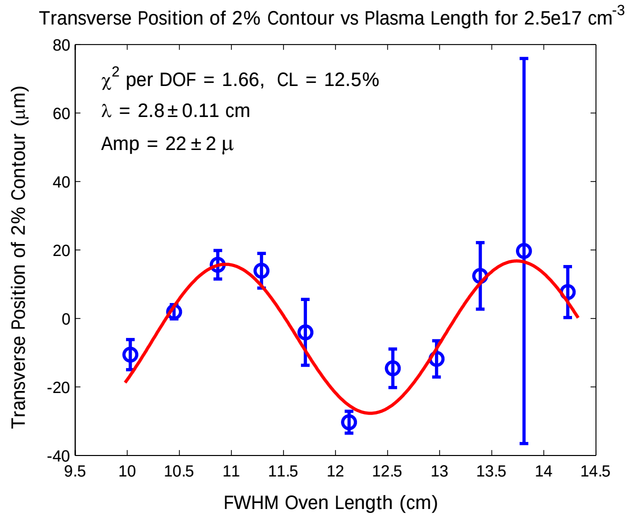

Figure 5.22: Horizontal position of the 2% contour versus plasma length and find that its oscillations could be consistent with simple betatron focusing of an initial offset in the accelerated tail.

And the observed 1.26 ± 0.05 cm period of oscillation in the energy gain is roughly half this value, as we would expect.

With all of this information, we fit the acceleration data not to a line, but we superimpose a sinusoid on top of the expected linear increase in energy with plasma length. If we fit the points again, but to a curve which is the sum of a line and a sinusoid, we get better agreement with the data. This is plotted in Figure 5.23.

There are now more parameters to the fit, and we show the most important ones on each graph, namely the wavelength of the oscillation and its amplitude. The wavelength we see for this particular fit is 1.26 ± 0.05 cm. This number is close to half of the betatron wavelength for our 28.5 GeV beam in a plasma of this density: 1.12 cm, and is close to half the value obtained by fitting to the transverse oscillations. This is exactly the relationship we expect if the transverse oscillations affect the acceleration.

Furthermore, if the oscillations in acceleration that we see are caused by this transverse motion of the tail particles, then we expect that the oscillation amplitude will be reduced in tandem with the decreasing overall acceleration as we investigate ever more inclusive contour lines. This is exactly what we see at right in Figure 5.24.

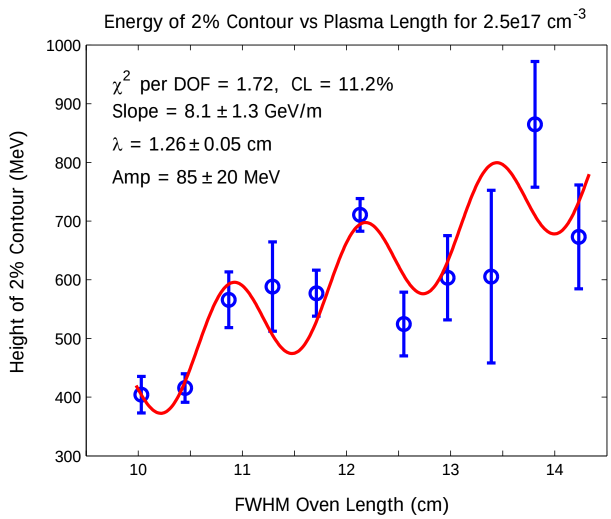

Figure 5.23: Plot of change in acceleration of the 2% contour vs oven length for plasma density 2.5×10^{17} cm^{-3} with fit to a line plus sinusoid.

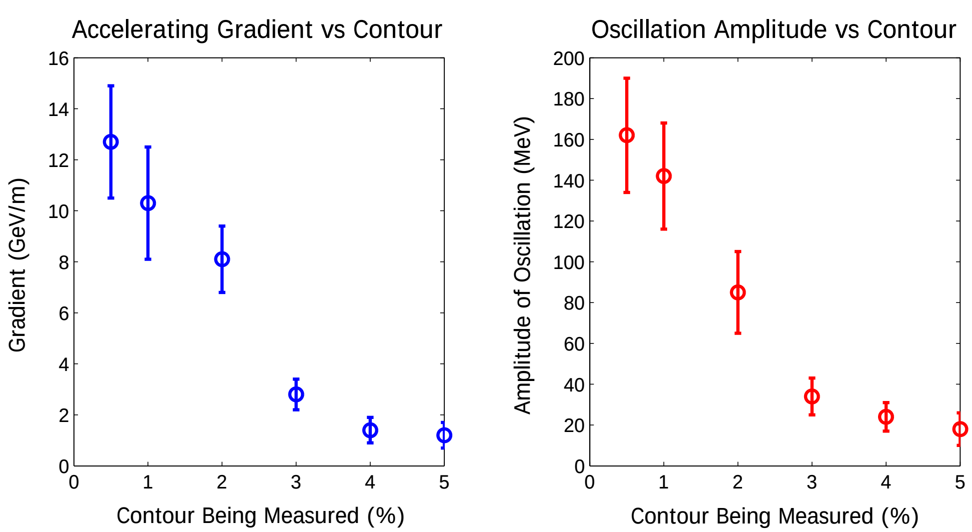

Figure 5.24: Plot of our accelerating gradient vs. the contour at which we choose to measure it for plasma density of 2.5×10^{17} cm^{-3}. Maximum achieved gradient is extrapolated to be about 14 ± 3 GeV/m.

% Contour	Gradient [GeV/m]	λ [cm]	Amplitude [MeV]	C.L.
0.5	12.7 ± 2.2	1.67 ± 0.08	162 ± 28	1%
1	10.3 ± 2.2	1.27 ± 0.05	142 ± 26	1%
2	8.1 ± 1.3	1.26 ± 0.05	85 ± 20	11%
3	2.8 ± 0.6	1.27 ± 0.06	34 ± 9	24%
4	1.4 ± 0.5	1.37 ± 0.11	24 ± 7	22%
5	1.2 ± 0.5	1.63 ± 0.14	18 ± 8	66%

Table 5.2: Various parameters of the gradient and observed oscillations in that gradient for the medium density plasma of 2.5×10^{17} cm^{-3}

If we look at the wavelength of the oscillation in acceleration amplitude from Table 5.2, we see that there is a group of similar values for the contours 1% through 4%, and the average value is about 1.3 cm. This is close to the expected period of oscillation for transverse effects, and gives further evidence that these oscillations cause the uneven acceleration as a function of oven length.

We see also that the oscillation amplitude changes in much the same way as the overall accelerating gradient, which again implies that this is caused by the above mechanism. We note that the overall gradient is actually lower than for the case of the lower density above. This is to be expected if the tail spends a significant amount of time outside of the main accelerating bucket. So these three lines of inference give confidence that we understand the effects on acceleration. As a side effect of this investigation, we can say that we do not see evidence of hosing growth, but transverse deflections can still be a problem for the overall achievable gradient.

5.3.3 Highest Density Plasma

Lastly, we investigate the highest plasma density achieved, 3.5×10^{17} cm^{-3}. As observed in § 5.2.2, we should again have the highest gradient.

If we again plot the transverse deflections for the 2% contour as a function of plasma density, we obtain Figure 5.25. The transverse oscillations are somewhat greater in amplitude than those observed with the intermediate plasma density. The machine was likely in a slightly different state the day this data was taken.

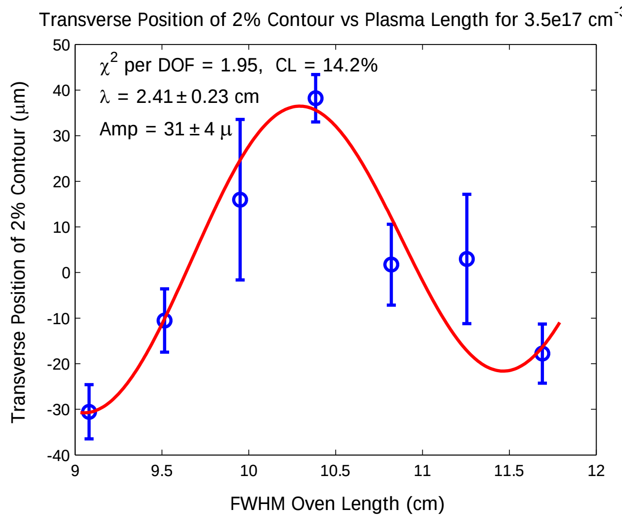

Figure 5.25: We plot the horizontal position of the 2% contour versus plasma length and find that its oscillations could be consistent with simple betatron focusing of an initial offset in the accelerated tail.

We expect that there will be oscillations in the energy, and that the amplitude of those oscillations will be greater than for the medium plasma density. We therefore use the line plus sinusoid fitting routine again. Figure 5.26 shows the result of finding the gradient for the 2% contour of shots into this high density plasma.

We see that the wavelength of oscillations in the transverse dimension is very close to twice the wavelength of the energy modulations that we observe, as expected with our understanding of what is happening inside the plasma. At 1.22 cm and 2.41 cm, both wavelengths are greater than expected theoretically, at 0.94 cm and 1.88 cm, respectively. This is probably because our understanding of the plasma length versus heater power is not perfect. If we refer to Table 4.1, we see that the value of the slope for 3.5×10^{17} cm^{-3} looks anomalously large. Incorrectly large values cause the wavelength to appear longer than it is (and reduce the measured gradient).

We plot the gradient and oscillation amplitude achieved vs. each contour investigated. This is presented for our high density plasma in Figure 5.27, where the estimated maximum gradient for the 0% contour would be at about 22 ± 5 GeV/m. If we assume that the differences we observe between the measured betatron wavelength and the theoretical value are because we have a poor calibration of the oven

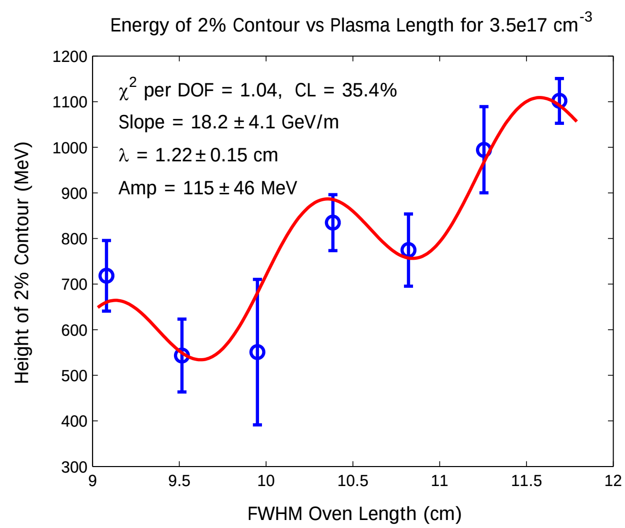

Figure 5.26: Plot of change in acceleration of the 2% contour vs oven length for plasma density 3.5×10^{17} cm^{-3} with fit to a line plus sinusoid.

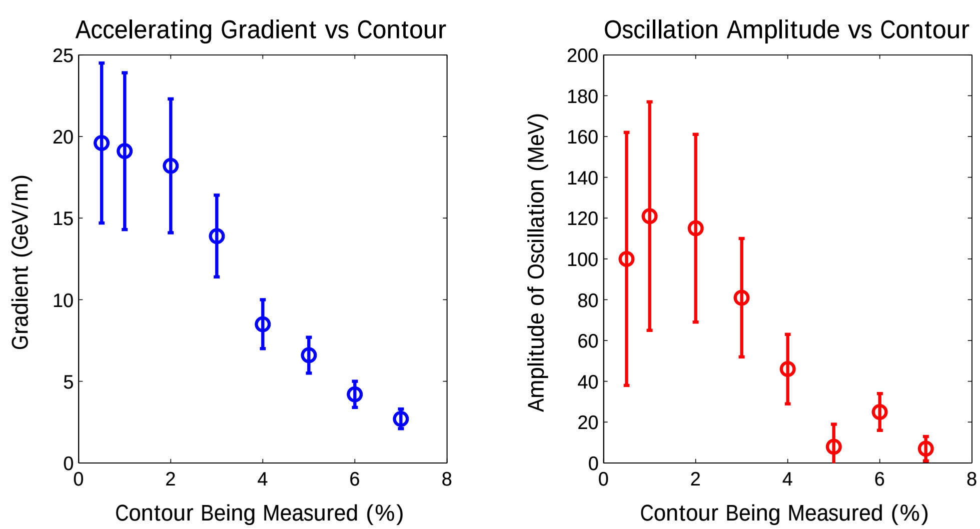

Figure 5.27: Plot of our accelerating gradient vs. the contour at which we choose to measure it for plasma density of 3.5×10^{17} cm^{-3}. Maximum achieved gradient is extrapolated to be about 22 ± 5 GeV/m.

% Contour	Gradient [GeV/m]	λ [cm]	Amplitude [MeV]	C.L.
0.5	19.6 ± 4.9	1.55 ± 0.22	100 ± 62	16%
1	19.1 ± 4.8	1.19 ± 0.14	121 ± 56	28%
2	18.2 ± 4.1	1.22 ± 0.15	115 ± 46	35%
3	13.9 ± 2.5	1.22 ± 0.13	81 ± 29	18%
4	8.5 ± 1.5	1.22 ± 0.13	46 ± 17	9%
5	6.6 ± 1.1	1.17 ± 0.47	8 ± 11	3%
6	4.2 ± 0.8	1.14 ± 0.07	25 ± 9	11%
7	2.7 ± 0.6	1.15 ± 0.24	7 ± 6	1.5%

Table 5.3: Various parameters of the gradient and observed oscillations in that gradient for the highest density plasma of 3.5×10^{17} cm^{-3}

length versus heater power, we can scale the gradient by the ratio of the observed betatron wavelength to that which we expect at a density of 3.5×10^{17} cm^{-3}. Doing so, we obtain a gradient of 28.5 ± 6.5 GeV/m, which is in good agreement with the more direct analysis presented in § 5.2.2.

Despite the increase in transverse oscillation amplitude compared to that for the intermediate plasma density, we see an increase in accelerating gradient compared to the case of 2.5×10^{17} cm^{-3} above.

5.3.4 Overall Results and Conclusions

We have demonstrated that energy spectrum measurements, coupled with detailed simulations of beams in the SLAC main accelerator, do, in fact, give an understanding of the longitudinal profile. The matching worked well, but should be even more reliable if the sources of blurring in the spectrum, primarily coming from the electron beam's significant β_x, were reduced. Nonetheless, this understanding then allows several investigations of the beam-plasma interaction.

Knowing the incoming energy of the particles which are accelerated gives a better understanding of their overall energy gain, and, therefore, the gradient. We can also use the energy spectra more simply to select a subset of nearly identical incoming bunches in order to compare the effect of different plasma densities and lengths on those bunches.

The most important benefit of this second technique is that we can measure the achieved gradient in a manner which does not depend on the regions in which the plasma density is changing. We have seen that the strongest acceleration is achieved when we use the highest density plasma available, although simple linear theory would predict that best acceleration for our bunches of about $\sigma_z \approx 18\,\mu$m should be with a plasma density of 2.5×10^{17} cm^{-3}. This indicates that continued progress in making denser plasmas could be more beneficial than initially thought.

In the course of these investigations, it became apparent that transverse oscillations of the tailmost particles are responsible for a decrease in accelerating gradient as well as oscillating energy gain to the particles. Further investigation is warranted to understand more fully these effects.

Due to the nature of the data available as of July 2004, where we could change the oven length only modestly, more accurate conclusions about gradients and hosing must await the case when the plasma length can be changed by larger amounts, and those experiments are underway as of the second half of 2005.

Chapter 6

Conclusions

With current technologies for particle acceleration nearing their theoretical maximum gradients, new accelerating techniques are necessary. Plasma acceleration with lasers has already demonstrated more than a thousandfold increase in gradient, but has long suffered from difficulties in propagating the laser over a significant distance.

Beam based plasma wakefield accelerators get around such difficulties, but can only achieve the same high gradients with very short electron bunches. Having created the necessary bunches of RMS length 40 fsec, we lose the ability to understand their longitudinal phase space directly, as no currently available technique can resolve such bunches in time. To solve this problem, we used an indirect method to understand the longitudinal phase space of our electron bunches. This indirect technique has proven crucial to understanding a variety of effects that happen with short electron beams driving a plasma wakefield.

As the technique is new, we have verified its accuracy by comparing with a variety of other ways of understanding the electron beams at SLAC.

We have seen several different ways to apply the phase space information to understanding acceleration, where we have applied the technique to measure incredibly high gradients of nearly 30 GeV/m over macroscopic distances. We found that the gradient does, in fact, increase with greater plasma density, though the gradient is weakened when transverse deflections become significant compared to the transverse size of the beam.

Understanding the phase space of short beams is not only useful for E164, but has already been applied in designs for the Short Pulse Picosecond Source (SPPS) experiment at SLAC, where short bunches are used to create intense X-Ray radiation in a wiggler, as a test for the Linac Coherent Light Source (LCLS). Researchers at DESY have also inquired about the technique for measuring the energy spread of beams in a non-destructive way, and it may well find broad applicability in future machines.

Continued progress in accelerating gradient requires ever more sophisticated methods for understanding the short electron bunches that we use, and we have demonstrated one such in E164.

Appendix A

Linear and Nonlinear Fitting to Data

We discuss the useful and generally applicable case of linear least squares fitting both for background and to motivate the discussion of the main topic of interest: nonlinear least squares fitting.

The linear case is nice because it admits of an exact analytic solution which can be used to check numerical methods we seek to employ for the nonlinear case.

A.1 Linear Least Squares

The term "linear least squares fitting" does not require that the function we are using for the fit be linear in the independent variable, e.g. x. Rather, the fit function must be linear in our *fit parameters*, which we will denote as α_i. To perform linear least squares fitting, we require that our function $F(x)$ have the form:

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \ldots + \alpha_m f_m(x) \tag{A.1}$$

In this case, the $f_i(x)$ can be anything we want, x^2, $\cos x$, $J_\nu(x)$, or whatever else, as long as each α_i multiplies a separate function of x.

A.1.1 Linear Fitting Example

Without proving any of the formulae, we outline the analytic method for finding the linear least squares fit to data with error bars by following the procedures and notation presented in Orear's monograph "Notes on Statistics for Physicists, Revised" [81].

The problem we often have is to fit experimental data with error bars to a theoretical curve. For the reader's reference, when that curve is just a line, the procedure of least squares fitting is sometimes referred to as "linear regression," a term coming from statistics in the social sciences and business world.

The best way to demonstrate the technique for finding the best fit to our data and for finding the errors in the fit parameters is through a concrete case. Orear's Example 6 poses the problem that we have four experimental data points which we wish to fit to a parabola. (This can also be referred to as "quadratic regression.") Namely:

$$\vec{x} = (-0.6, -0.2,\ 0.2,\ 0.6)$$
$$\vec{y} = (5,\ 3,\ 5,\ 8)$$
$$\vec{\sigma} = (2,\ 1,\ 1,\ 2)$$

where $\vec{x}$ is the vector of horizontal positions at which we took our data points, and is assumed to represent perfectly known values. $\vec{y}$ gives the experimentally measured values at the various x positions, and $\vec{\sigma}$ represents the quoted errors in the y values.

We seek to use a quadratic fit function $F(x)$, so we write it as

$$F(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 \quad \longrightarrow \quad f_1 = 1 \quad f_2 = x \quad f_3 = x^2 \tag{A.2}$$

With the three functions known, there is a straightforward method for calculating the fit parameters and their errors. The first step is to calculate what is generally known as the Hessian Matrix, **H**. Luckily, this is easy to do for linear fitting:

$$H_{ij} = \sum_a \frac{f_i(x_a) \cdot f_j(x_a)}{\sigma_a^2} \tag{A.3}$$

where a is an index referring to the various points. In this example, a runs from 1 to 4 such that we evaluate the functions f_i at each x. As i and j are interchangeable, $\mathbf{H}$ is clearly symmetric. For this example, as with all cases where $F(x)$ is composed of three subfunctions, we need only calculate six terms:

$$H_{11} = \sum_a \frac{1}{\sigma_a^2} \qquad H_{22} = \sum_a \frac{x_a^2}{\sigma_a^2} \qquad H_{33} = \sum_a \frac{x_a^4}{\sigma_a^2}$$

$$H_{12} = \sum_a \frac{x_a}{\sigma_a^2} \qquad H_{13} = \sum_a \frac{x_a^2}{\sigma_a^2} \qquad H_{23} = \sum_a \frac{x_a^3}{\sigma_a^2}$$

It is straightforward to verify that performing these sums will give:

$$\mathbf{H} = \begin{pmatrix} 2.5 & 0 & 0.26 \\ 0 & 0.26 & 0 \\ 0.26 & 0 & 0.068 \end{pmatrix} \tag{A.4}$$

The Hessian matrix is not directly useful until we take its inverse, $\mathbf{V}$. Doing this by hand is the worst combination of unpleasant and prone to calculational errors for all but the smallest of matrices. Thus, we normally let MATLAB perform the inversion with the delightfully simple command: `V = inv(H)`. We obtain:

$$\mathbf{V} = \begin{pmatrix} 0.664 & 0 & -2.54 \\ 0 & 3.847 & 0 \\ -2.54 & 0 & 24.418 \end{pmatrix} \tag{A.5}$$

Before we can calculate our fit parameters and their errors, we need to construct one final vector, $\vec{u}$. We have made no reference yet to the actual y values that we have measured, so it is reasonable that they enter into the calculation at this point:

$$u_i = \sum_a \frac{y_a \cdot f_i(x_a)}{\sigma_a^2} \quad \longrightarrow \quad \vec{u} = (11.25,\ 0.85,\ 1.49) \tag{A.6}$$

Armed with these various vectors and matrices, we can directly find the best fit

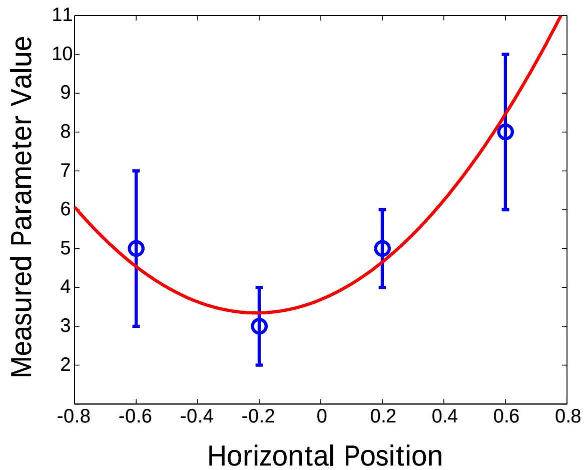

Figure A.1: Best fit plot. The blue points with error bars reflect the measured values and the red curve is the best fit function superimposed.

parameter values, often denoted by α_i^*:

$$\vec{\alpha}^* = \vec{u} \cdot \mathbf{V} \quad \longrightarrow \quad \alpha_1^* = 3.685 \quad \alpha_2^* = 3.270 \quad \alpha_3^* = 7.808 \tag{A.7}$$

To find the errors in the fit parameters, we just take the square root of the diagonal elements of our error matrix, $\mathbf{V}$:

$$\Delta\alpha_1 = 0.815 \quad \Delta\alpha_2 = 1.96 \quad \Delta\alpha_3 = 4.94 \tag{A.8}$$

If we are interested in the cross correlations between the fit parameters, they are simply the corresponding matrix elements of $\mathbf{V}$:

$$\overline{\Delta\alpha_1\Delta\alpha_2} = 0 \qquad \overline{\Delta\alpha_1\Delta\alpha_3} = -2.54 \qquad \overline{\Delta\alpha_2\Delta\alpha_3} = 0 \tag{A.9}$$

In summary, we have the best fit function which is shown in Figure A.1. We can quote only one or two significant digits, so write:

$$F(x) = (3.7 \pm 0.8) + (3.3 \pm 2.0)x + (7.8 \pm 4.9)x^2 \tag{A.10}$$

A.2 Interlude - Goodness of Fit

We have so far finessed the question of how good a job we have done in fitting our data to the assumed functional form. The traditional metric for how well the fit curve represents the data is called the χ^2, or "chi-square." This has the straightforward and intuitive definition that the differences between each data point and the fit function at that point are scaled by the known errors, then squared and summed:

$$\chi^2 = \sum_a \left[\frac{y_a - F(x_a)}{\sigma_a} \right]^2 \tag{A.11}$$

Generally more useful than the raw χ^2 is the normalized value, χ^2 per degree of freedom. As the reader recalls, the number of degrees of freedom, ν, is just the number of data points minus the number of free parameters in the fit. Thus χ^2/ν naturally builds in the concept that there must be more data points than degrees of freedom in the fit for us to be able to quote a meaningful result.

Depending on the number of degrees of freedom, we can determine how likely it is that we have chosen an appropriate function to fit our data. The intervals in Figure A.2 tell us this information. What the contours show is the percent chance that if we perform a subsequent identical experiment, we will get a χ^2/ν value greater than the one we have measured. The ideal case is where we are on the 50% contour, meaning that we have found the most likely function.

These "confidence level" curves can be calculated by using a generalization of the error function known as the Incomplete Gamma Function, $P(a, x)$ [82]. The following discussion is based on the treatment in Chapter 6, *Special Functions*. The Incomplete Gamma Function is calculated in much the same way as erf(x) except that the normalization is given by the Gamma Function. We recall that $\Gamma(a)$ is itself the generalization of the factorial function to non-integral a and is defined:

$$\Gamma(a) = \int_0^\infty e^{-t}\, t^{a-1}\, dt \tag{A.12}$$

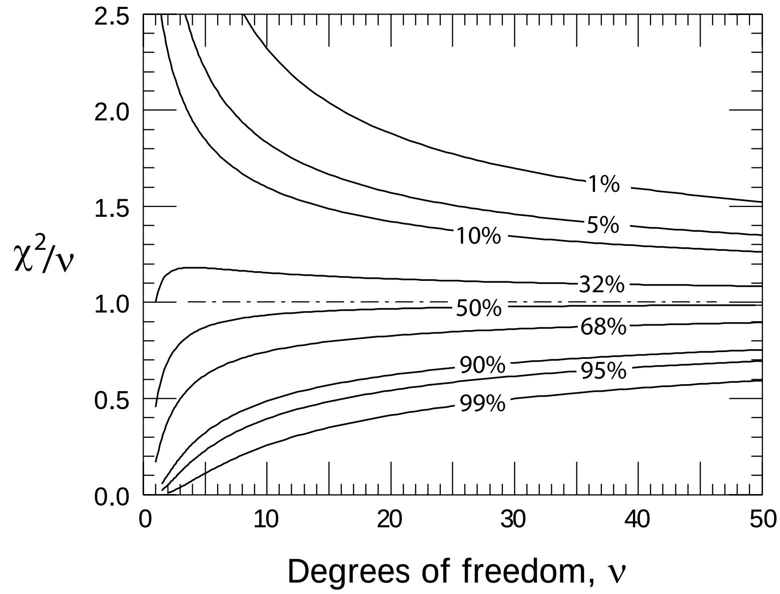

Figure A.2: Confidence interval curves for various χ^2/ν as a function of ν. In the example from § A.1.1, ν is just one, and the most trusted value for χ^2/ν would be about 0.5. Reproduced from [83].

With this normalization, we can define $P(a, x)$ to be:

$$P(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} \, t^{a-1} dt \tag{A.13}$$

We use this to find our Confidence Level, CL, by setting a to be $\nu/2$ and x to $\chi^2/2$ and by taking the complement of $P(a, x)$. So for a given raw χ^2 and number of degrees of freedom ν, our confidence level is given by:

$$\mathrm{CL} = 1 - P\left(\frac{\nu}{2}, \frac{\chi^2}{2}\right) \tag{A.14}$$

In MATLAB, we calculate $P(a, x)$ with `gammainc`, but must be careful because a and x are *reversed* in order relative to the standard mathematical notation. Thus, we might implement the above as `CL = 1 - gammainc(ChiSq/2,DOF/2)`.

In general, if we have a confidence level of only 0.1%, then our fit is not very meaningful. Either we are trying to fit to the wrong function, or we have underestimated the errors of each individual measurement. Conversely, if we are on the 99.9% contour, then the fit is "too good," and we have overestimated our errors or otherwise made a mistake. It is worth noting that when data is fudged to make it "better," a hallmark can be that the χ^2/ν is on a contour closely approaching the 100% confidence interval.

In the example from § A.1.1, the χ^2/ν is 0.35, placing us near the 60% contour, and indicating that the data are very much consistent with our assumed parabolic function. We note that being near the 50% contour does *not* in and of itself prove that we can exclude all other functional forms for the fit function.

We can only exclude other possible fit functions one by one. If the χ^2/ν associated with fitting to a specific function leads to a CL of 0.1%, then we say that that particular function does not match the data well. In this example, if we fit to a simple line, we obtain that χ^2/ν is 1.42, placing us at perhaps the 20% confidence interval. That is not a terrible result, so we cannot dismiss out of hand a linear fit to the data without real theoretical reasons to eschew this possibility.

It is nonetheless correct to say that the parabolic fit is a better match to the data. This difficulty in confidently choosing which is the correct fit function in this example is a direct result of the low number of degrees of freedom (1 or 2!), and adding even one more data point should enable a more confident discrimination between these two possible fit functions.

As we normally have a theory and therefore a hypothesis as to what functional form we would expect, the most robust result we can claim is simply whether or not we are confident that we have observed the functional form predicted in our hypothesis.

A.3 Theory Behind the Hessian Matrix

The most important part of determining the best fit and the errors on the fit parameters is the Hessian matrix, $\mathbf{H}$. We have seen *how* to construct this matrix, but we need to understand *why* we did what we did when we move to the case of nonlinear least squares fitting.

General background leading to the results I quote subsequently can be found in [82], specifically Chapter 14 on modelling of data. The following discussion is adapted primarily from §14.4, *Nonlinear Models*.

The general numerical approach to finding fit parameters is to give approximate values and then vary them repeatedly, calculating χ^2 for each combination of parameters. Minimizing χ^2 gives the best fit parameters as long as the search starts close enough to the global minimum. The Hessian matrix at any point can be thought of as the curvature matrix of the χ^2 merit function at our set of fit parameters, $\vec{\alpha}$. The curvature is simply one half times the second derivative matrix:

$$H_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial\alpha_i\partial\alpha_j} \tag{A.15}$$

At the risk of being repetitive, we again present the definition of χ^2 to show how its derivatives are calculated, and this time make explicit the dependence of F on the data and the fit parameters:

$$\chi^2(\vec{\alpha}) = \sum_a \left[\frac{y_a - F(x_a;\vec{\alpha})}{\sigma_a}\right]^2 \tag{A.16}$$

We take the gradient of χ^2 with respect to our fit parameters:

$$\frac{\partial\chi^2}{\partial\alpha_i} = -2\sum_a \frac{y_a - F(x_a;\vec{\alpha})}{\sigma_a^2}\,\frac{\partial F(x_a;\vec{\alpha})}{\partial\alpha_i} \qquad i = 1, 2, \ldots, m \tag{A.17}$$

Taking the second partial, we obtain:

$$\frac{\partial^2\chi^2}{\partial\alpha_i\partial\alpha_j} = 2\sum_a \frac{1}{\sigma_a^2}\left[\frac{\partial F(x_a;\vec{\alpha})}{\partial\alpha_i}\frac{\partial F(x_a;\vec{\alpha})}{\partial\alpha_j} - \left[y_a - F(x_a;\vec{\alpha})\right]\frac{\partial^2 F(x_a;\vec{\alpha})}{\partial\alpha_i\partial\alpha_j}\right] \tag{A.18}$$

As we can see, the components of $\mathbf{H}$ depend on both first and second partial derivatives, but the second derivatives are almost always ignored. Loosely speaking, this is for several reasons. First, the second derivatives are multiplied by the difference between the data and the fit function. An even modestly good fit makes this number relatively small. Second, the sign of these difference terms will randomly be positive or negative, causing them to tend to cancel one another. Third, including the second derivatives can make the fit even less reliable in the circumstance where the model is not an ideal fit or when there are random outlying points.

Recalling that our factor of one half in the definition of the curvature matrix cancels the 2 in front of our sum, we ignore the second derivative terms and give the approximated definition for the elements of **H**:

$$H_{ij} = \sum_a \frac{1}{\sigma_a^2} \left[\frac{\partial F(x_a; \vec{\alpha})}{\partial \alpha_i} \frac{\partial F(x_a; \vec{\alpha})}{\partial \alpha_j} \right] \tag{A.19}$$

It is immediately apparent that this is the general formula which gave us the method of constructing the Hessian matrix in our earlier example. We had required that F be a linear combination, $F = \sum \alpha_i f_i(x)$. Thus, the derivatives with respect to each α_i trivially return just the $f_i(x)$ such that the bracketed term in the above sum will always be $[f_i(x_a) \cdot f_j(x_a)]$ for linear least squares fitting.

Although nominally a second derivative matrix, our approximated Hessian is really constructed from combinations of first partial derivatives. Those partial derivatives are themselves the elements of a matrix known as the Jacobian, **J**. (Confusingly, the determinant of **J** is sometimes also called the Jacobian.) In terms related to the above, and recalling that a is the index of each data point, the definition of the elements of the Jacobian matrix is:

$$J_{ai} = \frac{\partial F(x_a; \vec{\alpha})}{\partial \alpha_i} \tag{A.20}$$

where i spans the number of fit parameters we are using and the derivative is evaluated at the data points x_a. Thus, in this formulation, the matrix must always be taller than it is wide, corresponding to a nonzero number of degrees of freedom.

Explicitly, we can now define the Hessian as:

$$H_{ij} = \sum_a \frac{1}{\sigma_a^2} \left[J_{ai} \cdot J_{aj} \right] \tag{A.21}$$

To be more concrete, the Jacobian for our earlier example uses the partials with respect to the various α_i:

$$\frac{\partial F(x;\vec{\alpha})}{\partial \alpha_1} = 1 \qquad \frac{\partial F(x;\vec{\alpha})}{\partial \alpha_2} = x \qquad \frac{\partial F(x;\vec{\alpha})}{\partial \alpha_3} = x^2 \tag{A.22}$$

Evaluating these partials our 4 data points, we obtain the simple matrix:

$$\mathbf{J} = \begin{pmatrix} 1 & -0.6 & 0.36 \\ 1 & -0.2 & 0.04 \\ 1 & 0.2 & 0.04 \\ 1 & 0.6 & 0.36 \end{pmatrix} \tag{A.23}$$

The first element of the Hessian matrix is the sum of the first column after we square each element and divide by σ_a^2. Similarly, H_{23} is given by the sum of the term by term products of column 2 and column 3, again divided by the corresponding σ_a^2.

In the linear case, this really just gives theoretical background for the method already discussed at the beginning of this appendix. Association with the concept of the Jacobian is useful because that is what MATLAB actually produces in doing numerical fits.

A.4 Nonlinear Least Squares Fitting

We can still sometimes calculate the Hessian even if $F(x_a;\vec{\alpha})$ is not linear with respect to the various α_i, and therefore the partial derivatives are difficult to find analytically. Better yet, we can let computers give a numerical approximation to the derivatives and never do any algebra at all.

As discussed in Chapter 5, our acceleration data from the plasma appears to have the functional form of a line plus a sinusoid. Fitting to this takes us away from linear theory, because several fit parameters appear inside the sine function, and we no longer have a linear combination. For readability, instead of using the notation α_i

for the fit parameters, we just use the obvious symbols in the fit function:

$$F(x) = (Ax + B) + C \sin\left(\frac{2\pi}{\lambda} x + \phi\right) \tag{A.24}$$

There is no straightforward way to solve this analytically, though one could perform a Taylor expansion on the functions with redefined fit parameters. This is, in general, a nightmare, so we let computers do what they do best: perform lots of calculations to give a numerical answer which will be close to the correct one.

MATLAB is widely used, and has the additional advantage that its syntax is generally straightforward and resembles pseudocode. So for this example, I will quote actual MATLAB code to provide illustration of the technique for nonlinear fitting including errors.

A good way to find the best fit numerically is to minimize the χ^2 directly. That is implicitly what is done with more user friendly routines in MATLAB, such as `lsqcurvefit`. Here, we need to use a less convenient, but more powerful minimization routine. In order to get all of the parameters needed for us to be able to quote errors in the fit parameters, we do some of the work explicitly ourselves.

To find the best fit parameters, we use the function `lsqnonlin`. Because we are doing a very specific task, it is worth noting that we cannot (as of MATLAB 7.0) use the closely related `lsqcurvefit`, because it passes data in a slightly different way.

The routine `lsqnonlin` has a variety of bells and whistles. Its basic purpose, however, is to minimize some function that the user provides with respect to various parameters whose seed values are also provided by the user.

We have to define the χ^2 function in a specific way so that MATLAB will calculate it correctly. Because `lsqnonlin` squares and sums our function for us, we calculate the individual terms at each data point which combine to make χ^2, calling them the scaled differences $\vec{D}$:

$$D_a = \frac{y_a - F(x_a; \vec{\alpha})}{\sigma_a} \tag{A.25}$$

If we compare the D_a with the $F(x_a; \vec{\alpha})$, we see that the only differences are that we have *already* divided by σ_a and that we will have an overall minus sign in the derivatives with respect to the α_i, as F is subtracted from y by convention. Thus, the terms of the Jacobian of $\vec{D}$ will be scaled by the $1/\sigma_a$ and will be the negative of the Jacobian of $F(x_a; \vec{\alpha})$. In constructing the Hessian, we must multiply a pair of Jacobian terms together, so the two minus signs will always cancel.

Our earlier definition of the Hessian of F (A.19) had a separate factor of $1/\sigma_a^2$ for each term, but that is now included in the Jacobian of $\vec{D}$. Having redefined the function we wish to minimize thus makes the Hessian matrix elements trivial to compute:

$$H_{ij} = \sum_a J_{ai} \cdot J_{aj} \tag{A.26}$$

In the previously mentioned case where we seek to match to a line plus sinusoid, we first have to write the function which will take the difference between each data point and the fit curve and then scale these differences by the errors. We require the vector of parameter starting conditions `StartParam` and the structure array `Data` which has the vectors $\vec{x}$, $\vec{y}$ and $\vec{\sigma}$ as its three fields. The function is:

```
\% Fit to a line plus a sinusoid with five parameters
function Differences =  LinePlusSinusoid(StartParam,Data)

\% Extract the data to be fitted
Horizontal = Data.Horizontal;
Vertical   = Data.Vertical;
Error      = Data.Error;

\% The initial conditions for the fit
Slope      = StartParam(1);
Intercept  = StartParam(2);
Amplitude  = StartParam(3);
Wavelength = StartParam(4);
Phase      = StartParam(5);
```

```
\% Calculate the function at the data points
Curve = Slope*Horizontal + Intercept + ...
   Amplitude*sin(2*pi*Horizontal/Wavelength + Phase);

\% Take the differences scaled to the errors, lsqnonlin  squares and
\% sums automatically in creating the Chi Squared.
Differences = (Vertical - Curve)./Error;
```

We now write a short script to call `LinePlusSinusoid` inside `lsqnonlin`. We assume that we have defined the vector `StartParam` and the structure array `Data`. We could have created a simple array with x, y, and σ, but this is more mnemonic.

In the input list of parameters, we just give empty sets for several of the parameters such as upper bounds that we do not wish to use as constraints.

To get parameters such as χ^2, we must give a sequence of variable names inside the brackets on the left of the equals sign into which `lsqnonlin` can write the results. These parameters must be given in order, so we still include variables for items we do not care about. For ease of comparison with MATLAB help, in the following code, we use the same names for output parameters as appear there and indicate which ones we care about by capitalizing them. `Resnorm` gives the raw χ^2 and `Jacobian` is the numerically calculated $\mathbf{J}$. It is now easy to calculate $\mathbf{H}$, invert it to make $\mathbf{V}$, and take the square root of the diagonal elements to find the errors on the previously returned fit parameters:

```
\%******* Fit to Line and Sinusoid with lsqnonlin *******
StartParam = [Slope0 Intercept0 Amplitude0 Wavelength0  Phase0];
[FitParams,Resnorm,residual,exitflag,output,lambda,Jacobian] = ...
    lsqnonlin(@LinePlusSinusoid, StartParam,[],[],[],DataToFit);

DOF = length(DataToFit.Horizontal) - length(StartParam)
ChiSqPerDOF = Resnorm/DOF
CL = 1 - gammainc(Resnorm/2 , DOF/2)
```

```
Slope          = FitParams(1);
Intercept      = FitParams(2);
Amplitude      = FitParams(3);
Wavelength     = FitParams(4);
Phase          = FitParams(5);

\%--------- Error Analysis on Fit Parameters -------------
for i = 1 : length(FitParams)
    for j = 1 : length(FitParams)
        H(i,j) = sum(Jacobian(:,i).*Jacobian(:,j));
    end
end
V = inv(H);
for k = 1 : length(FitParams)
    Deltas(k) = sqrt(V(k,k));
end
\%*********************************************************
```

With `FitParams` and `Deltas` in hand, we have the fit with its errors and are done with the statistics. It is now up to us to interpret the results.

To give an example of applying this code, we apply it to data discussed in Chapter 5 for our highest density plasma with 3.5×10^{17}. This code and a similar variant that fits simply to a line allows us to discriminate whether the line plus sinusoid or the simple line is superior. Our data vectors corresponding to `DataToFit.Horizontal`, `DataToFit.Vertical` and `DataToFit.Error` are:

$$\vec{x} = (0.37, 0.80, 1.22, 1.65, 2.08, 2.51, 2.93, 3.36)$$
$$\vec{y} = (780, 700, 1320, 1160, 1380, 1720, 1660, 1700)$$
$$\vec{\sigma} = (91, 64, 153, 124, 228, 67, 316, 60)$$

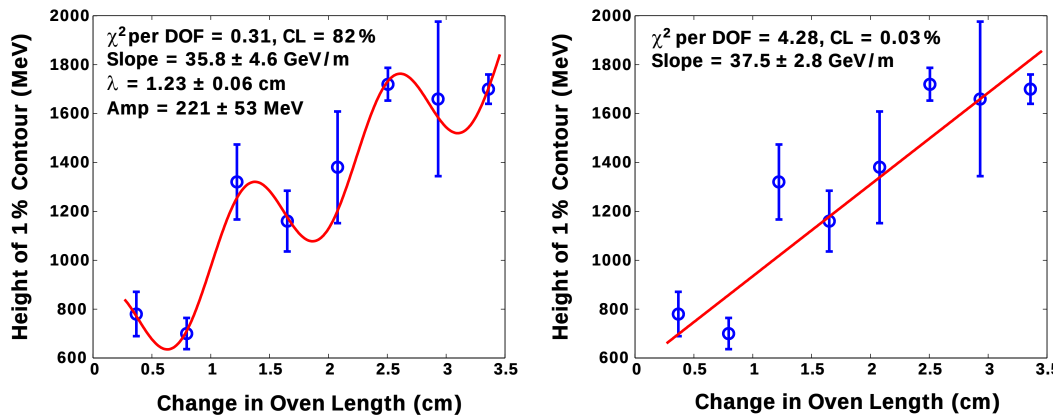

Figure A.3: Energy gain as we lengthen the plasma cell. Data fit much better to a line plus a sinusoid than a simple line.

To seed the fits, we just use `polyfit` to give us the linear portion's two parts. For the line plus sinusoid, we then use trial and error to find reasonable starting values for the amplitude, wavelength, and phase. When we get a reasonably close starting value, the fits always return to those shown in Figure A.3 for even substantial changes in any one of the fit parameters. For reference, we give the seed values for the line plus sinusoid fits. When fitting only to a line, our function is simpler than `LinePlusSinusoid` and we only seed it with the first two parameters:

$$(A_0, B_0, C_0, \lambda_0, \phi_0) = (353, 645, 300, 1.3, 5.0) \tag{A.27}$$

We plot the two ways of fitting this data side by side to compare. Clearly, the addition of the sinusoid makes for a vastly superior fit. In that case, there are 3 degrees of freedom, and the χ^2/ν of 0.31 gives us a degree of confidence of 82%. Perhaps we have slightly overestimated the errors, but this is a very believable value.

For the linear fit, with χ^2/ν of 4.28, the degree of confidence is 0.026%, and a line is extremely unlikely to represent the data fully.

Bibliography

[1] W. K. H. Panofsky and M. Breidenbach. Accelerators and detectors. *Review of Modern Physics*, 71:S121–S132, 1999.

[2] R. M. Barnett et al. Review of particle physics. particle data group. *Physical Review*, D54:128–131, 1996.

[3] J. W. Wang et al. Progress toward NLC/GLC prototype accelerator structures. In *International Linear Accelerator Conference (LINAC 2004)*, 2004.

[4] I. Wilson. The compact linear collider (CLIC). *Phys. Rept.*, 403-404:365–378, 2004.

[5] D. P. Pritzkau. *RF Pulsed Heating*. PhD thesis, Stanford University, 2001.

[6] D. P. Pritzkau and R. H. Siemann. Experimental study of rf pulsed heating on oxygen free electronic copper. *Phys. Rev. ST Accel. Beams*, 5:112002, 2002.

[7] S. S. Manson. *Heat Transfer Symposium*. University of Michigan Engineering Research Institute, 1953.

[8] R. H. Siemann. *Symposium on Electron Linear Accelerators: In Honor of Richard B. Neal's* 80^{th} *Birthday*. Stanford Linear Accelerator Center, 1997.

[9] A. E. Vlieks et al. *Breakdown Phenomena in High Power Klystrons*. SLAC-PUB-4546, Stanford Linear Accelerator Center, 1988.

[10] P. B. Wilson. Scaling linear colliders to 5 TeV and above. In *ITP Conference on Future High Energy Colliders*, 1996.

[11] G. Mourou et al. Laser-induced breakdown by impact ionization in SiO_2 with pulse widths from 7 ns to 150 fs. *Applied Physics Letters*, 64:3071–3073, 1994.

[12] B. C. Stuart et al. Laser-induced damage in dielectrics with nanosecond to subpicosecond pulses. *Physical Review Letters*, 74:2248–2251, 1995.

[13] R. L. Byer et al. Progress of the laser electron accelerator project at stanford university. Presented at IEEE Particle Accelerator Conference (PAC 2001), Chicago, Illinois, 18-22 Jun 2001. Pub. in 'Chicago 2001, Particle accelerator' 108-110.

[14] A. Modena et al. Electron acceleration from the breaking of relativistic plasma waves. *Nature*, 377:606–608, 1995.

[15] Y.B. Fainberg et al. Wakefield excitation in plasma by a train of relativistic electron bunches. *Plasma Physics Reports*, 20:606–612, 1994. and ref. therein.

[16] T. Tajima and J. M. Dawson. Laser electron accelerator. *Physical Review Letters*, 43:267–270, 1979.

[17] K. Nakajima et al. Proof-of-principle experiments of laser wakefield acceleration using a 1 ps 10 TW Nd:glass laser. In *Proc. AIP Conf. Advanced Accelerator Concepts (AAC 1995)*, pages 145–155, 1994.

[18] S. P. D. Mangles et al. Monoenergetic beams of relativistic electrons from intense laser-plasma interactions. *Nature*, 431:535–538, 2004.

[19] C. G. R. Geddes et al. High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding. *Nature*, 431:538–541, 2004.

[20] J. Faure et al. A laser-plasma accelerator producing monoenergetic electron beams. *Nature*, 431:541–544, 2004.

[21] L'OASIS Group, Lawrence Berkeley National Laboratory. Private Communication, 2005.

[22] K. G. Steffen. *High Energy Beam Optics*. John Wiley, New York, 1965.

[23] S. Wang et al. X-ray emission from betatron motion in a plasma wiggler. *Physical Review Letters*, 88(14):135004, 2002.

[24] T. Katsouleas. Plasma wakefields - linear theory. Lecture given at Joint US CERN Japan Russia Accelerator School (JAS 2002). Long Beach, CA, 2002.

[25] J. C. Stewart and K. D. Pyatt. Lowering of ionization potentials in plasmas. *The Astrophysical Journal*, 144:1203–1211, 1966.

[26] P. Debye and E. Hückel. *Z. f. Phys.*, 24:185, 1923.

[27] J. M. Dawson. Nonlinear electron oscillations in a cold plasma. *Physical Review*, 113:383–387, 1959.

[28] E. Esarey et al. Overview of plasma-based accelerator concepts. *IEEE Transactions on Plasma Science*, 24:252–288.

[29] W. K. H. Panofsky and W. Wenzel. Transverse deflection of charged particles in radiofrequency fields. *Review of Scientific Instruments*, 27:967, 1956.

[30] R. Keinigs and M. E. Jones. Two-dimensional dynamics of the plasma wakefield accelerator. *Physics of Fluids*, 30:252–263, 1987.

[31] W. Lu et al., 2004. Private Communication.

[32] T. Katsouleas. Physical mechanisms in the plasma wake-field accelerator. *Physical Review A*, 33:2056–2064, 1986.

[33] T. Katsouleas et al. Beam loading in plasma accelerators. *Part. Accel.*, 22:81–99, 1987.

[34] C. Joshi et al. High-energy density plasma science with an ultrarelativistic electron beam. *Physics of Plasmas*, 9:1845–1855, 2002.

[35] P. Muggli et al. Meter-scale plasma-wakefield accelerator driven by a matched electron beam. *Physical Review Letters*, 93:014802, 2004.

[36] P. Chen et al. Energy transfer in the plasma wakefield accelerator. *Physical Review Letters*, 56:1252–1255, 1986.

[37] R. D. Ruth et al. A plasma wakefield accelerator. *Part. Accel.*, 17:171–189, 1985.

[38] J. Krall and G. Joyce. Transverse equilibrium and stability of the primary beam in the plasma wakefield accelerator. *Physics of Plasmas*, 2:1326–1331, 1995.

[39] D. Whittum et al. Electron-hose instability in the ion-focused regime. *Physical Review Letters*, 67:991–994, 1991.

[40] C. R. Vidal. Heat-pipe oven: A new, well-defined metal vapor device for spectroscopic measurements. *Journal of Applied Physics*, 40(8):3370–3374, 1969.

[41] C. L. O'Connell. *Plasma Production via Field Ionization*. PhD thesis, Stanford University, 2005.

[42] N. B. Delone M. V. Ammosov and V. P. Krainov. Tunnel ionization of complex atoms and of atomic ions in an alternating electromagnetic field. *Sov. Phys. JETP*, 64:1191–1194, 1986.

[43] D. L. Bruhwiler et al. Particle-in-cell simulations of tunneling ionization effects in plasma-based accelerators. *Physics of Plasmas*, 1:2022, 2003.

[44] B. V. Podobedov. *Saw-Tooth Instability Studies at the Stanford Linear Collider Damping Rings*. PhD thesis, Stanford University, 1999.

[45] R. L. Holtzapple. *Longitudinal Dynamics at the Stanford Linear Collider*. PhD thesis, Stanford University, 1996.

[46] M. Sands. The physics of electron storage rings: An introduction. SLAC-0121, Nov. 1970. 172pp.

[47] K. L. F. Bane et al. High-intensity single bunch instability behavior in the new SLC damping ring vacuum chamber. Presented at 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High Energy Accelerators, Dallas, Texas, 1-5 May 1995, p. 3109-3111.

[48] J. Haissinski. Exact longitudinal equilibrium distribution of stored electrons in the presence of self-fields. *Il Nuovo Cimento*, 18B(1):72, 1973.

[49] R. Erickson (Editor). SLC Design Handbook. Stanford, USA: SLAC (1984) Nonconsec. Pag.

[50] K. L. Brown. A second order magnetic optical achromat. SLAC-PUB-2257.

[51] David C. Carey. Why a second order magnetic optical achromat works. *Nucl. Instr. Meth.*, 189:365, 1981.

[52] K. Bane and P. B. Wilson. Longitudinal and transverse wake potentials in SLAC. (talk). In High Energy Accelerators Proceedings, Geneva 1980, 592-596.

[53] V. P. Smirnov V. E. Balakin, A. V. Novokhatsky. VLEPP: Transverse beam dynamics. Proceedings of the 12$^{\text{th}}$ International Conference on High Energy Accelerators. Fermilab: 119, 1983.

[54] J. T. Seeman et al. Measured optimum BNS damping configuration of the SLC linac. Contributed to 1993 Particle Accelerator Conference (PAC 93), Washington, DC, 17-20 May 1993. Published in IEEE PAC 1993:3234-3236.

[55] K. L. F. Bane, A. Mosnier, A. Novokhatsky, and K. Yokoya. Calculations of the short-range longitudinal wakefields in the NLC linac. *eConf*, C980914:137–139, 1998.

[56] E. Keil. *Nucl. Instr. Meth.*, 100:419, 1972.

[57] K. L. F. Bane and P. Emma, 2005. Private Communication.

[58] P. Emma, L. Bentson, and P. Krejcik. A new bunch compressor chicane for the SLAC linac to produce 30-fsec, 30-kA, 30-GeV electron bunches. Contributed to the 8th European Particle Accelerator Conference (EPAC 2002), Paris, France, 3-7 Jun 2002. Published in 'Paris 2002, EPAC 02' 683-685.

[59] P. Emma et al. Measurements of transverse emittance growth due to coherent synchrotron radiation in the SLAC SPPS bunch compressor chicane. Prepared for Particle Accelerator Conference (PAC 03), Portland, Oregon, 12-16 May 2003. Published in IEEE PAC 2003:3129-3131.

[60] V. L. Ginzburg. Transition radiation and transition scattering. Proceedings, 16th International Cosmic Ray Conference, Vol. 14, Kyoto, 1979. Pages 42-50.

[61] X. Artru et al. Experimental investigations on geometrical resolution of electron beam profiles given by OTR in the GeV energy range. Laboratoire de l'Accelerateur Lineaire Report No. 97-30, 1997.

[62] V. L. Ginzburg and V. N. Tsytovich. Several problems of the theory of transition radiation and transition scattering. *Physics Reports*, 49(1):1–89, 1979.

[63] M. Castellano and V. Verzilov. Spatial resolution in optical transition radiation beam diagnostics. *Phys. Rev. Special Topics - Accelerators and Beams*, 1.

[64] A. G. Mozgovoi. The saturated vapour pressure of lithium, sodium, postassium rubidium, and cesium. *High Temperatures - High Pressure*, 19:425–430, 1987.

[65] Allegheny Ludlum Steel Corporation. *Stainless Steel Fabrication*. Allegheny Ludlum Steel Corporation, Pittsburgh 22, PA, 1959.

[66] M. Borland. ELEGANT: A flexible SDDS-compliant code for accelerator simulation. Advanced Photon Source LS-287, September 2000.

[67] P. Muggli et al. Coherent transition radiation to measure the SLAC electron bunch length. In *Proceedings of the 2005 Particle Accelerator Conference (PAC) Knoxville, TN, 16-20 May 2005*.

[68] M. L. Ter-Mikaelian. *High-energy Electromagnetic Processes in Condensed Media*. Wiley-Interscience, New York, 1972.

[69] M. A. Geitz. *Investigation of the Transverse and Longitudinal Beam Parameters at the TESLA Test Facility Linac*. PhD thesis, Universitat Hamburg, 1999.

[70] A. E. Siegman. *Lasers*. University Science Books, Sausalito, 1986.

[71] J. Seeman et al. SLC energy spectrum monitor using synchrotron radiation. Contributed to Linear Accelerator Conf., Stanford, CA, Jun 2-6, 1986. Published in Stanford LINAC86:441.

[72] J. Arthur et al. Linac Coherent Light Source (LCLS) conceptual design report. pages 7–36 to 7–41. SLAC-R-593.

[73] P. Sprangle et al. Laser driven acceleration in vacuum and gases. *AIP Conf. Proc.*, 398:96–105, 1997.

[74] E. R. Colby. *Design, Construction, and Testing of a Radiofrequency Electron Photoinjector for the Next Generation Linear Collider*. PhD thesis, The University of California, Los Angeles, 1997.

[75] R. H. Evans. *The Atomic Nucleus*. McGraw-Hill, New York, 1955.

[76] A. W. Chao (ed.) and M. Tigner (ed.). Handbook of accelerator physics and engineering. pages 215–217. World Scientific, Singapore (1999).

[77] J. D. Jackson. *Classical Electrodynamics, Third Edition*. John Wiley and Sons, New York, 1998.

[78] K. Bane P. Emma. LiTrack: A fast longitudinal phase space tracking code with graphical user interface. In *Proceedings of the 2005 Particle Accelerator Conference (PAC) Knoxville, TN, 16-20 May 2005.*

[79] M. J. Hogan et al. Multi-GeV energy gain in a plasma-wakefield accelerator. *Phys. Rev. Lett.*, 95:054802, 2005.

[80] M. M. Zhou et al. Modeling self-ionized plasma wakefield acceleration for afterburner parameters using QUICKPIC. Submitted to Particle Accelerator Conference (PAC 05), Knoxville, Tennessee, 16-20 May 2005.

[81] J. Orear. Notes on statistics for physiciscts, revised. *Cornell Laboratory for Nuclear Studies*, 82/511, 1982.

[82] W. H. Press et al. *Numerical Recipes: The Art of Scientific Computing (FORTRAN Version)*. Cambridge University Press, Cambridge, 1990.

[83] L. Montanet et al. Review of particle properties. particle data group. *Phys. Rev.*, D50:1173–1823, 1994.