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

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

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 $12\,\mu$m $\approx 40\,$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.

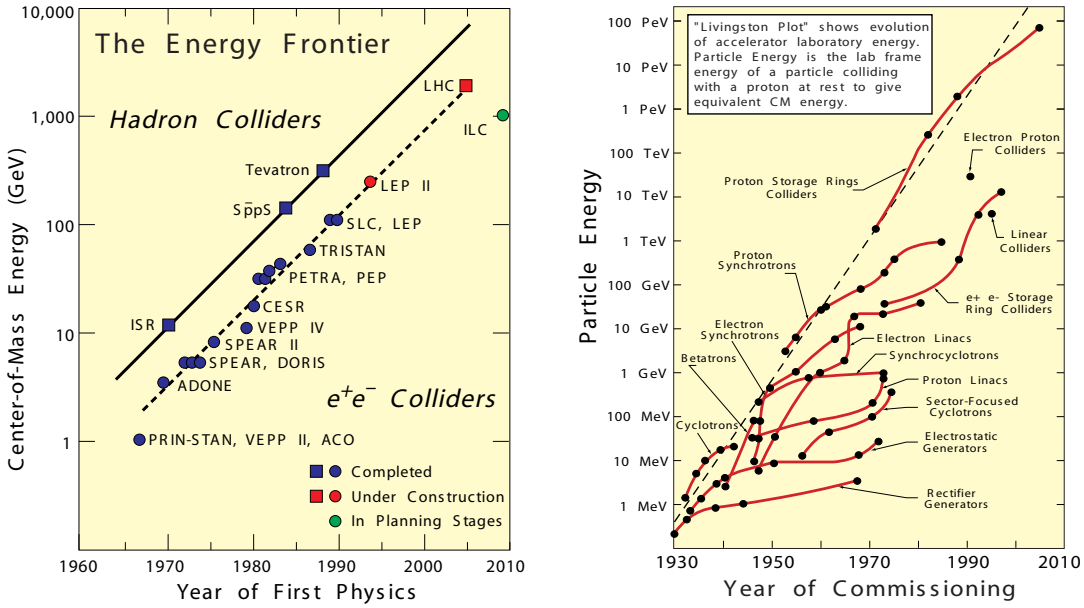

Figure 1.1: Two views of progress in collision energy over time, from [1]

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.

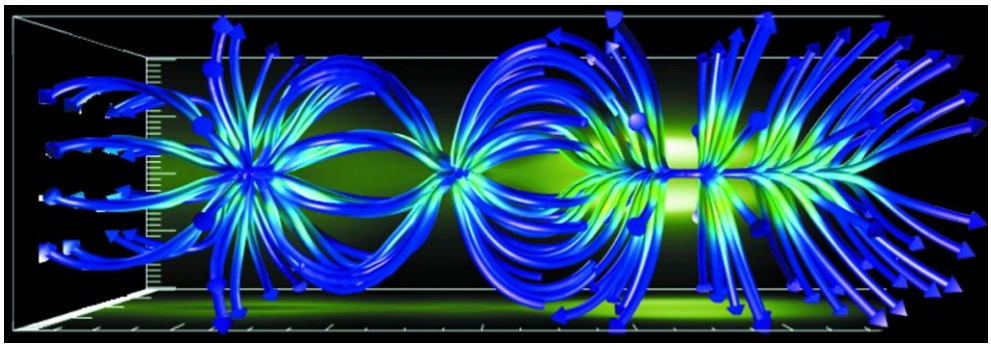

Figure 1.2: Simulation of electric fields due to plasma blowout from an electron beam moving to the right. Acceleration of electrons happens inside the left "bubble" created by the electric field lines. Graphic by Ricardo Fonseca.

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

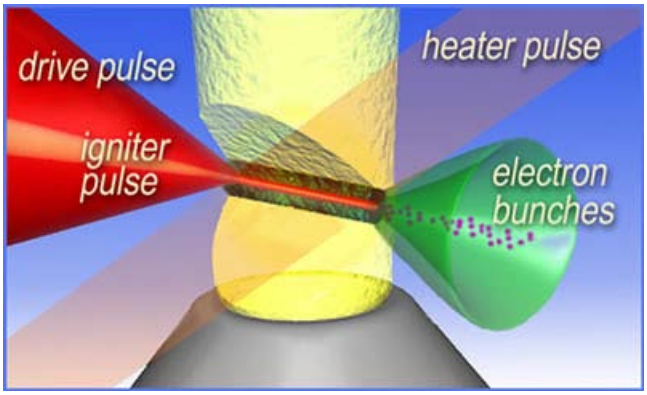

Figure 1.3: Laser acceleration scheme. An igniter pulse forms a "plasma wire" in a hydrogen jet. The heater pulse expands this, making a plasma channel to guide the following drive pulse, which accelerates electron bunches to relatively uniform energy. The green cone is Terahertz radiation from the plasma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

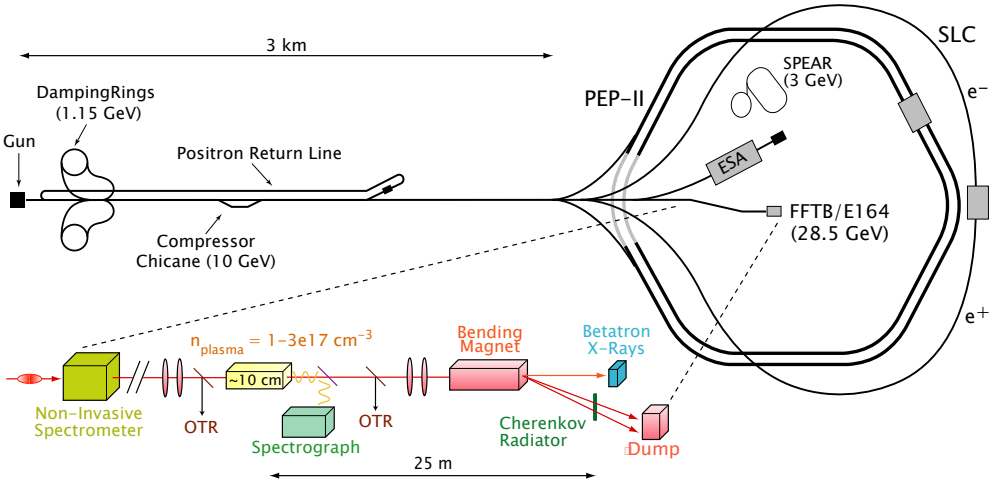

Figure 1.4: The experimental setup for E164

of the accelerator, the bunches can have an RMS length as short as 12 μm, 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 μm, 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,"

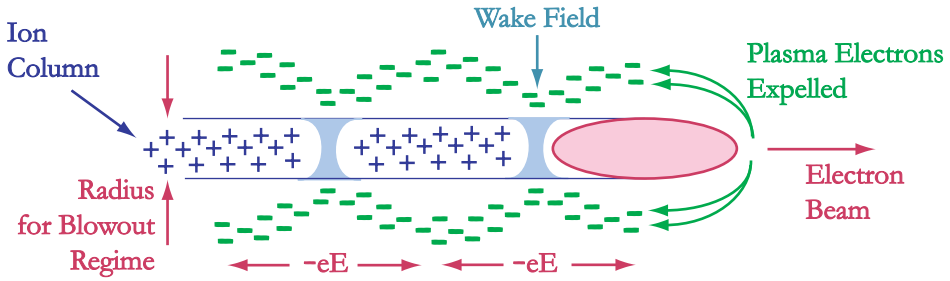

Figure 1.5: Schematic view of beam driven plasma wakefield accelerator.

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.

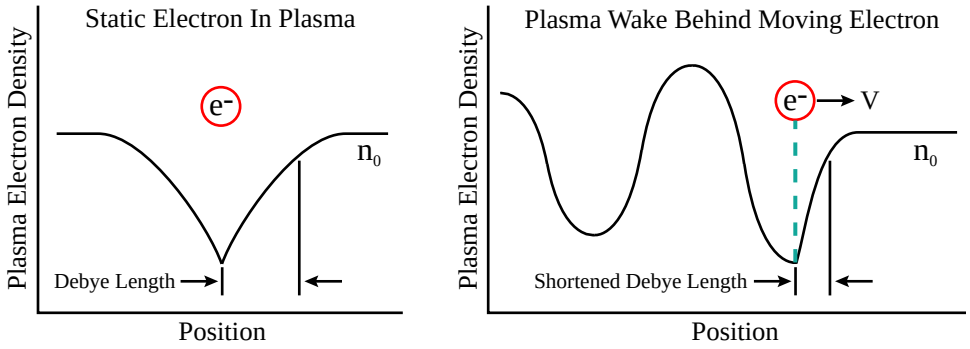

Figure 2.1: Electron Wake Schematic

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:

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

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:

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

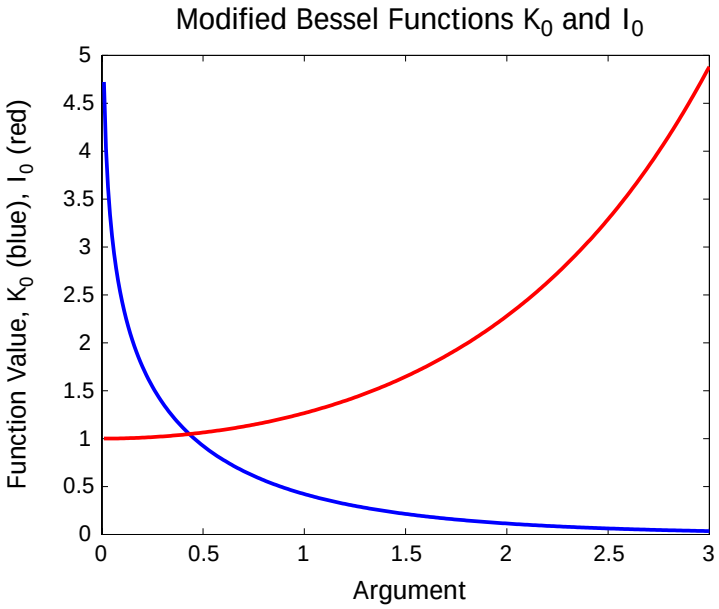

Figure 2.2: Plot of two modified Bessel Functions for arguments 0 to 3.

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 $v = \frac{k_p^2\sigma_z^2}{2}$:

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

We can find the maximum of the function of v: $\partial_v(ve^{-v}) = (1-v)e^{-v} = 0$ 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:

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

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:

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

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.

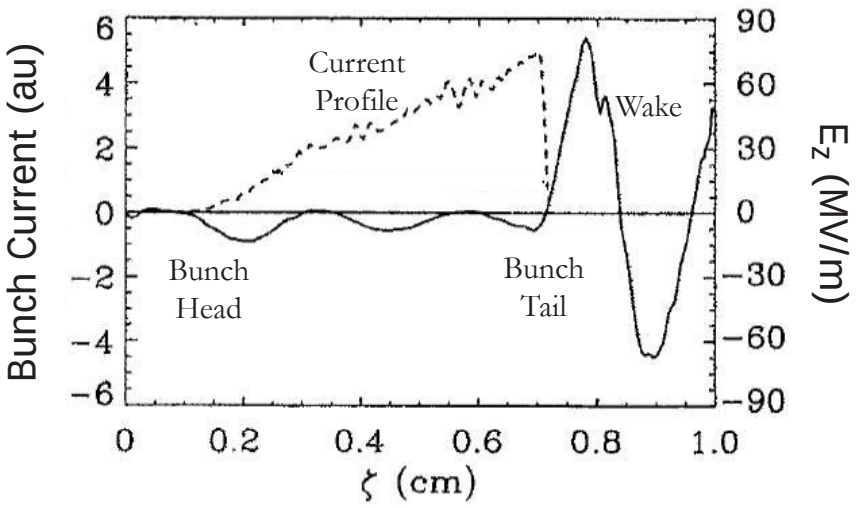

Figure 2.3: Ramped beam shows a strong wake behind. The dashed line is the current profile of the bunch, and the solid line shows the axial wakefield, which develops rapidly once the beam passes. From the simulation presented in [38].

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

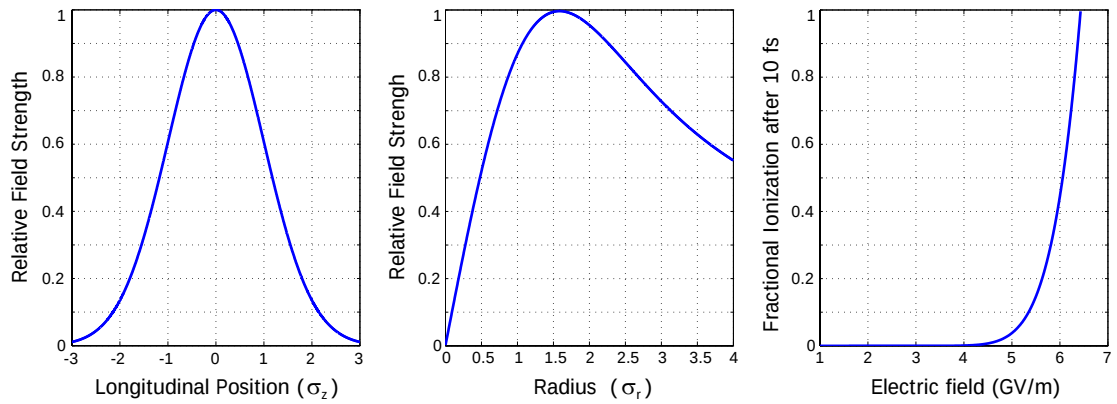

Figure 2.4: Relative field strength in a Gaussian bunch versus z and r. On the right is the fractional ionization of Lithium after 10 fs as a function of the electric field.

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 μm and compress to bunch lengths of about 20 μm, 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

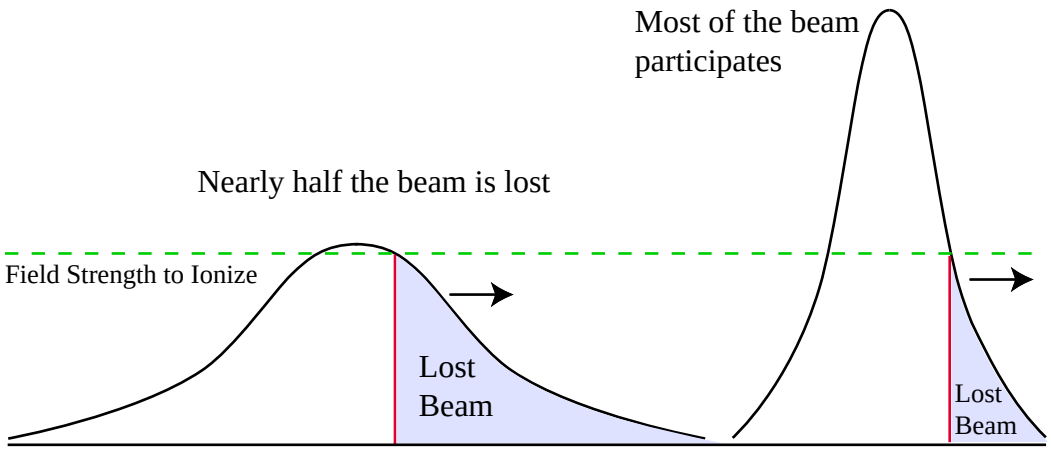

Figure 2.5: Dual Advantages of Longitudinal Beam Compression.

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

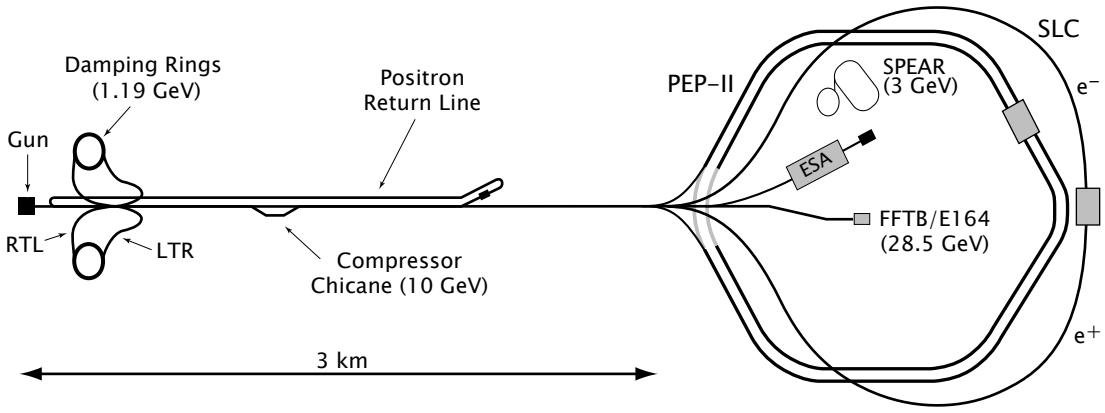

Figure 3.1: Diagram of SLAC with the FFTB facility in which E164 takes place

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

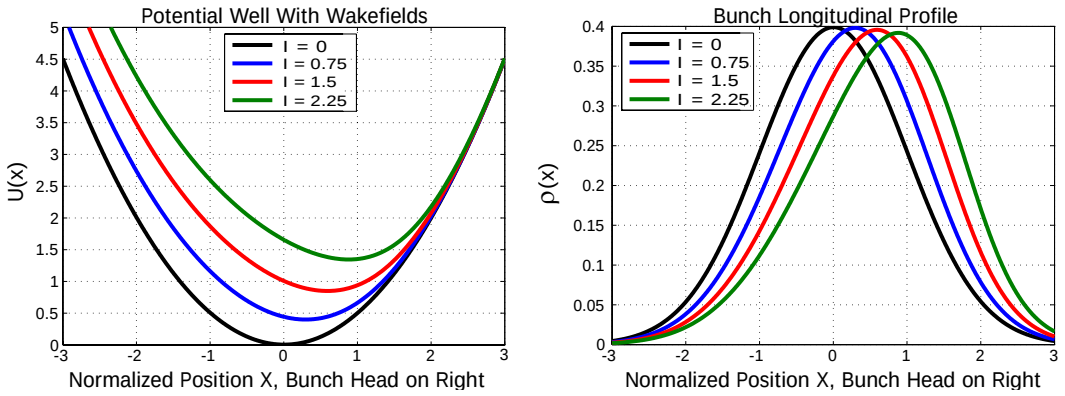

Figure 3.2: The effect of wakefields is to distort the potential well and cause the bunch to tilt toward the head [48].

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

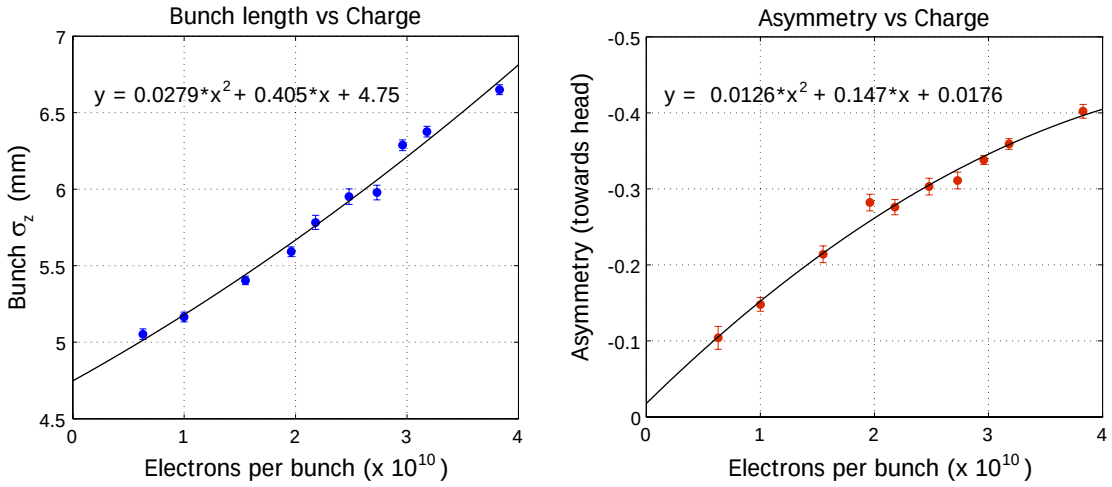

Figure 3.3: Due to wakes, the bunch increases in length and has greater asymmetry toward the head with increasing beam charge. Adapted from [45]

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.

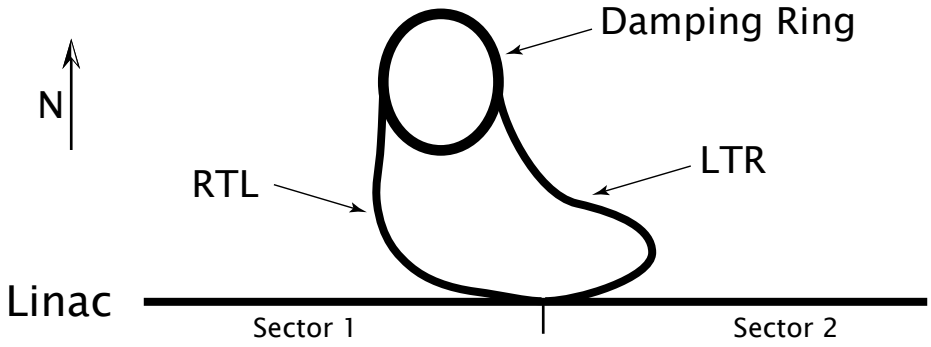

Figure 3.4: Beginning of the linac where the damping ring and RTL are located.

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

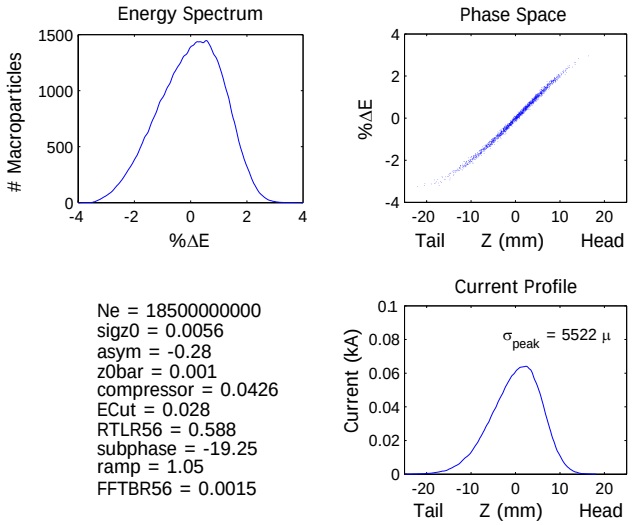

Figure 3.5: Simulation of the beam phase space after the compressor cavity. 100,000 macroparticles represent the beam and they can be histogrammed to provide the energy spectrum and current profile of the bunch.

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 $\mathbf{R}$ 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 $\mathbf{R}$ 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 $\mathbf{T}$, 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

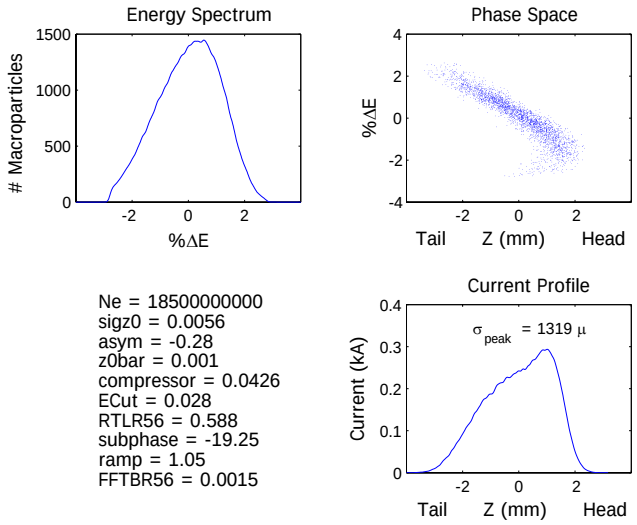

Figure 3.6: Simulated phase space of the beam after compression in the RTL.

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 $R_{56} = -0.588\,\mathrm{m}$ and $T_{566} = -1.054\,\mathrm{m}$. 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 $+20^\circ$ [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

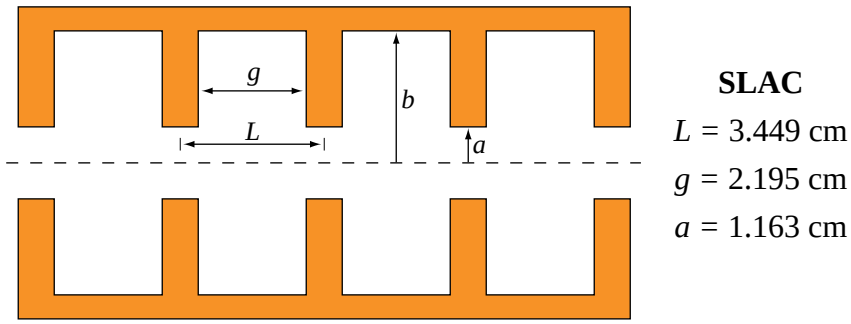

Figure 3.7: Four accelerator cells showing the parameters used to characterize the cell proportions as used in calculating the wake. Our ultrarelativistic electrons never have a chance to see the outer cavity edge, and b does not enter any wake calculations.

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:

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

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.

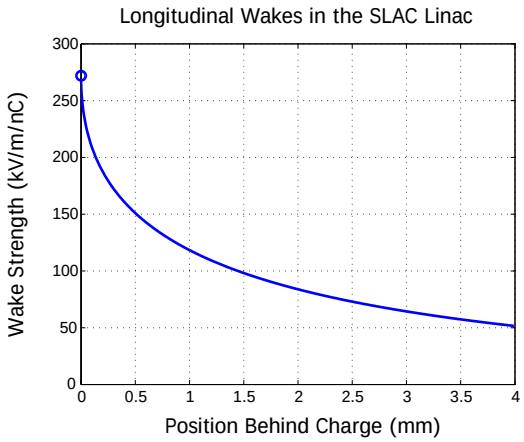

Figure 3.8: Plot of wake strength behind a charge in the SLAC linac. The empty circle shows the maximum field immediately behind the source charge. This calculation [57] was performed in developing the simulation code LiTrack discussed in Chapter 4.

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

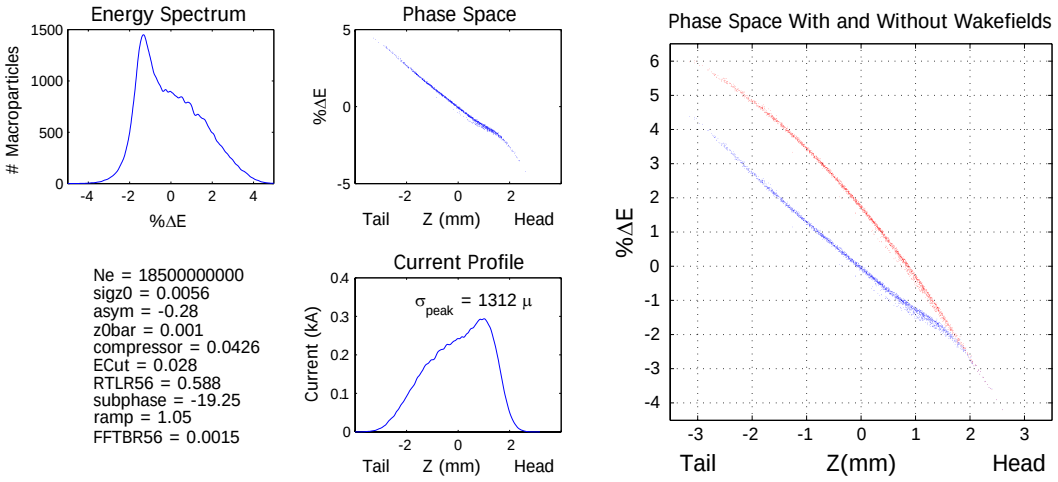

Figure 3.9: Phase space of the beam after traversing the first third of the linac. For comparison, the right hand side shows the beam energy if there were no wakefields in red compared to the actual energy when the wakes are included in blue. The peak loss from wakes is 180 MeV, or about 2%.

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

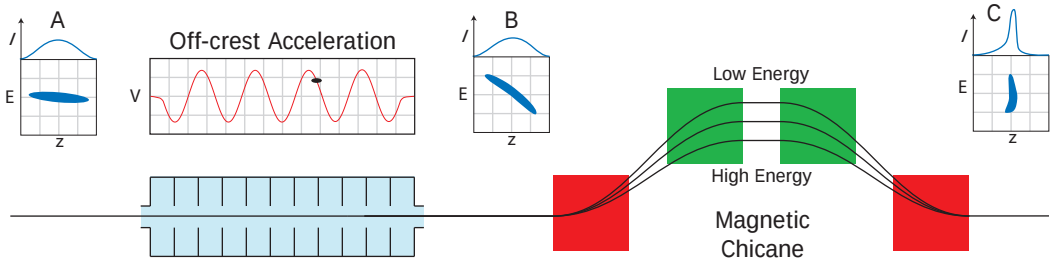

Figure 3.10: Diagram of the Sector 10 Chicane which compresses the bunch down to less than 100 μm using the energy chirp produced in Sectors 2-6.

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

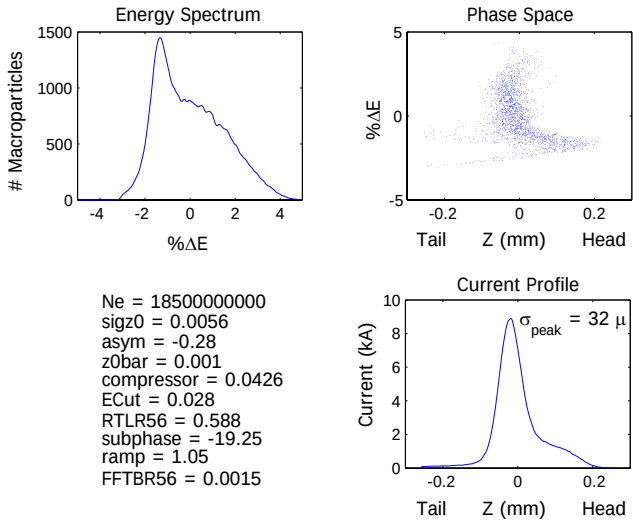

Figure 3.11: Phase space of the beam after the second stage of compression in the Sector 10 chicane.

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.

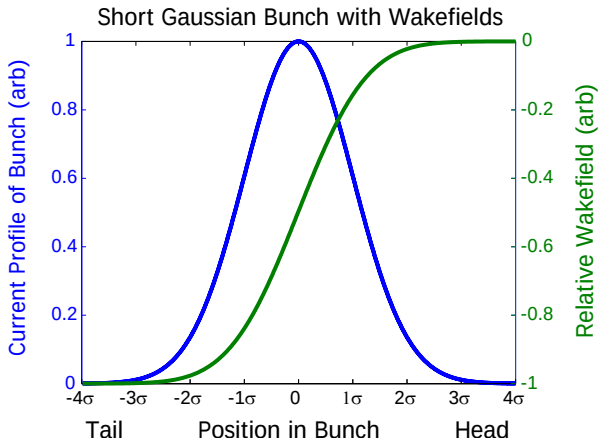

Figure 3.12: The longitudinal wake in a Gaussian bunch which is very short compared to the distance in which the wake decreases appreciably behind a given particle. The wake thus has the shape of the error function, with greatest deceleration in the tail.

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.

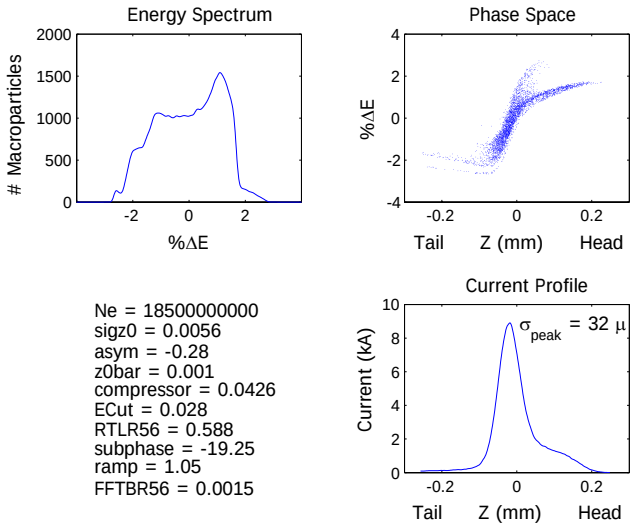

Figure 3.13: Phase space of the beam at the end of the linac.

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

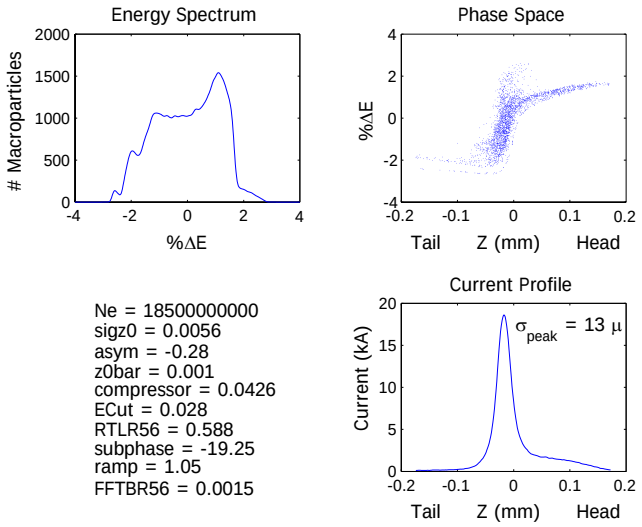

Figure 3.14: Final compressed phase space of beam before the plasma.

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.